

**UCLA**

**UCLA Electronic Theses and Dissertations**

**Title**

Big Data in Hollywood: The New Operating Logic of Media and Entertainment

**Permalink**

<https://escholarship.org/uc/item/6dr680v1>

**Author**

Stork, Matthias

**Publication Date**

2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Big Data in Hollywood:

The New Operating Logic of Media and Entertainment

A dissertation submitted in partial satisfaction of the  
requirements for the degree of Doctor of Philosophy  
in Film and Television

by

Matthias Stork

2020

© Copyright by

Matthias Stork

2020

## ABSTRACT OF THE DISSERTATION

Big Data in Hollywood:  
The New Operating Logic of Media and Entertainment

by

Matthias Stork

Doctor of Philosophy in Film and Television

University of California, Los Angeles, 2020

Professor Stephen Mamber, Chair

This dissertation traces the role of big data in shaping corporate decision-making, institutional practices, and cultural production in Hollywood's legacy media industry over the course of the 21st century. Specifically, the project examines the ways in which the major film studios and television networks integrated data as an industrial operating logic in order to navigate the growing influence of Silicon Valley's technology sector in a shifting digital media environment.

While Hollywood has long incorporated data from audience research to inform the business of cultural production, the dissertation focuses on an emerging set of industrial strategies to manage the exponential increase in digital information. In particular, I consider the industry's use of enterprise application software to *access*, *analyze*, and *apply* data across the

media and entertainment value chain, and the resulting implications for film and television production, marketing, and distribution.

By combining discourse analysis of trade and popular press reports with close readings of select software applications and a mapping of industry dynamics, this project defines the foundational parameters of Hollywood's vertical integration with the data economy. In doing so, I demonstrate that Hollywood gradually engineered a data-driven business model that expanded – and challenged – the operational dynamics of the legacy media business.

The dissertation of Matthias Stork is approved.

John T. Caldwell

Leah A. Lievrouw

Denise R. Mann

Stephen Mamber, Committee Chair

University of California, Los Angeles

2020

## TABLE OF CONTENTS

<i>List of Figures and Tables</i> .....	<i>vii</i>
<i>Acknowledgments</i> .....	<i>ix</i>
<i>Vita</i> .....	<i>xi</i>

<b>Introduction</b>	<b>Big Data in Hollywood</b> .....	1
	Approach and Framework.....	8
	Background and Methodology.....	13
	Overview and Organization .....	18

### Part I – Data Access

<b>Context</b>	<b>Data and Information Building</b> .....	24
	A Framework for Data Access.....	37
<b>Chapter 1</b>	<b>A Century of Access:</b>	
	<b>Building a New Information Infrastructure</b> .....	42
	Pipeline Building: From Data Generation to Fragmentation .....	56
	Data Licensing: Automation and Access .....	81
	Data Acquisitions: From Aggregation to Centralization .....	101
	Direct-to-Consumer: Toward Data Incubation .....	109

### Part II – Data Analysis

<b>Context</b>	<b>Data and Decision-Making</b> .....	123
	A Framework for Data Analysis .....	135
<b>Chapter 2</b>	<b>Analytical Actions:</b>	
	<b>Establishing a Culture of Analytics</b> .....	139
	Data as a Service: The Formation of the Analytical Enterprise. ....	156
	Data as Intelligence: The Rise of Data Metrics .....	186
	Data as Storytelling: The Aesthetics of Data Visualization.....	210

*Part III – Data Application*

<b>Context</b>	<b>Data and Cultural Production</b> .....	230
	A Framework for Data Application .....	245
<b>Chapter 3</b>	<b>The Algorithmic Age:</b>	
	<b>Optimizing the Legacy Media Business</b> .....	249
	Data-Driven Production: Validating Investments.....	266
	Data-Driven Marketing: Programming for the Audience.....	285
	Data-Driven Distribution: Maximizing the Window .....	313
<b>Conclusion</b>	<b>The Iterative Industry</b> .....	332
<b>Bibliography</b> .....		338



## LIST OF FIGURES

Figure 0.1	Research Approach and Framework	14
Figure 0.2	Software Technology Stack Model	16
Figure 1.1	Content Input / Data Output Model	27
Figure 1.2	Platform Ecosystem Model	29
Figure 1.3	Linear Distribution Model	35
Figure 1.4	Digital Distribution Model (Linear View)	36
Figure 1.5	Data Access Industrial Stages Overview	41
Figure 1.6	<i>The Blair Witch Project</i> (1999) Website Landing Page	56
Figure 1.7	Data Taxonomy (Consumer Data)	59
Figure 1.8	Data Taxonomy (Usage Data)	59
Figure 1.9	Digital Landscape Evolution	62
Figure 1.10	Data Ownership Structures	66
Figure 2.1	Legendary Entertainment Box Office Overview	142
Figure 2.2	Data Analysis Process (Information > Insight)	149
Figure 2.3	2012 Movie Year Box Office Overview	157
Figure 2.4	Twitter Platform Growth	159
Figure 2.5	Facebook Platform Growth	168
Figure 2.6	Data Enterprise Overview	177
Figure 2.7	Linear Distribution Channels	203
Figure 2.8	Digital Distribution Channels	204
Figure 2.9	Media Audience Concepts	206
Figure 2.10	Tableau Interface Views	215
Figure 3.1	Box Office Chart (2000-2019)	238
Figure 3.2	Film Admissions (2010-2019)	239
Figure 3.3	Broadcast TV Viewership Trends (2008-2018)	240
Figure 3.4	Pay-TV Subscriber Trends (2012-2018)	240
Figure 3.5	Home Entertainment Sales Trends (2014-2018)	241
Figure 3.6	Relativity Media Box Office Overview	250
Figure 3.7	Data Application Model	253
Figure 3.8	Script Validation Process	278
Figure 3.9	Data Application Tracking System	299
Figure 3.10	Audience Feedback Mechanisms	301
Figure 3.11	Data Application Algorithmic Modelling	302
Figure 3.12	Pay-TV Landscape (2009-2019)	322
Figure 4.1	Legacy Media Streaming Services Overview	334

## LIST OF TABLES

Table 0.1	Data Value Chain	10
Table 1.1	First-Party, Second-Party, Third-Party Data	34
Table 1.2	Data Volume and Data Variety	64
Table 1.3	Platform Types	75
Table 1.4	Platform Data	75
Table 1.5	Platform Data Sharing	78
Table 1.6	Platform Data Applications	90
Table 1.7	Data Startups	93
Table 1.8	Investments in Digital Publishers and Multi-Channel Networks	99
Table 1.9	Data Acquisitions	107
Table 1.10	Data Incubation	115
Table 2.1	The Spectrum of Industrial Decision-Making	125
Table 2.2	Data Analysis Model	128
Table 2.3	Data Analysis Leadership	178
Table 2.4	Legacy Media Measurement Framework	192
Table 2.5	Legacy Media and Digital Media Measurement Frameworks	197
Table 2.6	Platform and Digital Metrics	200
Table 2.7	Legacy Media Metrics and Digital Media Metrics	200
Table 3.1	Cultural Production Value Chain	259
Table 3.2	Production Investments	273
Table 3.3	Demand and Profitability Indicators	274
Table 3.4	Data Application Marketing Investments	308
Table 3.5	Data Application in Distribution	325
Table 4.1	The Data Value Chain in Streaming	335

## ACKNOWLEDGMENTS

I have been fortunate to receive advice and support from a remarkable group of people.

First and foremost, I would like to express my utmost gratitude to my committee chair, Steve Mamber. Steve is a distinguished scholar and brilliant teacher whose work has served as a constant inspiration to me. I was lucky enough to take a series of classes with Steve – covering issues in electronic culture, video games, and computer applications – which would form the foundation of my ongoing research. With the dissertation, he encouraged me to chart my own path from the start, offering ongoing feedback, insight, and reassurance through engaging discussions along the way. He was always present, open to all questions and ideas, and unwavering in his support. Without Steve, this project would not have been possible. I will always be grateful for his perspicacity, empathy, and friendship.

I am further indebted to my committee members and their invaluable support. Denise Mann has been an incredible mentor to me, ever thoughtful and inspirational. I remain in awe of her encyclopedic knowledge, keen wit, and uplifting spirit. John Caldwell is a true renaissance man and I am thankful for his wisdom and compassionate perspective throughout this project. Leah Lievrouw's unique scholarship and expertise provided substantial guidance, and her work ethic was nothing but inspiring. To my entire committee – *dankeschön*.

I also would like to thank the Cinema and Media Studies faculty for their support and guidance: Chon Noriega, Vivian Sobchack, Kathleen McHugh, Jasmine Trice, Janet Bergstrom, William McDonald, Maja Manojlovic, Vincent Brook, and Peter Lunenfeld. I owe tremendous gratitude to Barbara Dube and Barrett Korrerat for their immeasurable administrative support. And to Brian Clark – your help meant so much in so many ways, thank you.

I am further tremendously grateful for the fellowships and grants I received from the UCLA Graduate Division and the Department of Film, Television and Digital Media.

To all my fellow students and cohort members – thank you for your kindness and inspiration. Special thanks to James Fleury for everything.

To Thomas – merci, mon ami.

To James Gilmore, Josh Olejarz, and Kelly Lake – thank you for giving me a home away from home. My home is your home.

And, to my family, Mama, Papa, Katja, Paula, and Nelly – thank you for always being there for me. *In Gedanken an* Peter and Wolfgang. To my American family, Mari, Chris, and Chandler – thank you so much for all you do. And a special thank you to Ren.

Finally, to my wife, Hadley – this is dedicated to you. You make everything better.

## VITA

### Education

- 2015 C.Phil., Film and Television  
Department of Film, Television and Digital Media  
University of California, Los Angeles
- 2013 M.A., Film and Television  
Department of Film, Television and Digital Media  
University of California, Los Angeles
- 2010 B.A., English and French Foreign Language Education (*Staatsexamen*)  
Faculty of Educational Sciences and Modern Languages  
Goethe University Frankfurt, Germany

### Awards

- 2013–17 University Fellowship  
University of California, Los Angeles

### Select Research Experience

- 2011–16 Research Assistant  
Department of Film, Television and Digital Media  
University of California, Los Angeles
- 2014–15 Research Coordinator  
Transforming Hollywood Conference  
University of California, Los Angeles

### Select Publications

James N. Gilmore & Matthias Stork, eds., *Superhero Synergies: Comic Book Characters Go Digital* (Lanham, MD: Rowman & Littlefield, 2014)

## Introduction

### Big Data in Hollywood

This dissertation project traces the emergence of big data<sup>1</sup> as a media industry logic that increasingly informed corporate decision-making, institutional practices, and cultural production in Hollywood over the 21st century. Specifically, I examine how the major film studios and television networks integrated data as an industrial operating logic in order to navigate a shifting media environment.<sup>2</sup> To this end, rather than offering a historical trajectory, the study engineers a conceptual map to understand the impact of data on the way the legacy media industry has come to operate in the digital era, economically, culturally, and technologically.<sup>3</sup>

On the surface, the concept of data refers to a set of technological practices that have facilitated the automation and management of an exponential increase in digital information.<sup>4</sup> In particular, it delineates a complex digital infrastructure that consistently translates online activity into quantifiable information, thereby enabling a growing output of data.<sup>5</sup> The resulting industrial

---

<sup>1</sup> For the main body of this dissertation, I will rhetorically modify the term *big data* as simply *data* to align with conventional media industry usage.

<sup>2</sup> I define Hollywood's legacy media industry and the majors broadly as a group of leading film studios (e.g., Disney, Warner Bros., Universal, Sony, Paramount, Fox, Lionsgate), broadcast television networks (e.g., ABC, NBC, CBS, Fox), and premium cable channels (e.g., HBO, Showtime, Starz), situated within larger complex and shifting conglomerate structures, with a focus on the U.S. market (thus excluding international developments).

<sup>3</sup> The rationale of this project is not to historicize the present, but to offer conceptual ideas and frameworks to understand the role and impact of data in the media and entertainment industry from the late 1990s to the 2010s.

<sup>4</sup> I use digital information to refer to data that specifically relates to media and entertainment. The digital information increase is generally summarized under the 3 Vs: volume, variety, velocity. See, Lycett, Mark, "Datafication: making sense of (big) data in a complex world," *European Journal of Information Systems*, 22, 2013.

<sup>5</sup> This generative data infrastructure setup is generally described as the process of datafication. See, Viktor Mayer-Schoenberger & Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (New York: Eamon Dolan/Houghton Mifflin Harcourt, 2013).

data-base has been amplified by the ongoing commercialization of the Internet, grounded in a key set of underlying developments:

- *Always-On Connectivity*: The wholesale shift from dial-up to broadband access in the early 2000s and the ongoing development of mobile technologies effectively mainstreamed Internet availability across the U.S., enabling a consistent rise of online users and a concurrent increase in data output.<sup>6</sup>
- *Web 2.0*: The millennial shift from a static, non-interactive online environment to a digital infrastructure of interactive media effectively laid a foundation for increased online activity, with users moving from acting as consumers to creators of digital content, which further stimulated data output.<sup>7</sup>
- *Cloud Computing*: The mainstream shift from a localized data center infrastructure to on-demand cloud storage capabilities in the mid-2000s yielded a substantial increase in data output, driven further by the rise of new commercial business models that enabled the capture, consolidation, and commercialization of data across enterprise organizations.<sup>8</sup>

The confluence of technological advancements, economic considerations, and cultural practices effectively created a state of *data deluge*,<sup>9</sup> an industrial complex increasingly defined, driven,

---

<sup>6</sup> See, McCullough, Brian, *How the Internet Happened: From Netscape to the iPhone* (New York: Liveright, 2018).

<sup>7</sup> See, Bruns, Axel, *Blogs, Wikipedia, Second Life, and Beyond: From Production to Producership* (Frankfurt am Main: Peter Lang, 2008).

<sup>8</sup> See, Ruparelia, Nayan B., *Cloud Computing* (Cambridge: MIT Press, 2016).

and determined by digital information.<sup>10</sup> In this environment, data has come to operate as an industry driver, impacting the dynamics of corporate organizations, including their institutional thinking, corporate actions, and the associated commercial value of products and services.<sup>11</sup>

Accordingly, this project posits data as an industrial logic, marked by a multi-dimensional system of economic, cultural-organizational, and technological practices that increasingly inform operational industry norms in Hollywood. Rather than positioning the legacy media industry as a passive and static entity, subject to the innovations and disruptions brought about by the data industrial complex,<sup>12</sup> I examine a set of integration strategies the majors employed to leverage data as an industrial driver, differentiator, and determinant in a shifting digital media environment. Specifically, I explore the ways Hollywood engaged a network of enterprise application software solutions to *access*, *analyze*, and *apply* data across the media and entertainment value chain, informing the industry's approach to corporate decision-making and institutional practice in cultural production. Indeed, I argue that the legacy media industry negotiated a complex and layered software technology stack to integrate data, from cultivating

---

<sup>9</sup> Anderson, Chris, "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete," *Wired*, June 28, 2008, accessed June 6, 2020, <https://www.wired.com/2008/06/pb-theory/>

<sup>10</sup> See, *The Economist*, "The world's most valuable resource is no longer oil, but data," May 6, 2017, accessed June 6, 2020, <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

<sup>11</sup> For an analysis of data as an industry driver in Silicon Valley, see Wu, Tim, *The Attention Merchants: The Epic Scramble to Get Inside Our Heads* (New York: Vintage, 2017). For data's impact on socio-cultural and economic dynamics, see, Lohr, Steve, "The Age of Big Data," *New York Times*, February 11, 2012, accessed June 6, 2020, <https://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>

<sup>12</sup> See, Bilton, Nick, "Why Hollywood As We Know It Is Already Over," *Vanity Fair*, January 29, 2017, accessed June 6, 2020, <https://www.vanityfair.com/news/2017/01/why-hollywood-as-we-know-it-is-already-over>



partnerships with a range of technology companies to incubating internal product solutions and managing the organizational implications of technology.<sup>13</sup>

As such, Hollywood effectively engineered a new operating logic, organized around the access, analysis, and application of data through the holistic implementation of enterprise software solutions. To this end, the project tracks three overarching industrial shifts in how the legacy media industry has come to operate under the increasing impact of data.

First, I examine the legacy media industry's shift from *impressionistic to analytical modes of decision-making* as part of a larger, institutional practice of data-driven rationalization.<sup>14</sup> While Hollywood has long cultivated a creative operating model organized around instinct, intuition, and gut logic,<sup>15</sup> the integration of data at once complicated and challenged this approach by increasingly enabling analytical, measured, and rationalized industrial practices.

Second, I explore the industry's shift from *lacking to actively managing and controlling data* in a new media environment increasingly defined by alternative modes of production, competitive forces, and partnership models. A growing body of scholarship and commentary has positioned Hollywood's creative business in opposition to the data-driven logics of Silicon Valley,<sup>16</sup> positing a dynamic wherein the major studios and networks consistently play catch-up

---

<sup>13</sup> I use the concept of the software stack as a reference to characterize Hollywood's use of software to access, analyze, and apply data, with the majors effectively building a modular toolbox of capabilities and partnerships. For more information on the software stack, see, Bratton, Benjamin H., *The Stack* (Cambridge: MIT Press, 2016).

<sup>14</sup> See, Napoli, Philip M., "On Automation in Media Industries: Integrating Algorithmic Media Production into Media Industries Scholarship," *Media Industries Journal*, Volume 1, Issue 1, 2014.

<sup>15</sup> See, Michael D. Smith & Rahul Telang, *Streaming, Sharing, Stealing: Big Data And The Future Of Entertainment* (Cambridge: MIT Press, 2016).

to technologically-advanced data companies. This study acknowledges and analyzes differences between the two industrial sectors while employing a model of complex convergence that sees the legacy media industry actively work to adopt, adapt, and advance data-driven practices from the technology industry.<sup>17</sup>

And, third, I analyze the industry's overarching shift toward a new operating logic, from *pushing out content to pulling in data*, effectively leveraging data as an industrial driver of cultural production. As such, the integration of data enabled alternative and advanced modes of production, marketing, and distribution, thereby directly impacting the way the industry operates. While Hollywood has long incorporated audience research to inform industrial output,<sup>18</sup> “combining aggregate statistics [...] with data from small samples [...] and gut feel [...]”<sup>19</sup> the use of data introduced a new set of practices, adding unprecedented information volume, variety, and velocity, that require the deployment of new economic models, organizational dynamics, and technologies. Over the course of the early 21st century, the majors have increasingly integrated data holistically, investing in modes of automation to manage the legacy media business.

Hollywood's integration of data represents a complex, complicated, and non-linear process, marked by a suite of industrial experiments with enduring implications and ephemeral iterations. By emphasizing a conceptual rather than historical narrative, my goal is to construct a

---

<sup>16</sup> See, Lotz, Amanda D., *Portals: A Treatise on Internet-Distributed Television* (Ann Arbor: Michigan Publishing Services, 2017) and Johnson, Derek, ed., *From Networks to Netflix: A Guide to Changing Channels* (New York: Routledge, 2018).

<sup>17</sup> While convergence is often posited as an organic coming together of industries (see, Jenkins, Henry, *Convergence Culture* (New York, New York University Press, 2008)), I approach it as a complex integration of industrial forces at the economic, cultural-organizational, and technological level. Indeed, I argue that the legacy media industry had to actively negotiate the dynamics of the technology industry in order to integrate data.

<sup>18</sup> See, Wasko, Janet, *How Hollywood Works* (New York: Sage, 2003).

<sup>19</sup> Smith & Telang, 2016

framework that accounts for the legacy media industry's ongoing integration of data, outlining distinct continuities and change within the evolving industrial model, as opposed to marking a specific end state. As such, the project aims to highlight viable models of integration while simultaneously emphasizing notable challenges and fissures associated with industrial shifts.<sup>20</sup> To this end, I build on a dynamic catalog of research and scholarship chronicling the impact of macro-structural factors on the logics of Hollywood writ large, providing a high-level analytical view of industrial integration.<sup>21</sup> At the same time, I consider structural developments at the economic, cultural-organizational, and technological level, detailing the impact of data on corporate actions and institutional practices that define industrial output at film studios and television networks, as well as their overarching conglomerate structures. Finally, I work toward providing a differentiation from existing research by prioritizing the legacy media industry perspective. Existing work on the intersection of data and Hollywood has largely framed the subject through an analysis of the production cultures, organizational dynamics, and digital aesthetics of platforms, with a focus on streaming providers.<sup>22</sup> This project, meanwhile, aims to balance the scholarly narrative by focusing on the perspective, position, and performance of the old guard, the major film studios and television networks, effectively reverse-engineering the operational data dynamics of the legacy media industry to add to the scholarly discourse.

Given this focus, there are a number of considerations that lie outside the scope of this project. In particular, while I look at the impact of data on established decision-making practices

---

<sup>20</sup> On the value of studying industrial fissures and challenges, see, Nora, Draper, "Fail Fast: The Value of Studying Unsuccessful Technology Companies," *Media Industries Journal*, Volume 4, Issue 1, 2017.

<sup>21</sup> See, Timothy Havens & Amanda Lotz, *Understanding Media Industries* (Oxford: Oxford University Press, 2016).

<sup>22</sup> Netflix, in particular, has emerged as an emblem of this approach (see, Johnson, 2018 and Napoli, 2014). I equally draw on Netflix for contextual insight, yet further consider the role of the entire FAANG ecosystem.

in Hollywood, the goal is not to measure, quantify, or evaluate the resulting business implications for the legacy media industry (as such, it is not an economics study). Additionally, the purpose of this project is not to serve as a manual on how to optimize media industrial practice with data (it is not a business study or how-to guide). Finally, while I focus on software as a central component of modern media industrial dynamics, the project is not meant to examine the viability of software systems at work (it is not a computer science study).

In tracing Hollywood's integration of data as an industrial operating logic over the course of the 21st century, I set out to answer the following key questions:

- How has the legacy media industry negotiated the implications of data as a central operating logic for established modes of corporate decision-making, institutional practice, and cultural production?
- What industrial strategies have the major film studios and television networks employed to make data work within their legacy media organizations, and the industry as a whole?
- How has the integration of data impacted the industrial infrastructure of the legacy media industry, and what were the implications for the majors at an economic, cultural, and technological level?

Given the rapid evolution of digital technology, the answers to these questions are not absolute, but manifest across a dynamic spectrum of frameworks. My overarching goal is to offer a working conceptual toolkit to account for the implications of Hollywood's emerging data logic.

## Approach and Framework

To examine data as a complex industrial logic, I propose a three-dimensional framework that considers the economic, cultural, and technological affordances of data in the context of the media and entertainment industry. The framework maps these affordances against a core set of ideas that constitute the underlying foundation of this project on industrial integration.

First, I consider the idea of *data as capital*. Data is widely seen as the most valuable resource of the digital economy in that it provides organizations with unprecedented amounts of “user and profiling data,”<sup>23</sup> offering detailed information on online consumer actions, habits, and preferences, thereby enabling a customer-centric business approach with customized product output. Indeed, data companies have increasingly developed organizational structures and technological capabilities to manage a data-driven business model, effectively pulling in data to inform their business strategy.<sup>24</sup> While the legacy media industry has historically focused on pushing out content to maximize audience, market, and revenue share,<sup>25</sup> the rise of data as capital has shifted focus to the underlying information clusters inherent in established operational modes of cultural production. To this end, the project examines the process by which the major film studios and television networks have sought to unlock the capital of data.

Second, I consider the idea of *data as process*. Building on the notion of data as a form of economic, cultural, and technological capital, I argue that data effectively operates as an industrial process rather than an information object. Specifically, I examine how the major film

---

<sup>23</sup> Van Dijck, Jose, *The Culture of Connectivity: A Critical History of Social Media* (Oxford: Oxford University Press, 2013). Jose Van Dijck provides a critical perspective on the way platforms capture and commercialize data.

<sup>24</sup> See, Steven Rosenbush and Michael Totty, “How Big Data Is Changing the Whole Equation for Business,” *Wall Street Journal*, March 10, 2013, accessed June 6, 2020, <https://www.wsj.com/articles/SB10001424127887324178904578340071261396666>

<sup>25</sup> See, Wasko, 2003

studios and television networks have worked to access, analyze, and apply data to realize it as capital. This three-part structural approach is modelled on the actual data value chain of the digital economy and adapted from the operational dynamics of digital platforms (table 0.1). Mainstream platforms like Netflix, Amazon, and Google *access* data on millions of users through a proprietary technological infrastructure; *analyze* data to deliver strategic insights that drive business decisions; and *apply* the resulting insights to execute business decisions.<sup>26</sup> As such, the entire platform value chain is driven by the logic of data, with platforms actively pulling in, processing, and productizing data to manage their business. To this end, I posit that the majors vertically-integrated with the data value chain of the digital economy, adopting and adapting the processes of data access, analysis, and application to manage the legacy media business in the digital era, encountering and engaging various challenges along the way. The consideration of platforms in this framework adds yet another layer of complexity in that several platforms emerged as active players in the media and entertainment business in the 21<sup>st</sup> century, serving as distribution partners and programming competitors to the majors.<sup>27</sup>

---

<sup>26</sup> It is crucial to note that the three procedural stages are neither sequential nor linear, but occur in a complex dynamic that unfolds across multiple iterations.

<sup>27</sup> Facebook, Amazon, Apple, Netflix, and YouTube licensed programming from the majors, while later creating original programming to compete with them. See, Lev-Ram, Michal, “How Netflix Became Hollywood’s Frenemy,” *Fortune*, June 7, 2016, accessed June 6, 2020, <https://fortune.com/longform/netflix-versus-hollywood/>

**Table 0.1**

The Data Value Chain			
Stage	Access	Analysis	Application
<b>Process</b>	Monitor, collect, and organize data on user information and behavior	Turn data into insights that evaluate and inform business decisions	Apply the data-driven insights to execute a business decision
<b>Context</b>	The collection of data to make it available, usable, and actionable for analysis	The analytical interpretation of data to create business insights	The encoding/execution of data insights through specific algorithms

And, third, I consider the idea of *data as logic*. Hollywood’s integration of data as an industrial logic had several strategic and logistical implications on how the major film studios and television networks have come to operate, impacting corporate decision-making, institutional practice, and the dynamics of cultural production. As such, the data logic manifested across a holistic spectrum of industry layers, from top-down executive actions to bottom-up patterns of the larger organizational workforces. Effectively, I argue that data logic became media industry logic, increasingly shaping the legacy media industry’s economic model, cultural-organizational structures, and technological infrastructure. The majors did not undergo a wholesale transformation into platforms. Rather, they negotiated a dynamic industrial model that integrates the operational features of legacy media companies with select platform characteristics, yielding an increasingly data-driven media industry.

This framework further incorporates a set of hypotheses that dimensionalize the prevailing concept of data as digital information. Previous media studies research has generally framed data as “units or morsels of information [...]“ that form part of a larger industrial

context.<sup>28</sup> Expanding this overarching concept, I consider underlying structural properties and features of data that consistently feed into the contextual industry dynamics.

- *Data is multi-layered*: Data is generally presented as a mass of digital information. Yet, the emphasis on scale tends to obscure the underlying layers of information, which span a multitude of categories. Data can be classified by channel (e.g., web data, search data, social data, mobile data, video data), platform (e.g., Facebook data, Amazon data, Netflix data, Google data), or device (e.g., smartphone data, tablet data, streaming player data), among others. For this project, I will construct a working taxonomy of data, largely organized around the dominance of platforms in the digital economy, while primarily distinguishing between two core categories of data: *consumer* data (i.e., audience information, such as demographics and psychographics) and *usage* data (i.e., behavioral audience information, such as video views, likes, online comments, shares, clicks).
- *Data is messy*: Data is not a holistic centralized entity, but a wide range of information, distributed, fragmented, and decentralized across various digital channels, platforms, and devices. For this project, I primarily categorize data by channel type, from website data (i.e., derived from web portals) to search data (i.e., search engines), social data (i.e., social media), mobile data (i.e., mobile apps and games), and video/OTT data (i.e., streaming), while equally considering the role platforms play in the surrounding power dynamics of the digital economy. Based on this classification, I argue that platforms dictate certain economic, cultural, and technological affordances of data, from ownership

---

<sup>28</sup> See, Gitelman, Lisa, ed., *“Raw Data” Is An Oxymoron* (Cambridge: MIT Press, 2013).



structures to the economic model and associated capital, which can yield highly complex, complicated, and messy industrial connections.

- *Data is curated*: Platforms consistently run enterprise software to code and curate data, offering varying views of information that are at once public and proprietary, there-to-see and hidden from view. To this end, I differentiate between a user-centric and machine-centric view of data.<sup>29</sup> The user-centric view (i.e., the *frontend*) curates information for a specific audience. For example, the *like button* has become a visual reference point to express user interaction with a post across platforms, offering a consolidated and structured view of publicly available information. Almost any piece of digital content features a like button that showcases the number of interactions the piece has generated (i.e., the frontend data). Meanwhile, the machine-centric view (i.e., the *backend*) contains a much broader data set, unstructured and expansive, hidden from public view, which holds incremental details to contextualize the user-centric view (i.e., the demographics, geographic location, and time stamps associated with the users who liked a post or the underlying metadata associated with users' related online accounts).<sup>30</sup>
- *Data is automated*: Data is continuously accessed, analyzed, and applied by software applications across the industrial value chain. The concept of software-driven automation is not a self-contained phenomenon, but rather a combination of technological resources and cultural-organizational practices, driven by economic considerations. For this project,

---

<sup>29</sup> See, Manovich, Lev, *Software Takes Command* (London: Bloomsbury, 2013).

<sup>30</sup> The idea of curation builds on Lisa Gitelman's notion that data is never raw, but invariably informed by the platform infrastructure where it generates from. See, Gitelman, 2013

I specifically focus on modes of automation that visually output data in the industrial context. To this end, I place emphasis on software interfaces that surface, structure, and streamline data.<sup>31</sup>

- *Data is negotiated*: Data is an expansive information complex that is actively produced, pursued, and processed by a multitude of agents following economic, cultural, and technological considerations. In this project, I make the argument that the legacy media industry has actively negotiated data access, analysis, and application through the use of software, engagement with various technological and strategic stakeholders, and management of corresponding industrial dynamics.

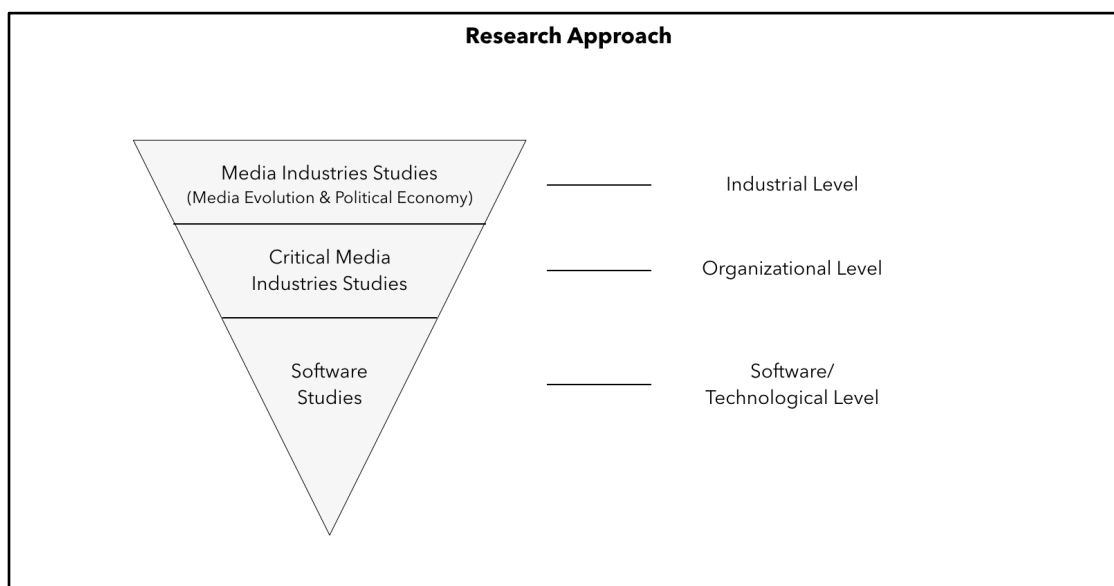
Collectively, this feature set aims to underline that data is invariably an industrial construct that depends on various internal and external factors. As such, I offer the argument that data is not a unique, infallible industrial trick that elevates business performance by default, but merely another logic at play in Hollywood's approach to managing the legacy media business.

## **Background and Methodology**

The project takes an interdisciplinary research approach that synthesizes media industries studies with select principles of software studies. The overarching goal is to critically examine the industrial practices that enable Hollywood's integration of data while, at the same time, analyzing the underlying technological processes that encode and enact it through new economic and cultural-organizational models. I propose a three-part model in which the industry negotiates a confluence of internal and external forces across several layers (figure 0.1).

---

<sup>31</sup> See, Manovich, Lev, *The Language of New Media* (Boston: MIT Press, 2001) and Bolter, Jay David & Grusin, Richard, *Remediation: Understanding New Media* (Cambridge: MIT Press, 2000).



**Figure 0.1:** The project synthesizes media industries and software studies to examine multiple layers of industrial integration through a politic-economic, cultural-organizational, and technological lens.

In media industries studies, I specifically draw on ideas from media evolution and, more broadly, political economy. As such, I focus my analysis on how the legacy media industry adapted to shifting environmental circumstances.<sup>32</sup> In particular, I place emphasis on the ways the majors have adapted existing legacy media processes to the economic, cultural, and technological affordances of data. To this end, the project offers a narrative of continuity and change, tracing the industry’s forward propulsion to experiment with new data practices against a set of established modes of operation. I further employ a political-economic lens to highlight the dynamics between Hollywood and Silicon Valley or, more specifically, the relationship between legacy media organizations and a suite of technology companies, from digital platforms to upstarts, startups, and third-party intermediaries. My goal is to sketch an industry portrait that is

---

<sup>32</sup> See, Napoli, Philip M., *Audience Evolution: New Technologies and the Transformation of Media Audiences* (New York: Columbia University Press, 2010).

inherently rhizomatic, with the legacy media industry forging opportunistic partnerships and connections that are highly dependent upon larger industrial circumstances.<sup>33</sup>

Additionally, the project adds a micro-layer of critical media industries studies to highlight the impact of data at the organizational level and thereby balance the broader focus on large-scale industrial developments.<sup>34</sup> In particular, I examine the emerging *data centers* within film studios and television networks, a growing set of divisions, departments, and deputies tasked with leading the internal processes and practices to integrate the access, analysis, and application of data within rigidly existing and fluidly expanding corporate infrastructures. Rather than providing a small set of comprehensive case studies from select legacy media organizations, I focus on a multitude of interrelated use cases across the entire legacy media industry, constructing a network of small-scale examples designed to reflect, reverse-engineer, and, at times, recast the broader macro-structural developments at play in the industry. In this regard, the project aims to provide a full-stack analysis of data as a media industry logic, connecting industrial phenomena from top-down to bottom-up.

The use of software studies, meanwhile, contextualizes the broader media industrial analysis by placing emphasis on the underlying processes that enable, encode, and enact data-driven industrial practices.<sup>35</sup> Specifically, I focus on Hollywood's integration of a larger data technology stack, a suite of enterprise application software solutions that *establish access to larger databases* (i.e., application programming interfaces, or APIs), *analyze massive amounts of*

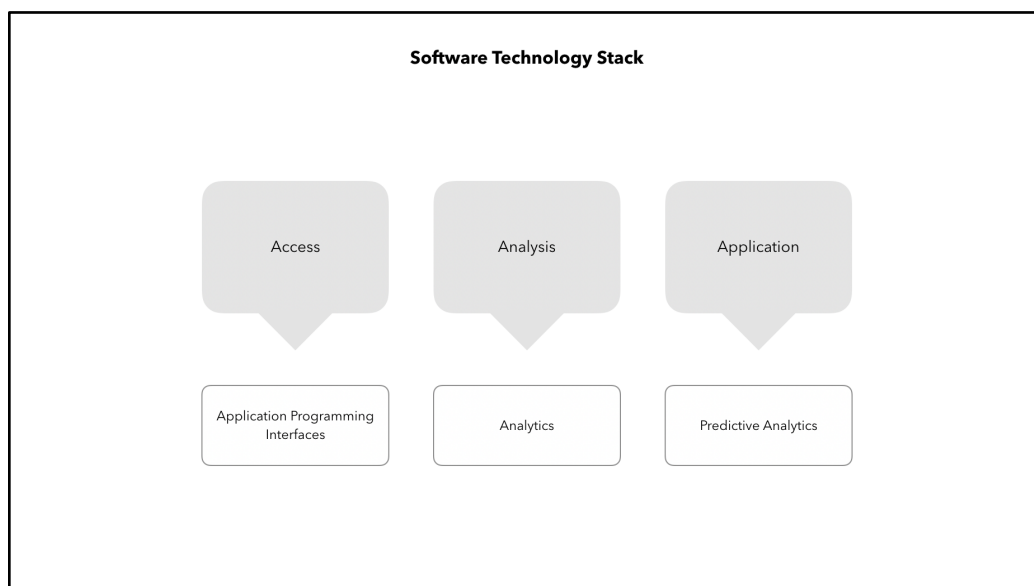
---

<sup>33</sup> See, Caldwell, John T., "Para-Industry, Shadow Academy," *Cultural Studies*, Volume 28, Issue 4, 2014.

<sup>34</sup> See, Timothy Havens, Amanda D. Lotz & Serra Tinic, "Critical Media Industries Studies: A Research Approach," *Communication, Culture & Critique*, 2, 2009.

<sup>35</sup> Following Noah Wardrip-Fruin, the project aims to focus on the "the actual processes that make digital media work, the computational machines that make digital media possible." See, Wardrip-Fruin, Noah, *Expressive Processing: Digital Fictions, Computer Games, and Software Studies* (Cambridge: MIT Press, 2009).

information to generate strategic insights (i.e., analytics software), and apply data-driven insights to execute corporate decisions for legacy media companies (i.e., predictive analytics software) (figure 0.2). The overarching goal is to document the operational flow of data across the legacy media industry and analyze how the majors *make sense, use, and value* of data through the implementation, iteration, and institutionalization of software. As such, the project examines the unifying and differentiating features of data software, with a focus on the dynamic between data input (*the backend infrastructure*) and data output (*the frontend interface*), offering a holistic rendering of data and software connectivity in the industrial context of the media and entertainment industry.<sup>36</sup>



**Figure 0.2:** The software stack maps software applications across key stages of the data value chain, including access (i.e., API software), analysis (i.e., analytics software), and application (i.e., predictive analytics software).

The project further engages a varied set of methodologies to examine the interplay and interconnection of industry, data, and software. Specifically, I draw on discourse analysis,

---

<sup>36</sup> By synthesizing media industries studies with software studies, I build on the idea of “reverse-engineering” disciplinary connectivity to go beyond established research approaches to the media industries’ convergence with technology (see, Vonderau, Patrick, “Industry Proximity,” *Media Industries Journal*, Volume 1, Issue 1, 2014).

software analysis, and visual analysis. Discourse analysis aims to decode deep texts<sup>37</sup> from across the corporate spectrum in an effort to negotiate perspectives on data logic with larger theoretical ideas.<sup>38</sup> To this end, I utilize an expansive set of materials, from journalistic trade articles to public financial documents, corporate promotions, and earnings reports. Additionally, I source public information from a range of new media formats, such as podcasts, blog posts, corporate websites, forum discussions, online panels, webinars, video transcripts, and recorded conference presentations. Finally, I incorporate qualitative and quantitative industry data from a variety of publicly available and reported databases covering a range of information sets, including box office statistics (e.g., *Box Office Mojo*), ratings data (e.g., *Nielsen*), and home entertainment revenues (e.g., *MPAA* reports), digital consumer behavior and viewership trends (e.g., *eMarketer*), corporate funding, transactions, and mergers & acquisitions activity (e.g., *Crunchbase*), workforce data (e.g., *Bureau of Labor Statistics*, *World Economic Forum*), as well as data on *data* (e.g., *Pew Research* reports on Internet and platform usage). The rationale for incorporating this data is to contextualize and connect ideas on industrial integration with larger media and entertainment industry trends.

Software analysis examines the technological properties of software (*how it works*), the corresponding visual manifestation (*what it looks like*), and the resulting implications in the larger industrial context (*where it impacts the industry*). To this end, I study software in two ways. First, I utilize various forms of documentation, including descriptions from corporate

---

<sup>37</sup> To account for the *filtering* (i.e., promotional spin) of corporate texts, I put them in dialogue with journalistic and data sources to arrive at a critical cross-examination of industry positions, perspectives, and performance. I further build on the idea of industry texts as interfaces that facilitate a critical outside-in view into industrial dynamics. See, Ortner, Sherry B., "Access: Reflections on Studying Up in Hollywood," *Ethnography*, 11, 2, 2010.

<sup>38</sup> For a model of this structural approach, see, Caldwell, John T., *Production Culture: Industrial Reflexivity and Critical Practice in Film and Television* (Durham: Duke University Press, 2008).

websites, details from SDKs (software development kits), API (application programming interface) indexes from developer sites, as well as press releases. Second, I focus on the material properties of the visual software interface that consolidates and communicates data to a range of end users, and thereby informs the larger integration of data at the organizational level of legacy media companies. While access to enterprise software is largely limited to paid subscriptions, I analyze interfaces via free-to-use online videos, webinars, images from press releases, and open-access materials from public forums.

Finally, visual analysis is focused on analyzing the way data is communicated through the visual interface of software applications, supporting the broader process of making sense, use, and value of data within the legacy media industry. I further use visual analysis in a practical way, visually mapping the operational dynamics of software in a variety of conceptual figures and tables to identify and illustrate the oft-invisible system of IT software infrastructure, thereby tracing the connective tissue and integration points of the major film studios and television networks, technology companies, and software applications in the emerging data software stack.

Together, the interdisciplinary background and methodological approach aim to offer a multi-dimensional perspective on data as an industrial logic in the legacy media industry.

## **Overview and Organization**

The dissertation follows a customary structure. As such, it is divided into three parts, which collectively trace Hollywood's integration of data as an industrial logic over the course of the early 21st century, covering the period from the late 1990s to the late 2010s in a non-linear, conceptual sequencing dynamic. Each part covers a critical stage of Hollywood's integration with the data value chain, from *data access* to *analysis* and *application*.

Each part further comprises two key elements: an introductory *context* section and a comprehensive *chapter*. The context section is designed to outline the industrial circumstances of data integration and propose a conceptual framework for analysis. In particular, it works to provide a foundational perspective on Hollywood's complex convergence with the technology industry (i.e., mapping the ways the majors adopted and adapted Silicon Valley data practices from platforms) and offer a toolkit to examine the corresponding data integration strategy. Since the project is primarily focused on analyzing Hollywood's integration of data through the perspective of legacy media companies, the context section pulls out a brief preliminary perspective on the technology industry and its impact on Hollywood, specifically the role of platforms, which remains essential to contextualize the data dynamics of the majors. Effectively, the context section serves as a short introduction for each main chapter.

The chapter format forms the main body of each part and examines how the majors accessed, analyzed, and applied data as an industrial logic. Each chapter puts the respective contextual framework into action, exploring the strategic developments, operational patterns, and industrial use cases that define the dynamics of each integration stage, from data access to analysis and application.

The rationale behind this modular organizational structure is to preview and unpack the larger conceptual ideas at play in Hollywood's industrial integration of data (*context focus*) while afterwards providing a deeper contextual analysis through the lens of the legacy media industry (*chapter focus*).



## Part I – Data Access

The context section, titled *Data and Information Building*, outlines some of the key industrial dynamics that define data access in the digital economy. As such, it contextualizes the majors' data integration as an exercise in information building by negotiating various data stakeholders, from platforms to third-party technology companies. It further maps out a contextual framework designed to define the access mechanisms by which the majors made data increasingly *available*, *accessible*, and *actionable* within an expanding data access infrastructure.

Chapter 1 then examines the legacy media industry's investment in building a new information infrastructure to unlock data access by moving from lacking to *generating*, *licensing*, and *owning* data access. Specifically, it places emphasis on the integration of three key access mechanisms. *Data Generation* covers the industry's response to managing the exponential increase of information across the digital landscape. *Data Licensing* highlights a multi-dimensional partnership model forged between the majors and a suite of technology companies to negotiate data access at multiple levels. *Data Ownership* showcases the majors' emerging attempts to build their own data access infrastructure, increase their data capital share, and lessen reliance on technology players.

## Part II – Data Analysis

The context section, titled *Data and Decision-Making*, outlines the impact of analytical organizational cultures and modes of decision-making on the legacy media industry.

Specifically, it introduces the underlying industrial processes that informed the majors' move toward establishing a data analysis system. It concludes by outlining a framework to account for new organizational structures focused on data analysis within the legacy media industry.

Chapter 2 then examines the legacy media industry's formation of an emerging culture of data analysis by way of three key industrial developments. *Data as a Service* focuses on the technological capabilities and human resources that increasingly enabled the majors to translate data information into data insights, thereby facilitating analytical modes of decision-making across their organizational structures. *Data as Intelligence* expands on this new infrastructure by contextualizing the rise, development, and impact of new data metrics, specifically with regard to Hollywood's way of measuring its core value levers, namely content performance and audience behavior. *Data as Storytelling* demonstrates how data analysis software encodes, enables, and enacts the industry's new measurement system across organizational structures through an industrial analytics interface.

### Part III – Data Application

The context section, titled *Data and Cultural Production*, introduces the technology industry's practice to apply data as a driver of cultural production in a new media landscape. It then sketches a framework to account for the majors' increasing reliance on data-driven algorithms and automation to manage the legacy media business. In particular, it contextualizes the industry's efforts to sustain and supercharge established modes of cultural production with data in an effort to mitigate the risk, challenge, and uncertainty deriving from the digital economy.

Chapter 3 then examines the legacy media industry's application of data across the media and entertainment value chain to preserve the economic and operational dynamics of the legacy media business. To this end, it examines three new operational modes of data-driven cultural production, designed to reframe the creative business through an analytical, rational, and risk-averse lens. *Data-Driven Production* examines the application of data to optimize pre-production

and greenlighting decisions. *Data-Driven Marketing* focuses on the optimization of marketing campaign operations and efficiency. *Data-Driven Distribution* analyzes the strategies to optimize audience turnout and tune-in.

The ultimate goal of this organization is to draw a conceptual map detailing key developments that made data one of the media and entertainment industry's core logics and, by extension, offer a contributing perspective to help guide scholarly discourse in examining the legacy media industry's ongoing integration of data as part of a larger complex and continuing convergence between Hollywood and Silicon Valley.

## Part I – Data Access

## Context

### Data and Information Building

“The future lies in data [...] The world is all about data.”<sup>39</sup>

Rich Greenfield, Media and Tech Analyst

In a conversation on the Recode Media podcast, Wall Street analyst Rich Greenfield argued that the entrance of technology platforms, such as Netflix, Amazon, and Google, has fundamentally changed the legacy media industry, for several reasons. First, platforms have introduced a direct-to-consumer business model, which fundamentally differs from the legacy media industry’s wholesale model. The major film studios and television networks traditionally partnered with intermediaries, including cable companies, telecommunications providers, exhibitors, retailers, and advertisers, to license and distribute content across channels. The platform approach, by contrast, removed this set of intermediaries to distribute content directly to consumers. Second, platforms have inserted themselves into the wholesale model of the legacy media industry by taking on the role of digital distributors, delivering content from the majors across digital channels. And, third, platforms have expanded beyond distribution into programming, effectively becoming competitors to the majors.

This industrial transformation is inherently linked to platforms’ exclusive access to data. Indeed, the evolving state of the digital media economy is increasingly defined by data access, with platforms accessing data across their ecosystems to create a competitive business advantage over media incumbents. Netflix is frequently cited as a paradigmatic expression of this model.

---

<sup>39</sup> Kafka, Peter, “BTIG Analyst Rich Greenfield on Recode Media,” *Vox*, June 15, 2017, accessed April 3, 2019, <https://www.vox.com/2017/6/30/15904416/transcript-btig-analyst-rich-greenfield-recode-media-peter-kafka-podcast>

The streaming company delivers film and television programming via its website and app to consumers, removing traditional intermediaries to directly access the resulting data. Netflix then analyzes the data to gain insight into behavioral viewing patterns and preferences and applies the insights to inform programming decisions. In this scenario, access operates as the foundational element of Netflix's business model.<sup>40</sup>

By contrast, data access is generally viewed as an unknown variable and disruptive element for the legacy media industry given film studios and television networks' reliance on partnerships with intermediaries which has outsourced any direct relationship with consumers.<sup>41</sup> Since the majors do not own any direct distribution outlets, they lack the ability to access any data on consumers and their interaction with content in the shifting media landscape. Indeed, Hollywood's wholesale model has invariably focused on *content-push* rather than *data-pull*, effectively disintermediating the industry's access to the data stream of the digital economy.

This contextual outline offers a useful starting point to examine Hollywood's industrial integration of data access. One, data access is an underlying factor of the digital economy where platforms distribute content directly to consumers and own access to the resulting data. Two, data access is a foundational element in how platforms operate, leveraging access to data in order to inform, implement, and iterate decision-making in their business. And, three, data access is a feature of the consumer-centric technology business, whereas legacy media companies are generally removed from the digital information flow. To this end, I argue that Hollywood's

---

<sup>40</sup> Technically, even Netflix works with intermediaries as its app is featured across digital outlets, including smart TVs, digital media players, and distribution platforms. Yet, Netflix still largely controls the resulting data.

<sup>41</sup> Prior to the 1948 Paramount decree, film studios owned theaters directly and via controlling partnership stakes. See, Balio, Tino, *The American Film Industry* (Madison: University of Wisconsin Press, 1985).

integration of data access can be conceptualized as an industrial exercise in building up information to address a *data gap*.

Over the 21st century, data has emerged as a central output of the media and entertainment industry. Indeed, the legacy media industry has expanded its distribution approach from linear channels to digital platforms, beginning with websites and extending into search (e.g., Google searches), social media (e.g., Facebook pages, Twitter accounts), mobile (e.g., mobile games, second-screen apps), and over-the-top, Internet-distributed video, or OTT (e.g., YouTube videos, Netflix streams). While the digital landscape allowed the majors to deliver content in new ways, increase audience reach, and connect with consumers across non-linear channels, it further worked to establish a foundation for the ongoing generation of data on media audiences and their interaction with content.<sup>42</sup> Effectively, the shift to digital distribution has made data an incremental ingredient of the legacy media business, ever-present, all-encompassing, and ubiquitous.<sup>43</sup> As film and television paratexts migrated across the digital landscape, platforms recorded the interaction between consumers and content as quantifiable information, automatically translating *content input* into *data output* (i.e., data as a generative product of the digital sphere).<sup>44</sup> As a result, media content came to generate growing sets of data by leaving a “trail of audience feedback.”<sup>45</sup> In this sense, Hollywood’s digital distribution

---

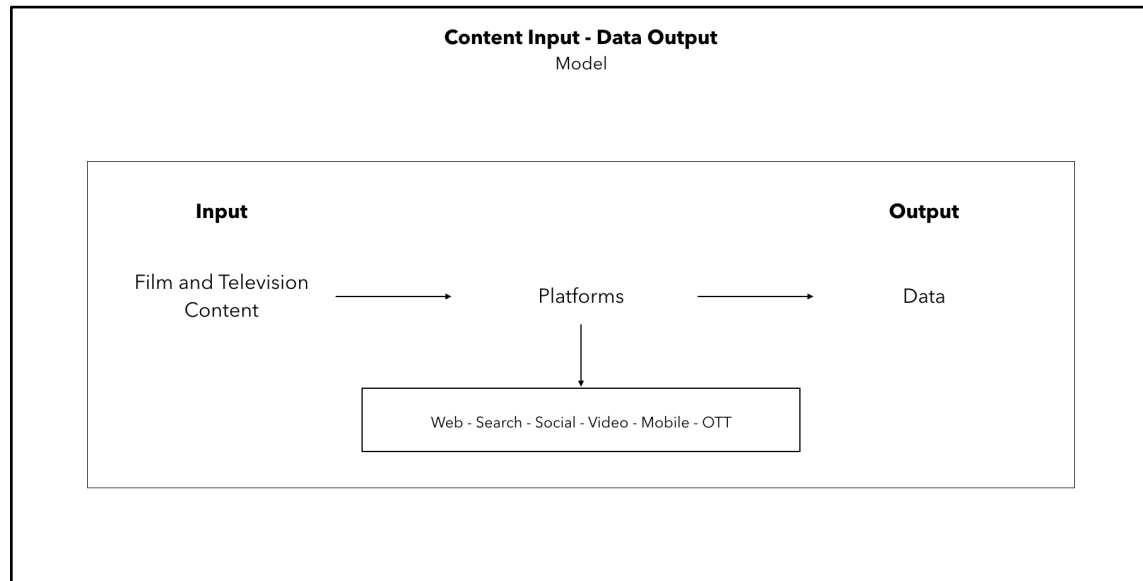
<sup>42</sup> See, Hesmondhalgh, David, *The Cultural Industries* (New York: Sage, 2019).

<sup>43</sup> See, Mayer-Schoenberger & Cukier, 2013 and Lycett, 2013

<sup>44</sup> The data output is dynamic rather than static, consistently informed by, and dependent upon, the platform from where it is generated. See, Gitelman, 2013

<sup>45</sup> Napoli, 2010

network came to act as a data generation mechanism, where content input generates an infinite canvas of data output (figure 1.1).



**Figure 1.1:** Content input and data output model illustrating how platforms across the digital landscape translate film and television content input into quantifiable data output.

Yet, while film studios and television networks have actively contributed to the generation of data by distributing content across the digital landscape, their access to the resulting data output has remained limited. Indeed, data access is largely proprietary and restricted to the platforms where the data is generated. Technology platforms invariably own the data output, even if generated by the content input from the majors.<sup>46</sup> Consider the following: when Netflix subscribers watch a Disney movie on Netflix, Netflix has access to the resulting data, not Disney; when fans watch a trailer for a new *DC Comics* movie on iTunes, the data is visible to Apple, not WarnerMedia (owner of DC Entertainment); and, when Amazon customers watch *Game of Thrones* via the HBO channel app on Amazon Prime Video, HBO misses out on

---

<sup>46</sup> For example, Netflix built a repository of film and television data on licensed programming from the majors.



the data in favor of Amazon.<sup>47</sup> The result is a media environment where data is universally available, yet only selectively accessible. While data is framed as a universal entity of information, access is highly fragmented and selective.<sup>48</sup> Platforms own access to data generated across their digital infrastructures, while the majors actively generate data from their content, with only limited access.

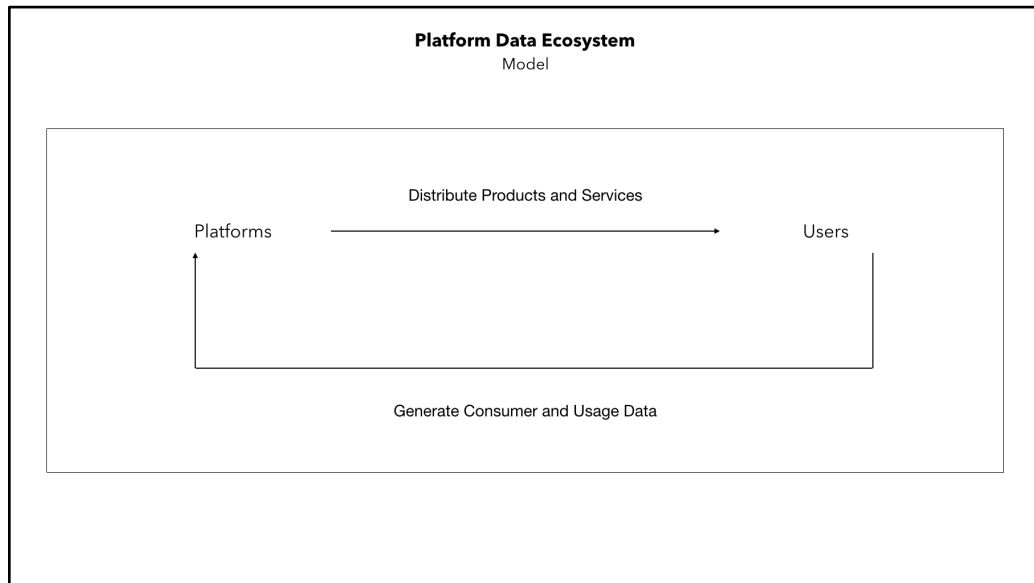
Hollywood's integration of data access thus inevitably raises questions about ownership. Data ownership structures effectively define, drive, and determine the dynamics of access. Technology platforms like Facebook, Apple, Amazon, Netflix, and Google (i.e., FAANG) can be classified as data owners. Data owners collect data through a direct-to-consumer business model that maintains a direct line between the platform (i.e., the underlying technology stack and the user interface) and users. Platforms thus have access to proprietary data drawn from their digital product portfolios (figure 1.2). As part of that construct, platforms equally have access to a wide range of data that is related to and generated by the majors, since film studios and television networks distribute content across platforms. As a result, platforms effectively manage and control the legacy media industry's underlying digital information sets. The resulting industrial environment marks a complex dual structure. In their role as content distributors, platforms have enabled the majors to distribute their content digitally, yet, in their role as data owners, platforms have consistently kept the data the majors generate for themselves, enabling them to access a vast database on film and television information that cuts across all the studios and networks.

---

<sup>47</sup> This procedural phenomenon is by no means a hard-and-fast rule, but a directional guideline. Indeed, data ownership structures are complex, highly dependent on technological, economic, and regulatory factors.

<sup>48</sup> See, Wallenstein, Andrew, "Big Data: Media Embracing the Most Detailed Information About You Yet," *Variety*, September 25, 2013, accessed June 5, 2019, <https://variety.com/2013/biz/news/big-data-media-embracing-the-most-detailed-information-about-you-yet-1200665847/>

Platforms have effectively come to own the data on Hollywood’s consumer and content base across the digital landscape.<sup>49</sup>



**Figure 1.2:** Platforms are vertically-integrated ecosystems that pull in data through a feedback loop from users via the distribution of digital products and services (i.e., social networks, online video, apps, games).

In this regard, data access effectively constitutes a form of digital industry capital.<sup>50</sup> Indeed, platforms manage, mediate, and monetize access across the digital landscape to benefit their business, in several ways. One, they share access to their data with external companies in order to promote and validate use of their platform products and services. As such, they aim to incentivize companies to buy advertising (i.e., by targeting advertisers), develop content (i.e., by targeting publishers), and create new products (i.e., by targeting software developers) on their digital infrastructures.<sup>51</sup> Two, platforms license access to third-party technology intermediaries,

---

<sup>49</sup> For a comprehensive overview of the platform business model, see, Cusumano, Michael A., Gawer, Annabelle & Yoffie, David B., *The Business of Platforms: Strategy in the Age of Digital Competition, Innovation, and Power* (New York: Harper Business, 2019).

<sup>50</sup> See, Wu, 2017

which work to capture, consolidate, and curate data from a range of digital channels and platforms for streamlined access. Effectively, these intermediaries pay a licensing fee to access data from platforms and make it available to external companies for a comprehensive data view that single platforms are unable to provide. Licensing partnerships with intermediaries make platform data more widely available and enable individual platforms to further promote and validate their data to external companies.<sup>52</sup> And, three, platforms actively control access by sharing and licensing data selectively, limiting the quality and quantity of data external companies can access.<sup>53</sup> Thus, platforms share, license, and control data access through a software mechanism, using application programming interfaces (APIs) to enable external companies to plug into their databases. While they provide access, it is always selective, designed to be fragmented and limited, enabling platforms to maintain control over data.<sup>54</sup>

Accordingly, data access represents an industrial construct driven by economic, cultural, and technological considerations. The number of players involved is complex and layered, yet can be distilled through the lens of data as capital, effectively focusing on companies that own, share, and license data:

---

<sup>51</sup> The major film studios and television networks effectively act as advertisers, publishers, and developers.

<sup>52</sup> The practice of data sharing and licensing is not universally valuable to platforms. Advertising-based platforms derive value from providing external data access to showcase the results of their platform to advertisers. Subscription platforms, by contrast, have no business incentive to share or license data with external partners.

<sup>53</sup> At times, platforms have even removed access entirely. See, Lowensohn, Josh, "Netflix will close its public API to some developers in November," *The Verge*, June 13, 2014, accessed March 5, 2018, <https://www.theverge.com/2014/6/13/5808424/netflix-will-close-its-public-api-to-some-developers-in-november>

<sup>54</sup> It is important to note that the data access of platforms is also limited, not absolute. Platforms only own the data in their own digital ecosystems. No one data agent owns total access to all data in the digital economy.

- *Platforms*: Platforms own and use proprietary data to manage their business model. They share and license data to partners and third parties to grow and expand their business. They restrict and gate data in order to protect their business. They promote and editorialize data as a way to pitch their business to partners and consumers as well as differentiate it from competitors.
- *Third-party technology intermediaries*: These software companies synthesize, consolidate, and package data from platforms through paid licensing partnerships and make the information available to external companies in exchange for a fee. As such, they effectively offer access to data across platforms. While platforms own the underlying data, they are unable to provide this type of cross-platform access since they can only draw data from their own infrastructure, not from other platforms. To this end, platforms benefit from partnering with third-party intermediaries as a way to further monetize data access.
- *Studios and networks*: The majors partner with platforms and third-party intermediaries on data access, effectively maintaining a network of business partnerships to access data across the digital landscape.

In this constellation, platforms, third-party intermediaries, and the majors effectively trade, broker, and capitalize access in a complex and interdependent data economy, fueled by the logic of data as capital.

The idea of a data access economy reinforces the data gap between platforms and the majors. While platforms are characterized by data access, Hollywood's infrastructure is defined by the lack thereof. While the platform business model is built on data input (i.e., processing data

to iterate digital products and services), the majors have long traded in a model that emphasizes content output (i.e., maximizing the release of content across distribution channels to drive share). And while platforms connect directly with consumers, studios and networks have traditionally outsourced this relationship to a range of intermediaries. Thus, the business model of the legacy media industry has emphasized *pushing out content over pulling in data*.

This data access gap has created a pressure chamber for Hollywood, casting doubt on the legacy media industry's ability to negotiate the digital era, from safeguarding a profitable media business to competing with platforms. Investors and analysts have expressed caution and uncertainty about the future data state of the legacy media business, which has impacted media stock performance on Wall Street. As Greenfield expressly noted, "[...] here we have an entire industry that's basically getting beaten up day by day because they have no data, they have no direct relationship with the consumer."<sup>55</sup> In effect, data access has created an existential dichotomy between Silicon Valley digital platforms and Hollywood legacy companies. Platforms own data, are data-rich, and have information access while the majors lack data, are data-less, and have no viable information at their disposal. The result is an entrenched oppositional constellation wherein platforms have been positioned as more competitive than the majors due to their data access, while the legacy media industry has been characterized as cut off from the digital data stream, lacking sufficient economic models, cultural-organizational practices, and technological capabilities to unlock the capital of data access. Yet, upon closer inspection, the dynamics of data access and ownership prove more intricate and layered than opposing ends of a spectrum. Indeed, the major studios and networks have displayed a far more active role in managing and negotiating the data access gap.

---

<sup>55</sup> Kafka, 2017

First, the legacy media industry's lack of data access is primarily defined through the lens of *first-party data* (table 1.1), rather than a universal stream of digital information. As such, the majors lack access to proprietary data sets owned by platforms. Indeed, given the legacy media industry's long-standing lack of a digital platform infrastructure, the majors only have limited access to first-party data.<sup>56</sup> To address this data gap, the majors have developed a variety of strategies to license, acquire, and incubate first-party data.

Second, the majors have maintained varying degrees of access to *second-party data*. Second-party data is select information that is shared by platforms as part of business partnerships. For example, retailers such as Target and digital marketplaces like Apple have shared transaction-level data (i.e., sales figures) with the majors in home entertainment. Additionally, ad-supported platforms like Facebook and YouTube share select data on digital advertising campaigns. In both cases, platform partners offer up data to showcase the value of their services and incentivize the majors to continue investing in their partnership. To expand access to second-party data, the majors have cultivated a range of partnerships with platforms.

And, third, the legacy media industry has maintained a system of *third-party data* access that consistently delivers external information input. Third-party data represents a custom consolidation of data accumulated from proprietary and public sources by third-party intermediaries. In the digital era, these companies generally access data by leveraging software technologies like APIs, platform licensing agreements, and long-standing industry partnerships, and then resell access in form of packaged research reports or a software subscription.<sup>57</sup> The

---

<sup>56</sup> Public data refers to information that is visible across the frontend of the digital infrastructure and can thus be seen, read, and scraped directly from digital platforms. Some select examples include YouTube's video views, the number of search results displayed from a Google search, and Facebook's like button.

<sup>57</sup> See, Matt Murphy & Steve Sloane, "The Rise of APIs," *TechCrunch*, May 21, 2016, accessed July 5, 2019, <https://techcrunch.com/2016/05/21/the-rise-of-apis/>

majors have long cultivated partnerships with analog third-party intermediaries, notably market research companies, and have effectively extended this approach to the digital era, partnering with a wide set of third-party technology companies to access a mix of public and proprietary data from the digital landscape.<sup>58</sup>

**Table 1.1**

First-Party Data	Second-Party Data	Third-Party Data
Owned	Shared	Licensed
Data is owned by platforms which exercise control over storage and usage	Select data is shared by platforms as part of business agreements with partners	Select data is licensed by platforms to intermediaries which, in turn, license data to external companies

Accordingly, the legacy media industry has cultivated an information infrastructure with layers of data access. As such, the industry is less defined by a general data gap, than by an active negotiation to access first-party data through a direct relationship with consumers. Due to the long-standing reliance on intermediaries to deliver content, the majors’ access to first-party data has been structurally disintermediated. As Smith and Telang note, “[the creative industries] have had very little first-hand information about consumers’ behavior.”<sup>59</sup> In the legacy model of distribution, access to first-party data did not provide a competitive advantage for the majors since the entire model was designed to push out content, rather than pulling in data. Yet, the rising viability of the direct-to-consumer model, as demonstrated by the growing success of platforms, has made first-party data a key variable in managing the media business. As such,

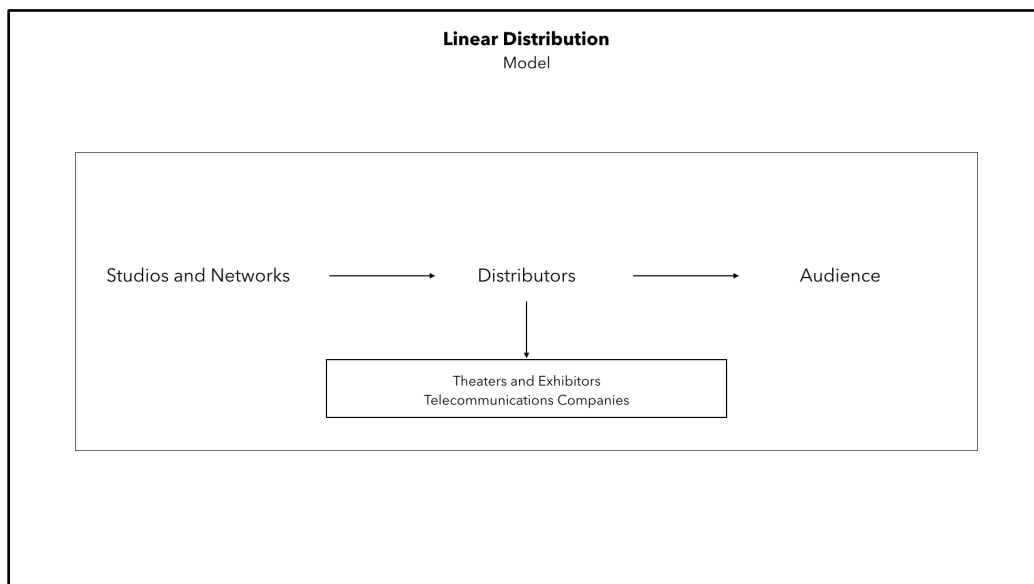
---

<sup>58</sup> The majors have employed research firms to conduct audience research since the 1930s, specializing in focus groups, surveys, and questionnaires to provide aggregate statistics rather than free-flowing digital data. See, Buzzard, Karen, *Tracking the Audience: The Ratings Industry From Analog to Digital* (New York: Routledge, 2012).

<sup>59</sup> Smith & Telang, 2016

studio and network executives have acknowledged the industry’s access limitations. Michael Lynton, former Chief Executive Officer of Sony Pictures, summarized the situation pointedly, “We don’t have that direct interface with the American public.”<sup>60</sup>

The legacy media industry has consistently negotiated this missing link to the audience. In linear distribution, the majors relayed first-party data access to analog intermediaries (figure 1.3).<sup>61</sup> In film, theatrical exhibitors maintained the direct transactional relationship with the audience. In television, telecommunications providers managed partnerships with advertisers (i.e., network television) and subscribers (i.e., cable television). In retail, retailers sold content and merchandise directly to consumers through physical stores and digital channels.



**Figure 1.3:** Linear intermediaries act as the connection point between the majors and the audience.

The digital era expanded the data dynamics of the linear distribution model. Indeed, the majors have come to negotiate a complex set of industrial relationships with platforms to access

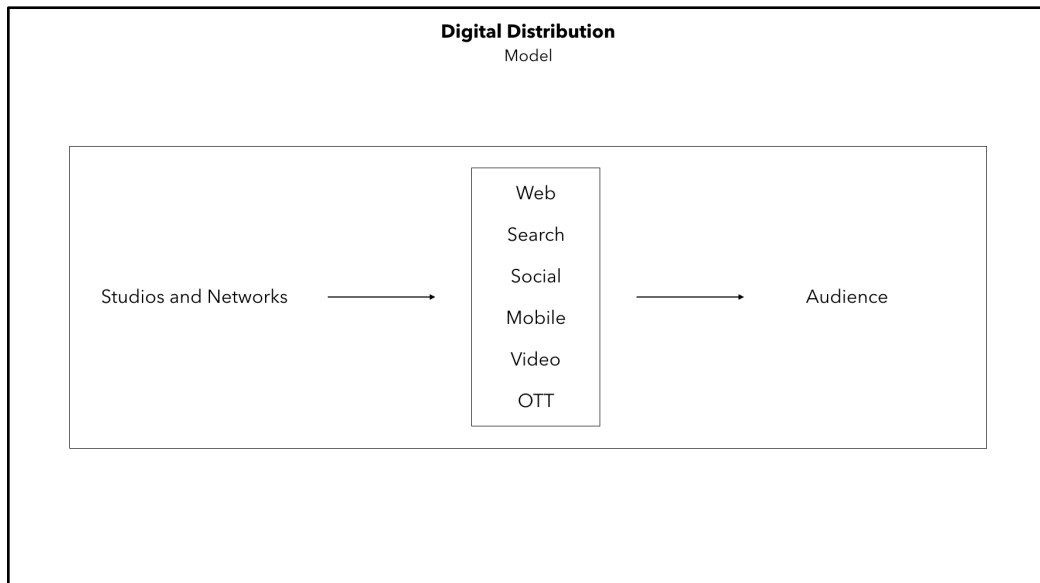
---

<sup>60</sup> Quoted from Ibid. Also see, CNN Press Room, December 19, 2014

<sup>61</sup> Analog intermediaries rarely collected data on consumers since they emphasized revenue over information. See, Wasko, 2003 and Balio, 1985



first-party data. Data access is no longer distributed across a small set of analog intermediaries, but a multitude of platforms that operate across several channels (e.g., web, search, social, mobile, online video, OTT) (figure 1.4).



**Figure 1.4:** Digital distribution expands the infrastructure of intermediaries across various channels and platforms, enabling multiple connection points between the majors and the audience.

The digital distribution infrastructure does not represent a unifying, comprehensive, and all-inclusive meta-database. Instead, it operates as a complex data spectrum defined by various stakeholders. In this framework, data access is not straightforward, but fragmented, disintermediated, and messy. Data is owned in parts by a small set of platforms, shared and licensed selectively among platforms and intermediaries, and accumulated, synthesized, and packaged by third-parties. As such, data access is constantly in flux, which effectively dissolves the static dichotomy of owning versus lacking data access. Rather, data access is about consistently negotiating varying data types, sources, stakeholders, and underlying ownership structures. Platforms represent data owners; the major studios and networks act as data generators, creating new information on the audience and their interaction with content; and

various technology companies function as data intermediaries, capturing, sharing, and licensing information. Data is captured, collected, and commercialized among all these players. It is shared, licensed, and sold; synthesized, consolidated, and packaged; brokered, negotiated, and traded. Data access is not a static state, but a dynamic process.

To this end, Part I aims to address the following questions on Hollywood's industrial integration of data access:

- What industrial developments and business strategies shaped Hollywood's integration of data access?
- What are the cultural, economic, and technological considerations of data access and how did they impact corporate machinations in Hollywood?
- How did the legacy media industry manage and negotiate data access in an environment controlled by platforms, governed by data ownership structures, and defined by digital economics?

To answer these questions, it is crucial to conceptualize access as a dynamic, shifting, and evolving process of capturing, collecting, and commercializing data. Access is not a linear state, but a complex industrial convergence of multiple stakeholders and practices.

### **A Framework for Data Access**

Access represents a dynamic process of information building that involves a core set of cultural, economic, and technological elements. These elements help to examine how the majors have created, coordinated, and commercialized data access. Put differently, they serve to illuminate

the ways data access has manifested within organizational structures of legacy media companies.

To this end, I propose the following elements of data access:

- *Data Availability*: While data is frequently characterized as a universal entity, it is not universally available. Data is generated, captured, and stored across a wide array of platforms, which makes access inherently limited, fragmented, and restricted. Platforms use software to make data digitally available to organizations and control the level of access through technology. Thus, data is not a physical resource, but a digital currency that can be accessed and made available through technology. It is thus vital to consider the technological dimension of data access and how data is made available within corporate structures.
  
- *Data Usability*: Adopted from software engineering theory, the concept of usability refers to a seamless, easy-to-use software design that enables any user to interact with software.<sup>62</sup> Data usability thus relates to the practice of making data access seamless and easy-to-use within media organizations, which includes several factors. One, media organizations consist of multiple businesses with varying data output (i.e., a film studio generates distribution data, a television network generates different sets of distribution and advertising data). Making data usable requires sharing data access across departments and teams, establishing a practice to minimize data silos and mitigate fragmentation. Two, media organizations have traditionally worked with legacy data sets, such as output from surveys, polls, and focus groups, and are thus not generally accustomed to the scale and detail of digital data. Establishing usability therefore involves making data relatable,

---

<sup>62</sup> See, Bratton, 2016

visible, and easy-to-understand for users, creating a “level of data organization, which is made visible and accessible to a user and thus becomes part of his/her mental model.”<sup>63</sup>

As such, making data usable involves a cultural-organizational component that is interrelated with technology.

- *Data Actionability*: Data operates as a form of capital in the digital economy. Yet, the value of data is not universal, but specific to stakeholders and their business logic. Put differently, platforms value the types and sources of data that carry the most benefit for their business, the highest degree of capital. For example, Netflix is focused on accessing OTT data on viewing behavior and preferences to operate its video business, yet does not derive the same value from social media data since it is not relevant to its business model. Likewise, Facebook prioritizes the demographics and psychographics of social data over any other data source in alignment with its goal to create value for its advertisers. Establishing access thus involves a consideration of the economic value of data for an organization. Rather than emphasizing access to a massive bulk of data, the industry is driven by a focus on specific actionable data.

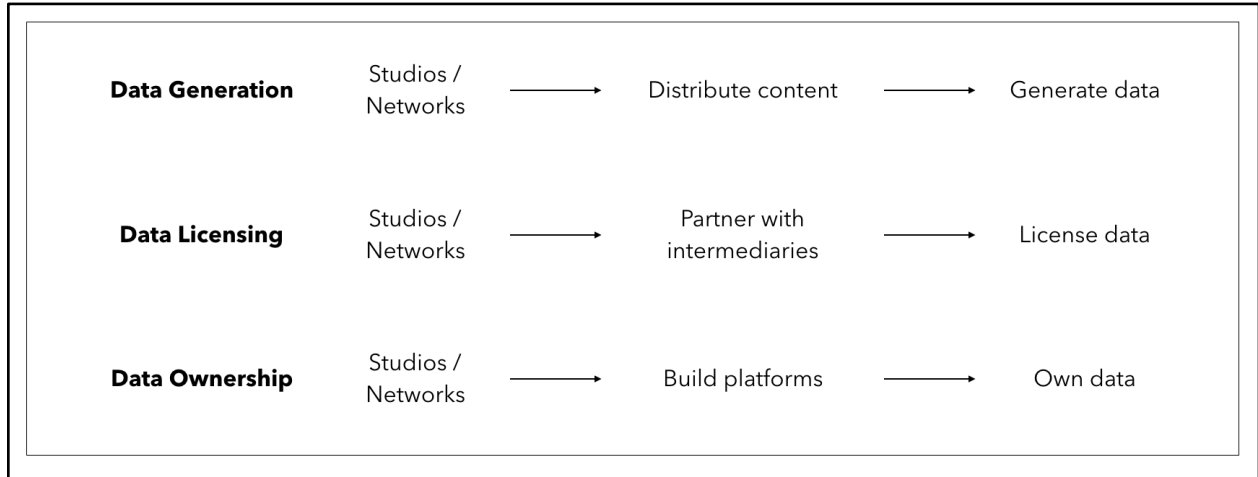
Data access operates across all three elements as a dynamic, layered, and complex industrial process. As such, Part I traces the industrial developments and strategies that have made data technologically *available*, culturally-organizationally *usable*, and economically *actionable* within the legacy media industry. Furthermore, it examines the ways in which the industry has negotiated this process through an analysis of key stages of access:

---

<sup>63</sup> Manovich, 2013

- *Data Generation*: From the late 1990s onward, the legacy media industry has seized upon a growing digital infrastructure to market and distribute films and television shows online. This focus on digital marketing and distribution has created a mechanism for ongoing data generation. As studios and networks release content across the digital landscape, they generate growing amounts of data across platforms. As such, the industry has come to consistently generate data, making it *available* across the digital economy.
- *Data Licensing*: The legacy media industry has developed partnerships with a wide range of third-party technology companies to leverage software applications in an effort to gain *usable* data access. Platforms and digital intermediaries use different mechanisms to make a variety of data sources available to the majors, including providing access to public data sets, sharing data related to business transactions (i.e., marketing campaigns for films and television shows), and licensing data for a fee. Accordingly, the majors have developed partnerships to access an integrated suite of first-party, second-party, and third-party data.
- *Data Ownership*: The legacy media industry launched its own set of digital, direct-to-consumer platforms, adopting core tenets of the platform business model, such as subscription and ad-supported monetization and a digital video interface, to create a working foundation to own first-party data on their content and audience, thereby making data access increasingly *actionable* for the media business.

In this sense, Part I traces Hollywood's development from a state of data disintermediation (i.e. the majors lacking data access) to a complex state of *generating, licensing, and owning* data access (figure 1.5).



**Figure 1.5:** The major film studios and television networks unlocked data access in three interrelated stages, actively moving from generating to licensing and owning data access.

These stages are conceptual rather than historical. As such, they exist symbiotically, influencing each other's development. While the rise of new technologies has facilitated distinct data access capabilities within the digital landscape, technology is only one factor in Hollywood's integration of data access. Data generation, licensing, and ownership are driven by a network of cultural, economic, and technological factors, and result from the majors' proactive, strategic decision-making as well as reactive, ad-hoc adjustments to external factors.

In effect, Hollywood's data access development underscores the legacy media industry's convergence with platforms in the digital economy. Film studios and television networks did not operate like platforms and thus lacked the corresponding capabilities to access data in the digital economy. Yet, the majors have adopted and adapted a distinct set of data access mechanisms, enabling them to emulate, embed, and engage the business logic of platforms. As such, they have increasingly taken control of the industrial process that is data access.

## Chapter 1

### A Century of Access: Building a New Information Infrastructure

“The industry is essentially a B2B business. They don’t have access to nor collect much data about their end-consumer.”<sup>64</sup>

Yves Bergquist, USC Entertainment Technology Center

At the turn of the century, America Online (AOL) took a step toward a new media future by acquiring Time Warner, Inc. in an unprecedented \$182 billion deal,<sup>65</sup> combining the country’s leading tech giant with the media industry’s largest conglomerate to create “the world’s first fully integrated media and communications company for the Internet Century [...]”<sup>66</sup> The revolutionary merger bridged the media and tech industries, giving the newly formed entity a market capitalization of \$350 billion, which made it one of the most valuable companies in the world. *AOL-Time Warner* became an early symbol for the convergence of old and new media, illustrating the integration of two companies that shared common business goals, yet highly differentiated cultural-industrial identities. Time Warner offered decades-long expertise in film and television cultural production (along with news and publishing), while AOL signaled the future of Internet-distributed digital content, grounded in a complex technological infrastructure.

---

<sup>64</sup> Weaver, Erik, “Three Ways AI will Solve Hollywood’s Data Problem,” *Western Digital Blog*, May 11, 2017, accessed June 5, 2020, <https://blog.westerndigital.com/ai-will-solve-hollywoods-data-problem/>

<sup>65</sup> See, The Economist, “Who’s afraid of AOL Time Warner?,” *Media Giants*, January 24, 2002, accessed June 5, 2020, <https://www.economist.com/business/2002/01/24/whos-afraid-of-aol-time-warner>

<sup>66</sup> Time Warner Press Release, “AOL & Time Warner Will Merge To Create World’s First Internet-Age Media & Communications Company,” January 10, 2000, accessed June 5, 2020, <https://www.warnermediagroup.com/newsroom/press-releases/2000/01/10/aol-time-warner-will-merge-to-create-world-s-first-internet-age>

- *Time Warner*: The largest media and entertainment conglomerate in the world, Time Warner owned a breadth of content assets, including the Warner Bros. movie studio, the Turner cable network, premium cable channels HBO and Cinemax, news network CNN, music publishing company Warner Music Group, the Time, Inc. magazine empire, as well as a massive vault of classic and contemporary IP with in-built global brand recognition, including rights to the *Batman*, *Harry Potter*, and *Lord of the Rings* franchise IP. As such, Time Warner represented a leading creative conglomerate, delivering content and programming at a global scale, effectively supplying what audiences watch, read, and listen to, and accounting for a large share of consumer engagement with legacy media channels in film, broadcast, and cable television.
  
- *AOL*: The world's leading Internet service provider (ISP) and digital bulwark, AOL owned a walled garden of numerous online assets,<sup>67</sup> including the AOL search engine, web browser Netscape, web mapping service MapQuest, online service provider CompuServe, instant messenger platform AOL Instant Messenger (AIM), a library of online games, and over 20 million dial-up subscribers, as well as a bundling partnership with software giant Microsoft, which ensured global exposure of its service offering to a rapidly growing digital customer base. As such, AOL managed a wide range of interconnected online services, capturing the majority of digital consumer activity by enabling users to go online, search for information and content, message and connect with other users, play games, and watch video programming.

---

<sup>67</sup> Some of the company's assets followed a pay-to-play business model, with AOL charging a subscription fee.



The unprecedented merger of old and new media was rationalized by executives and shareholders under the promise of synergy,<sup>68</sup> namely the convergence of Time Warner's content assets and production capabilities with AOL's digital distribution infrastructure. The vision was to elevate, extend, and expand the business by converging content and distribution to achieve more scale, reach more consumers, and maximize revenue. In this scenario, AOL-Time Warner's combined audience reach, content assets, and global delivery system would enable the company to scale advertising revenue (i.e., sell more ads across legacy and digital media channels), subscription revenue (i.e., incentivize more consumers to sign up for AOL's walled garden of paid services as well as Time Warner's roster of paid legacy media products like magazine and Pay-TV subscriptions), and consumer engagement (i.e., keep audiences firmly locked into AOL-Time Warner's newly developed ecosystem of content and services). Time Warner's content catalog, coupled with its presence in 20% of U.S. households through a robust cable television business, was the glue to AOL's digital media infrastructure with over 20 million users, and the promise of a growing global audience in the offing as more and more consumers went online. In effect, AOL-Time Warner was positioned as a vertically-integrated media and technology company that creates, programs, and delivers content across a proprietary infrastructure of legacy and new media channels with unprecedented scale and stickiness.

AOL-Time Warner's core competitive differentiator was said to be the convergence of content and distribution, the synergistic combination of media and technology into a new seamless and holistic business ecosystem of experiential services. Yet, this vision of digital revolution was grounded in an industrial factor that has received less attention, namely the

---

<sup>68</sup> See, Meehan, Eileen R., "'Holy Commodity Fetish, Batman': The political economy of the commercial intertext." In: *Hollywood: Cultural Dimensions: ideology, identity and culture industry studies*, edited by Thomas Schatz (New York: Routledge, 2004).

company's direct connection to the audience and, as such, a unique mechanism to access data in a way previously unavailable to legacy media companies.

Indeed, AOL-Time Warner, at its core, formed a new infrastructure of information, an ecosystem where data would travel seamlessly across multiple business divisions, revealing information that would inform and optimize the company's portfolio of products. AOL's online service channels would enable access to a range of consumer information, including basic demographics, such as age, gender, and location, email addresses, and credit card details, as well as behavioral and usage information, capturing clicks, sign-ups, and time spent on digital services. Time Warner, meanwhile, would add a content layer on top of the information infrastructure, enriching the underlying data with metadata detail (i.e., information on the type and genre of content consumers are engaging with), while simultaneously using content as a way to incentivize more usage of AOL's online services, thereby fueling the creation of more data. In this sense, AOL-Time Warner operated a pipeline of information on how consumers watch and interact with content online, effectively gathering data on what programming *works*. It is the access to data, the consolidation of information on the company's end-users (e.g., online users, film viewers, television watchers, magazine readers), the direct connection to the audience, that set the vision of AOL-Time Warner apart from its competitors at the time.

The legacy media industry has long been disconnected from the audience, working with intermediaries to distribute content across multiple channels, from film to television, home entertainment, and downstream channels (i.e., retail merchandising), in order to maximize revenue.<sup>69</sup> Even leading media conglomerates of the time, like Time Warner, Viacom, and The Walt Disney Company, which operated across multiple channels with considerable audience

---

<sup>69</sup> See, Wasko, 2003 and Basin, Ken, *The Business of Television* (New York: Routledge, 2018).

reach and a global presence, invariably leveraged their scale to manage content output rather than information input. Effectively, legacy media companies focused on pushing out content rather than pulling in the available data across the digital landscape, which made the media industry a proverbial black box. Indeed, legacy media companies have long operated as a B2B (i.e., business-to-business) industry, part of a larger supply chain of intermediaries, relying on partners to deliver content directly to consumers. Technology companies like AOL, by contrast, have consistently cultivated a B2C approach (i.e., business-to-consumer), delivering their products directly to consumers, rather than relying on intermediaries, thereby deploying their digital scale to collect, consolidate, and commercialize information on consumers. In this sense, AOL-Time Warner signaled a model where legacy media companies could build on the emerging digital landscape to integrate into the evolving information infrastructure, connect to a broader audience, and establish access to data.

AOL-Time Warner thus represents one of the first major *data access transactions* of the 21st century, a data-driven corporate convergence designed to create, consolidate, and capitalize information access. The company effectively positioned its convergence of media and technology capabilities as a data access infrastructure. Time Warner's content catalog would give AOL access to information on how consumers interact with film and television programming, informing the dynamics of cultural production (i.e., what types of content and programming should be made in order to ensure audiences would watch it). In turn, AOL's online presence would give Time Warner access to digital consumer information, offering a gateway to understand what consumers want to watch. Together, the company would effectively own access to digital consumer background and behavior data. In this sense, the vision of AOL-Time Warner reconfigured the new media industry from a black box into a glass box of

information, with data and information as the core driver of the company's business. In laying out the vision for AOL-Time Warner, Steve Case, AOL Chairman and newly elected Chief Executive at the time, seized upon this narrative of innovation through information (my emphasis): "AOL Time Warner will be a **laboratory** where we find the right **formulas** for pioneering **transformations** of industries."<sup>70</sup>

The Wall Street Journal vividly sketched the industrial reality of this vision, a digital world where AOL-Time Warner administers, captures, and records all consumer activity:

"This is a glimpse of the home of the future: Mom, in the den, logs onto AOL from a flat-panel TV to download a Warner Bros. movie. Dad, in the kitchen, uses an AOL information appliance to find a recipe from Time Inc.'s Cooking Light magazine. Their daughter is in her bedroom reading a Time Warner e-book on a handheld tablet while gabbing on a Web-enabled cell phone. And all of them are zapping instant messages to each other as they discuss which HBO show they will watch that night."<sup>71</sup>

In this conceptual model of seamless synergy,<sup>72</sup> AOL-Time Warner would be able to access data on consumer interaction with content across multiple different contexts, generating a wealth of information that, theoretically, could be monetized in two principal ways. One, by keeping consumers in the digital ecosystem, increasing usage of AOL-Time Warner products and services and thereby selling more ads and subscriptions to an engaged audience. Or, two, by

---

<sup>70</sup> Julia Angwin & Martin Peers, "AOL-Time Warner Megamerger Creates a Web, Media Behemoth," *Wall Street Journal*, December 15, 2000, accessed June 5, 2020, <https://www.wsj.com/articles/SB976835137406328524>

<sup>71</sup> Ibid

<sup>72</sup> One of the first realizations of this vision can arguably be found in *The Lord of the Rings* (Warner Bros., 2001-2003) franchise, which pioneered an early online ecosystem of digital content and experiences. See, Thompson, Kristin, *The Frodo Franchise: The Lord of the Rings and Modern Hollywood* (Berkeley: University of California Press, 2007) and Holson, Laura M. & Lyman, Rick, "In Warner Bros.' Strategy, A Movie Is Now a Product Line," *New York Times*, February 11, 2002, accessed June 6, 2020, <https://www.nytimes.com/2002/02/11/business/warner-brothers-strategy-movie-now-product-line-making-franchise-films-that.html>

creating new products and services, such as digital entertainment, to drive incremental growth across ads and subscriptions. In either case, AOL-Time Warner would rely on data access to drive their business forward. For the company's executive leadership, this vision was the beginning of the next Internet revolution, an era wherein data access represents a substantial form of capital for new media businesses. "We're still just scratching the surface," Steve Case noted upon the announcement of the merger, signaling the downstream impact of this new information infrastructure.<sup>73</sup>

The AOL-Time Warner merger proved to be neither a black box nor a glass box, but Pandora's Box, with unintended consequences to the company's bottom-line and lasting implications for the entire digital media industry. The merger resulted in significant valuation loss, divisional spin-offs, executive departures, and talent turnover. Scholars, pundits, and former executives have cited a wide range of reasons for the implosion, from cultural clashes between media and tech workforces (i.e., opposing ways to envision, execute, and operate the business) to economic headwinds (i.e., the lasting effects of the dot-com crash in the early 2000s) and a lack of technological innovation (i.e., AOL's slow shift from dial-up to broadband Internet technologies). Steve Case, in later reflections, acknowledged the laundry list of challenges, yet also insisted that the merger's unraveling was primarily a result of failed execution.<sup>74</sup> He argued that the industrial vision of AOL-Time Warner, the strategy to build an information

---

<sup>73</sup> Press Release, Time Warner

<sup>74</sup> See, Gunther McGrath, Rita, "15 years later, lessons from the failed AOL-Time Warner merger," *Fortune*, January 10, 2015, accessed June 5, 2020, <https://fortune.com/2015/01/10/15-years-later-lessons-from-the-failed-aol-time-warner-merger/>

infrastructure to access data as a core driver of business capital, has come to define the core tenets of the digital media industry over the course of the following decade.<sup>75</sup>

Indeed, AOL-Time Warner laid the foundation for a century-defining business model that was subsequently adopted, refined, and mainstreamed by a new suite of tech giants, rising from the AOL-Time Warner crater. Platforms like Facebook, Amazon, Apple, Netflix, and Google (i.e., FAANG) built on an ever-expanding digital infrastructure that allowed consumers to connect with ever-growing repositories of online content and services, capturing, recording, and storing the resulting information as data to drive their business. Platforms established an ecosystem to consistently track consumer information, utilizing their data access to iterate, modify, and launch digital products and services, tailoring, customizing, and personalizing their offerings to ensure consumers stay engaged, consistently locked into their digital ecosystems, thereby creating a scalable revenue generator.

As such, platforms effectively created, consolidated, and capitalized a direct connection with the audience. While their paths varied, they followed the principal steps of a homegrown Silicon Valley playbook: *raise and burn cash to grow and monetize a user base*. Platforms initially used external funding and cash infusion from venture capital firms (VCs) and private investors to build and scale their digital ecosystems, diversify their product and services portfolios, and attract and grow an audience of millions of users, thereby establishing a working information infrastructure with access to massive amounts of data. Platforms then implemented business models to monetize the access to data, initially selling advertising with advanced targeting capabilities (i.e., allowing advertisers to plug into their data access and reach specific

---

<sup>75</sup> Indeed, AT&T acquired Time Warner in 2018 in a strategic play to leverage data access for advanced content and advertising commercialization. See, Aaron Pressman & Adam Lashinsky, “Data Sheet—Why AT&T’s Time Warner Deal Brings the Future of Television Closer,” *Fortune*, June 13, 2018, accessed June 5, 2020, <https://fortune.com/2018/06/13/data-sheet-att-time-warner-antitrust-analysis/>

audience segments not available through legacy media channels) and later adding subscription models to ensure recurring revenue (i.e., using data access to tailor products and services, including content, to ensure consumers stay engaged and subscribed). Platforms thus leveraged data access as an existential form of capital, drawing on their information infrastructure as the cornerstone of their business.

Data access as capital has become a central component of the new digital media environment, driven by a suite of cultural, economic, and technological developments, including:

- *Internet Commercialization*: Following the dot-com crash, the Internet re-emerged as a global economic infrastructure, consistently expanding access across the globe through the maturation of broadband and mobile technologies. Consumers and businesses migrated online, and social and commercial activities were re-encoded into the developing digital landscape. At the beginning of the 21st century, about 120 million Americans were online (42%). Two decades later, the number had risen to over 312 million (90%). At the same time, in 2000, Americans spent an average of 9.4 hours per week online. By 2018, that number had grown to 23.6 hours.<sup>76</sup>
- *Data Economics*: Companies that capitalize data have become a crucial factor of the economy, with nearly all representatives of the S&P 500 citing data as a key component of their business over the 21st century. The FAANG consortium has consistently outperformed the S&P 500, with data playing a critical role in the platform companies' rise, expansion, and consolidation.

---

<sup>76</sup> Condliffe, Jamie, "The average American spends 24 hours a week online," *MIT Technology Review*, January 23, 2018, accessed June 5, 2020, <https://www.technologyreview.com/f/610045/the-average-american-spends-24-hours-a-week-online/>

- *Digital Technologies*: The rise, expansion, and maturation of new digital consumer technologies (e.g., the smartphone, tablets, digital video players) has introduced an everyday network where data is consistently generated, collected, and stored, yielding an exponential output of data.
- *Cloud Computing*: The commercialization of cloud computing has enabled the development and expansion of enterprise-level technologies that generate, collect, and store data at scale.<sup>77</sup>

Platforms seized upon the new media environment to build proprietary data access, effectively owning all consumer information across the digital landscape. In this environment, Google has become synonymous with web and search data, Facebook with social data, Apple with mobile data, Amazon with e-commerce data, YouTube with online video data, and Netflix with OTT streaming data. While all these platforms have consistently captured, consolidated, and commercialized a spectrum of information, part of their global brand identity and business capital is built on access to specific data types.

Yet, as AOL-Time Warner has demonstrated, the concept of data access as capital has its limitations and is not a fail-safe strategy. The digital landscape is filled with high-profile cases of failed platform models, from early pioneers like Yahoo (1994) and Friendster (2002) to emerging players like MySpace (2003).<sup>78</sup> These companies had massive information infrastructures and owned access to data on millions of consumers, yet still were unable to capitalize on it with

---

<sup>77</sup> Examples of cloud computing services include Amazon Web Services (launched in 2006), Google Cloud Platform (launched in 2008), and Microsoft Azure (launched in 2010).

<sup>78</sup> See, Cusumano, Michael A., Gawer, Annabelle, & Yoffie, David B., "A Study of More Than 250 Platforms Reveals Why Most Fail," *Harvard Business Review*, May 29, 2019, accessed June 5, 2020, <https://hbr.org/2019/05/a-study-of-more-than-250-platforms-reveals-why-most-fail>



lasting success. Reasons for this development are manifold, ranging from simple launch timing to the quality of the technological product and experience and the economic resources of each company as well as macro-economic and socio-cultural developments, which illustrates that data access is but one of many factors that drive the dynamics of the digital media economy. Still, data access arguably constitutes a crucial driver of competitive differentiation, enabling platforms, at least conceptually, to anticipate, iterate, and implement change.

The dynamic of social media platforms Friendster and Facebook offers an enlightening use case in this context. Upon launch, both platforms displayed fairly similar product features and services, enabling the creation of digital profiles and online chat. Friendster launched a year ahead of Facebook and initially generated more growth, allowing the company to scale its data access. Yet, Facebook introduced a new social feature called *News Feed*, a landing page that automatically updates what each individual user sees based on the behavioral data from their friends list, which helped the company significantly grow engagement and expand its audience by drawing users away from Friendster. Facebook did not only own data access, it actively leveraged access to consumer information in order to drive and evolve its business. News Feed was a product engineered through data access and designed to expand data access through more users and increased engagement. In this regard, data access is not a static entity, but a dynamic process that derives capital from constant iteration and innovation.

As the digital economy became embedded into all industrial sectors, including media and entertainment, the industrial development of data access as capital posed a significant shift for the legacy media industry, for several reasons. Hollywood has long operated as a B2B economy, emphasizing partnerships with intermediaries over direct consumer relationships, focusing on externally-controlled rather than internally-managed content distribution. As a result, the

industry has focused on pushing out content over pulling in data. Furthermore, the reliance on distribution intermediaries has effectively disintermediated film studios and television networks from having a direct connection to their audience. Theaters have maintained the direct relationship with moviegoers, cable companies owned the connection to television viewers, and retailers managed the physical and digital transactions of home entertainment and film and television merchandise. In this sense, the majors did not have an operational infrastructure of information access.

While the legacy media industry lacked proprietary access to information, the majors have long managed external mechanisms to access data. From the late 1920s and early 1930s on, the industry has contracted with third-party research vendors to gather audience research, pulling in consumer and usage data from focus groups, surveys, and questionnaires.<sup>79</sup> In this sense, the majors have consistently maintained an external infrastructure of information, albeit largely grounded in the aggregate statistics of surveys and polls rather than the expansive flow of digital data. The rise of data access as capital has put this outsourcing model in stark relief. Indeed, the majors' reliance on outside resources and representative samples over a direct connection with the audience has effectively made studios and networks "data broke" and "data poor" in the digital landscape, thus unable to access information on their customers.<sup>80</sup> Notably, upon the release of *Star Wars: The Force Awakens* (Disney, 2015), arguably one of the modern century's most successful global film releases, Walt Disney Company's former CEO Bob Iger acknowledged the lack of proprietary, first-party data access due to the company's reliance on

---

<sup>79</sup> See, Wasko 2003

<sup>80</sup> See, Think with Google, "Lights! Camera! Data! How insights help 20th Century Fox Film reach the right audiences," *Google*, August 2018, accessed June 5, 2020, <https://www.thinkwithgoogle.com/marketing-resources/data-measurement/data-insights-film-marketing/>

distributors: “We don’t have any idea who went to see *Star Wars* in the cinemas [...]. [O]ur access to the consumer is very limited.”<sup>81</sup>

Yet, contrary to wide-ranging claims,<sup>82</sup> Hollywood was not caught in a constant state of disintermediation, where the majors were eternally disconnected from audience information, unable to own data access in the digital media economy. Indeed, the majors have actively negotiated varying stages of access, building on existing industrial tactics, adopting new strategies, and adapting to external developments in the shifting digital media landscape. Consequently, they have increasingly integrated with the new digital information infrastructure to manage access. Hollywood’s integration of data access thus constitutes a complex course of action, marked by proactive decision-making and reactions to external factors, with the industry working to coordinate, consolidate, and control multiple layers of access.

Accordingly, this chapter examines Hollywood’s negotiation of three interrelated stages of data access and their role in establishing the industry’s new information infrastructure over the course of the 21st century. All stages represent complicated and challenging dynamics, involving a suite of industrial strategies and developments. As such, they operate as interconnected and interdependent phenomena rather than sequential processes with specific time stamps.

- *Data Generation*: The major film studios and television networks have taken active steps to distribute their content across the digital landscape, experimenting with a variety of programming output to reach digital audiences. By releasing content across platforms, the

---

<sup>81</sup> Hazelton, John, “Big film, big data: how analytics is shaping the business,” *ScreenDaily*, June 16, 2016, accessed June 5, 2020, <https://www.screendaily.com/features/big-film-big-data-how-analytics-is-shaping-the-business/5104922.article>

<sup>82</sup> For a summary on Hollywood’s lack of data access, see, Mishra, Devendra, “Hollywood’s Digital Blind Spots,” *Graziadio Business Review*, Volume 18, Issue 2, 2015.

majors have fueled an environment that consistently generates data on the media audience and its interaction with film and television content, effectively making specific media and entertainment data available in the digital economy. As a result, the majors have contributed to an ever-pulsating stream of data, tantamount to an active, yet complicated and layered information access pipeline.

- *Data Licensing*: The majors have developed a suite of partnerships with technology companies to license data access. These partnership deals have enabled studios and networks to access substantial amounts of consumer and usage data, making information more usable within legacy media organizations. At the same time, they introduced complex dynamics of access, including limitations around what information is made available and to what extent.
  
- *Data Ownership*: The legacy media industry has gradually explored, experimented, and institutionalized the business logic platforms in an effort to own data access, an industrial development that has made data more actionable for the majors, directly affecting their business model. Specifically, the majors have pursued two models of ownership:
  - *Data Acquisitions*: The majors made strategic investments and acquired specific digital media and technology companies to establish data access capabilities from the top down.
  
  - *Data Incubation*: The majors launched their own platforms to build their own information infrastructure and establish data access from the ground up.

Collectively, these stages illustrate how Hollywood has consistently negotiated the dynamic and complex process of data access, managing multiple layers of the information infrastructure to establish an updated operating model of access for the digital era.

### **Pipeline Building: From Data Generation to Fragmentation**

In the late 1990s, Daniel Myrick and Eduardo Sanchez, a team of budding documentary filmmakers, created a website to promote their independent debut feature about a group of kids getting lost in a Maryland forest in search for a mysterious mythical figure. The website's design and graphical user interface (GUI) featured a simple black background, displaying the title in sketched letters surrounded by glowing flashes of light. The mid-section of the page displayed four hyperlinks, opening up on a small set of paratexts, such as behind-the-scenes photos, audio files, and video footage, providing cryptic background information on the characters, setting, and premise, positioning the film as a found-footage documentary based on real events (figure 1.6).<sup>83</sup>



**Figure 1.6:** The landing page of *The Blair Witch Project* website featured a set of hyperlinks that led to a string of paratexts functioning as an online transmedia extension of the film's fictional narrative.

<sup>83</sup> See, Bereznak, Alyssa, "The 'Blair Witch' Extended Universe: How a Tiny Indie Film Became a Horror Sensation—and Invented Modern Movie Marketing," *The Ringer*, March 28, 2019, accessed June 6, 2020, <https://www.theringer.com/movies/2019/3/28/18280988/blair-witch-movie-marketing-1999>

<sup>84</sup> See *The Blair Witch Project* website, accessed June 5, 2020, <https://www.blairwitch.com/project/main.html>

The site was built on Flash, a computer developer software that enabled an early-stage multimedia and interactive experience at relatively low cost. When the website went live, it drew a base of about 10,000 users, many of whom reportedly scanned the content, shared their impressions on the site's online bulletin board, and signed up for the newsletter to keep up-to-date on future developments. It was a simple, small-scale grassroots marketing initiative designed to showcase a work-in-progress, a collection of digital promotional content, and an early instance of the Internet acting as a digital marketing and distribution channel. In effect, it was one of many movie websites at the time, a growing, yet still relatively small trend in the industry.<sup>85</sup> Then, the filmmakers premiered the film at the Sundance Film Festival.

*The Blair Witch Project* (Artisan, 1999) received a vibrant reception at the festival, generating solid reviews and transcending word-of-mouth. The festival buzz extended online, with the film's website becoming a massive discursive catalyst, reportedly increasing its user base by 10x to 100,000 and setting off one of the most vital media discourses in the history of film. Fans engaged in spirited online conversations, exchanging theories about the film's premise, characters, and ending, debating its status as a record of real events, and tracking any new information released by the filmmakers across the site's message board. The website activity became a leading indicator for the film's cultural buzz, channeling demand among young audiences, and convincing the studio, Artisan Entertainment, to implement a platform release and roadshow the film across the country. *The Blair Witch Project* became a cultural

---

<sup>85</sup> See, Malinowski, Erik, "'Space Jam' Forever: The Website That Wouldn't Die," *Rolling Stone*, August 19, 2015, accessed June 5, 2020, <https://www.rollingstone.com/movies/movie-news/space-jam-forever-the-website-that-wouldnt-die-70507/> and Hutsko, Joe, "Behind the Scenes Via Movie Web Sites," *New York Times*, July 10, 2003, accessed June 6, 2020, <https://www.nytimes.com/2003/07/10/technology/behind-the-scenes-via-movie-web-sites.html>

phenomenon, earning \$248 million at the box office against a \$60K budget,<sup>86</sup> spawning sequels and multimedia off-shoots, from books to comic books and video games, dominating the cultural conversation in online forums, and inspiring a new generation of filmmakers and films (e.g., Oren Peli and the *Paranormal Activity* franchise).<sup>87</sup>

*Blair Witch* is widely celebrated as a revolutionary piece of horror filmmaking, footnoted as a historical box office success story, and credited with patenting the tenets of modern digital marketing. It also constitutes a foundational element in Hollywood's data access infrastructure, illustrating how the legacy media industry's investment in digital content marketing and distribution created a system for the systematic generation of data. Indeed, the case of *The Blair Witch Project* offers insight into how content input became a driver of data output.

*The Blair Witch Project* website effectively functioned as an early "audience feedback mechanism,"<sup>88</sup> tracking, translating, and quantifying audience activity into consumer and usage data (figures 1.7-1.8).

- *Consumer Data*: Consumer data represents background information on the audience. It is largely qualitative information and includes demographics (i.e., age, gender, and location) as well as comments (i.e., feedback on the movie's plot, setting, and characters in form of comments on the site's message board). Consumer data is typically provided when users sign up for a platform to provide information for an initial user profile.

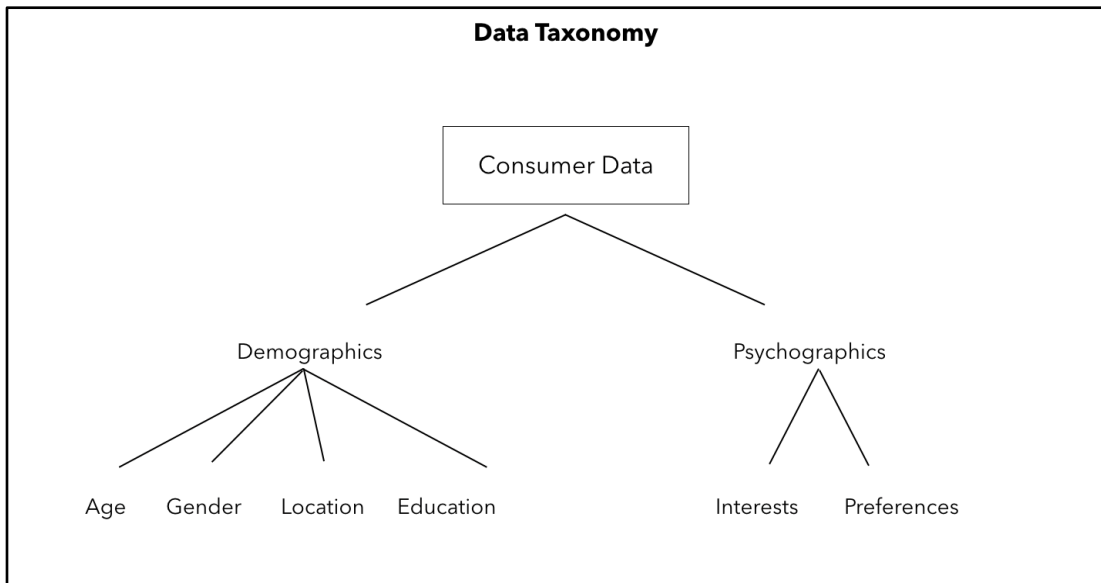
---

<sup>86</sup> See, Interview Series (conducted by Phil Hoad), "How we made The Blair Witch Project," *The Guardian*, May 21, 2018, accessed June 5, 2020, <https://www.theguardian.com/culture/2018/may/21/how-we-made-the-blair-witch-project>

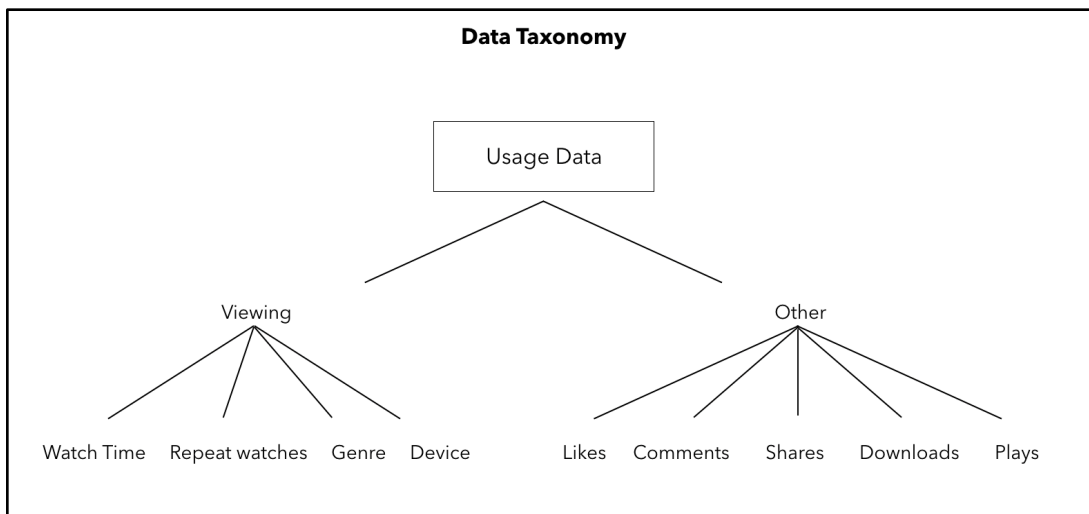
<sup>87</sup> See, Crucchiola, Jordan, "Charting *The Blair Witch Project*'s Influence Through 10 Horror Films That Followed," *Vulture*, September 16, 2016, accessed June 5, 2020, <https://www.vulture.com/2016/09/10-horror-movies-inspired-by-the-blair-witch-project.html>

<sup>88</sup> Napoli, 2010

- *Usage Data*: Usage data represents information on user behavior and the interaction with digital content. It is largely quantitative information and includes the number of website visits, clicks, video views, and comments on a website.



**Figure 1.7:** Consumer data includes information on the background of the audience. This information is generally captured up-front, during the sign-up process for an online product service, and through ongoing product use.



**Figure 1.8:** Usage data details the behavioral actions and patterns of users in their interaction with digital content.



Effectively, the website transformed digital consumer activity (i.e., viewing videos, clicking on hyperlinks, commenting on the message board) into measurable information. Put differently, it quantified consumer interaction with content (*input*) into new strands of digital information (*output*).

As such, the website operated as a data access enablement tool, generating information on the audience (i.e., consumer data) and their interaction with content (i.e., usage data), effectively translating audience feedback into measurable, quantifiable results. The site's collection of textual, audio, and video paratexts inspired audience activity, which, in turn, generated more data. The resulting information had a lasting business impact for the film. Access to the website data reportedly gave Artisan Entertainment an opportunity to “gauge [...] the extent of potential viewer interest”<sup>89</sup> and informed the studio's decision to platform the film release, capitalizing on the film's online buzz over time.

It should be noted that executives at the time arguably did not receive multi-layered spreadsheets or data dashboards containing volumes of information given the limitations of data infrastructures at the time. Indeed, the website's data was limited to the number of visitors and comments left on the message board, offering but a glimpse of the website's activity and popularity. Nevertheless, the website's numbers-display, showcasing word-of-mouth and engagement, provided direct access to data on digital consumer activity, quantifying consumer visits and comments in an unprecedented way. In effect, the website made available information that had previously been unknown to the major studios and networks, which had only processed audience information through aggregate statistics and representative samples in external research

---

<sup>89</sup> Telotte, J. P., “The *Blair Witch Project* Project: Film and the Internet,” *Film Quarterly*, 54 (3), 2001.

reports, compiled from surveys and polls, rather than tracing and measuring direct audience activity across the digital landscape.<sup>90</sup>

*The Blair Witch Project* thus marks a catalyst in Hollywood's systematic effort to leverage online channels as audience feedback mechanisms to market and promote content in the growing digital landscape.<sup>91</sup> As audiences migrated online, the legacy media industry consistently invested in digital marketing initiatives across an array of platforms to promote their content and programming to digital audiences.<sup>92</sup> The holistic, industry-wide investment in digital marketing created a foundation for the constant generation of data on the film and television industries. The rise of media paratexts<sup>93</sup> across digital platforms drew engagement from audiences, clicking, sharing, commenting, and interacting with films and television shows in an online environment, which consistently translated the input of audience activity into data output, thereby yielding industry-specific consumer and usage data. In other words, every digital marketing campaign generated data on the audience and their interaction with content. As such, Hollywood's digital marketing turn laid the foundation for a pulsating stream of data generation, a rich pipeline of digital information.

---

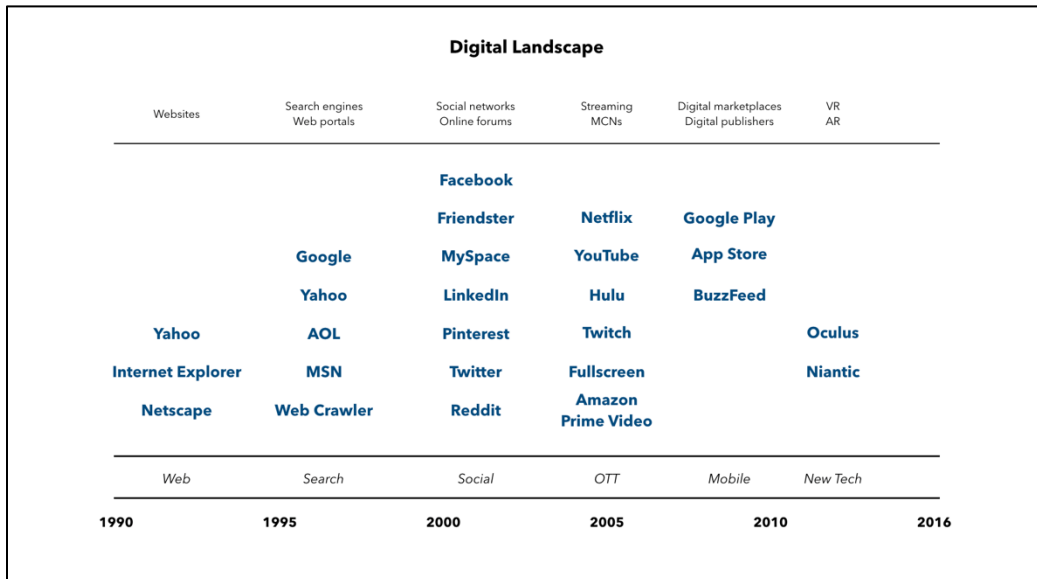
<sup>90</sup> This is not to imply that the studio and filmmakers actively looked at data from the website. Rather, it is intended to showcase the growing availability of data, in addition to traditional audience research.

<sup>91</sup> For an overview of Hollywood's growing use of digital marketing channels, see Grainge, Paul, *Brand Hollywood: Selling Entertainment in a Global Media Age* (New York: Routledge, 2007) and Wasko, Janet, *Hollywood in the Information Age: Beyond the Silver Screen* (Cambridge, UK: Polity, 2013).

<sup>92</sup> See, Balio, Tino, *Hollywood in the New Millennium* (London: British Film Institute, 2013) and Balio, Tino, "Adjusting to the New Global Economy: Hollywood in the 1990s." In: *Film Policy: International National, and Regional Perspectives*, edited by Albert Moran (New York: Routledge, 1996).

<sup>93</sup> See, Gray, Jonathan, *Show Sold Separately: Promos, Spoilers, and Other Media Paratexts* (New York: New York University Press, 2010).

Effectively, by investing in digital marketing across an ever-expanding set of platforms, the majors built a generative data pipeline on top of the digital landscape, a procedural feedback system that turned content and audience input into data output, capturing, recording, and storing increasing amounts of information on the legacy media business across an expanding suite of audience feedback mechanisms. Indeed, as the digital landscape developed, adding new platforms, expanding from websites to include search engines and web portals, social networks and online forums, video streaming services, digital marketplaces, and emerging technologies,<sup>94</sup> the data pipeline came to generate a growing volume (i.e., the number of consumers interactions) and variety (i.e., the type of data generated, such as web data, social data, or video data) of information (figure 1.9).



**Figure 1.9:** The digital landscape saw the emergence of digital channels that captured audience feedback, from websites to search, social media, online video, mobile games & apps, and emerging technologies.

<sup>94</sup> I define emerging technologies as a suite of digital products, tools, and experiences that had not yet reached a mainstream audience (e.g., virtual reality and augmented reality).

Over the course of the 21st century, the majors seized upon an ever-expanding suite of digital platforms to market their content, dynamically adjusting to new stages of the digital landscape, expanding from web to search, social, OTT/streaming, mobile media, and new technologies, such as virtual reality (VR) and augmented reality (AR).<sup>95</sup> While the digital landscape underwent constant technological evolution, Hollywood's data pipeline functioned as a consistent generative response mechanism, turning increased content input into ever-growing and varied data output. As such, Hollywood has consistently generated data across all stages of the digital landscape.

In this regard, Hollywood's digital pipeline created the foundation for an information infrastructure that captures, records, and stores consumer and usage data across platforms. The use of platforms as audience feedback mechanisms established a consistent stream of data output, which effectively made data available within the media industries. While studios and networks may not have actively considered, registered, or looked at the information surplus, they now operated in an environment where data was always actively generated, flowing through the digital landscape.

The majors primarily deployed platforms as channel partners to drive marketing and distribution of their content across the digital landscape, yet the use of platforms effectively established a new information infrastructure where data on the interaction between audiences and content is consistently generated.

The output of Hollywood's data pipeline has become increasingly complex and complicated, yielding an intricate web of information, for several reasons. First, Hollywood's

---

<sup>95</sup> Sample investments from the legacy media industry include film and television websites, Google search cards for new movies and television shows, social media pages, as well as film and television-themed mobile apps and games. See, *The Economist*, "Hollywood and its audience look to the Internet," January 11, 2010, accessed June 6, 2020, <https://www.economist.com/news/2010/01/11/a-happy-ending>

data pipeline operated across a broad range of platforms with differentiated data infrastructures. While the data generated across digital platforms share a set of consistent elements (i.e., consumer demographics are generally captured across all platforms), platforms also generate different types of data.<sup>96</sup> For example, social media platforms capture increasing levels of detail in consumer data, such as relationship status, favorite movies/television shows, and professional/educational information, while platforms specializing in mobile media emphasize usage data focused on mobile devices (e.g., laptops, smartphones, tablets). Similarly, while most platforms developed video capabilities, streaming platforms such as Netflix capture more complex layers of video viewing data than websites (e.g., Netflix does not only capture video views, starts, and sessions, which is common for digital video players, but records watch time, session time, skips, and other factors of the viewing experience). As such, Hollywood’s data pipeline generated data that differed in volume and variety by platform (table 1.2).

**Table 1.2**

	<b>Data Volume</b>	<b>Data Variety</b>
<b>Description</b>	Volume refers to the quantity of data output, specifically the number of behavioral interactions between audiences and content	Variety refers to the quality of data output, specifically data types
<b>Examples</b>	The number of clicks, video views, comments, shares, etc.	Web data, search data social data, mobile data, online video data, OTT data

Second, while Hollywood’s data pipeline generated data across all platforms, the resulting information infrastructure did not constitute a holistic data pool. Rather, it represented a stream of data that flowed through multiple platform rivers across an ocean of information. Put differently, the data that the majors generated through digital marketing initiatives across

---

<sup>96</sup> Indeed, platforms follow specific data policies, prioritizing specific data types. For example, Netflix has consistently emphasized video viewing data over other types of information (e.g., demographics).

platforms was not captured, recorded, and stored in one central database. In effect, the data was invariably captured, recorded, and stored where it was generated, on individual platforms. As such, data on the film and television industry was inherently fragmented across platforms.

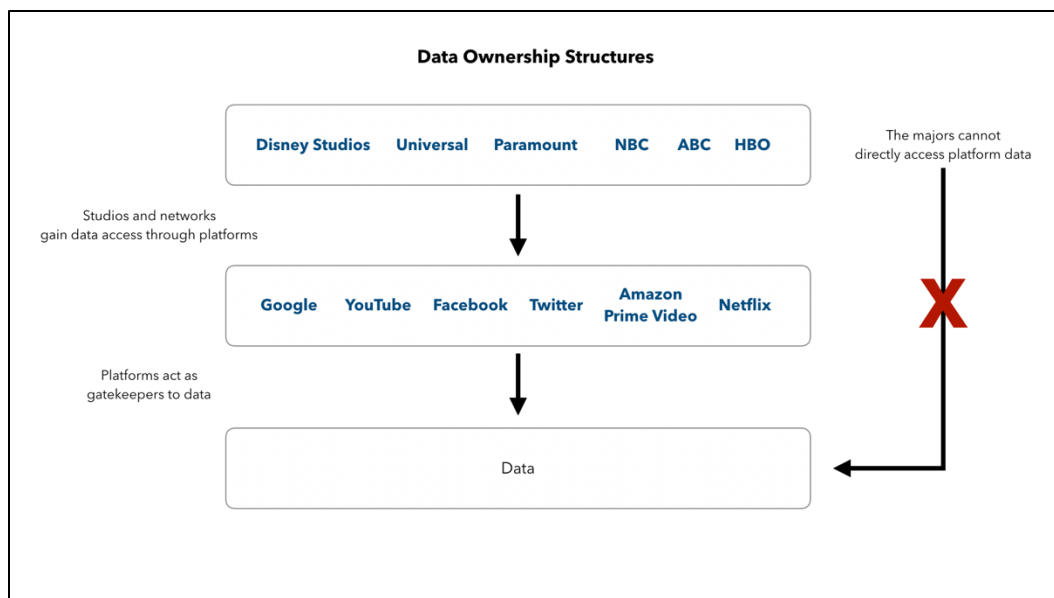
And, third, while Hollywood's data pipeline generated information across the digital landscape and thus made data available within the legacy media industry, the majors did not have universal access to the information infrastructure, which remained fragmented across platforms. Indeed, platforms did not only capture, record, and store the data generated by the majors, they also effectively owned it. As a result, the data generated by the majors was available across the digital landscape, yet not directly usable by the majors, meaning they could not seamlessly access and use the data within their organizational structures.

Admittedly, this notion of data access is not a hard-and-fast rule as data ownership structures are complex, nuanced, and constantly evolving. For example, considering websites, the majors could access and use web data since websites are built on web servers, and thus not owned by platforms, which enabled seamless access to the underlying web data layer for website owners. In the case of *The Blair Witch Project*, Artisan Entertainment could thus theoretically access the underlying data generated through the website given that it owned the URL (though, at the time, access to data was less feasible and seamless as it would be several years later with the emergence of data analysis software tools). As such, the studio had access to all the data generated on the website. The website effectively gave Artisan access to first-party data, the entire data generated on the website from visitors and their interaction with content.<sup>97</sup>

---

<sup>97</sup> At the time, data access software was not yet a mainstream phenomenon. Analytics tools like Google Analytics launched in the mid-2000s. As such, automated data access, from a technological perspective, was not yet a turnkey practice. In this regard, I present access as more theoretical and conceptual, rather than practical.

By contrast, the 2016 sequel, titled *Blair Witch* (Lionsgate, 2016), illustrates a more complicated data access infrastructure, reflecting the shifting digital media landscape and the proliferation of platform-based audience feedback mechanisms. The sequel expanded its marketing presence beyond the original website to various other platforms, including a Facebook page. While the Facebook page generated more data volume and variety given the scale of its multi-million user base, data access proved more limited. Indeed, Lionsgate, the new studio, only had access to a limited amount of data generated on the Facebook page because Facebook, the platform owner, kept the majority of its data proprietary, only making a small set of data available to platform users. To this end, studios and networks are only able to access data that platforms make available, but the majority of the information remains controlled by the platforms themselves (figure 1.10).



**Figure 1.10:** Platforms own and control their data – the data is not directly accessible to the majors.

In this regard, Hollywood’s data access pipeline has made data available within the industry by constantly generating information output from content input. At the same time, the

resulting data is fragmented across platforms, with usability (i.e., the use of the information within organizational structures) governed by complex data ownership structures. In this regard, the formation of Hollywood's data pipeline required the majors to navigate a complex set of industrial circumstances and relationships in the digital landscape, adapting to the dynamics of audience feedback mechanisms and their data ownership structures, specifically websites, platforms, and emerging technologies.

### Websites and Pipeline Construction

As the first digital marketing and promotional tool, the website represents the prototypical data generator in Hollywood's evolving data access pipeline and information infrastructure. Indeed, the industrial development of the website has fundamentally shaped the legacy media industry's data output, due to a variety of factors:

- *Audience Scale:* At the time of *The Blair Witch Project*, only 23% of U.S. citizens regularly visited the Internet.<sup>98</sup> Two decades later, the number of active Internet users has grown to 90% of the domestic population. The usage and proliferation of websites has grown exponentially, driven by the commercialization of email clients (e.g., Apple Mail, AOL Mail, Gmail), instant messaging software (e.g., ICQ, AOL Instant Messenger, Google Talk/Chat), and search engines (e.g., Yahoo, Google, Bing), which helped establish websites as a mainstream mechanism to structure and navigate the digital landscape. While website usage has shifted due to the introduction of new digital communications platforms, including the shift to a mobile web, the website has remained

---

<sup>98</sup> Pew Research, "The Internet News Audience Goes Ordinary," January 14, 1999, accessed June 5, 2020, <https://www.people-press.org/1999/01/14/the-internet-news-audience-goes-ordinary/>



a central component of digital consumer activity. As such, website usage has consistently generated data output.

- *Production Capabilities*: Early film and television websites were assembled by small, nimble teams of IT enthusiasts and burgeoning professionals within large corporate structures focused largely on legacy marketing efforts (i.e., television, radio, publishing, out-of-home) rather than the developing digital landscape. The website of *The Blair Witch Project* was reportedly put together by the filmmakers, while the website for the contemporary studio blockbuster *Space Jam* (1996, Warner Bros.) was designed and built by a group of “five outcasts,”<sup>99</sup> early digital marketers operating independently, outside of the corporate eye, since senior leadership did not yet prioritize the Internet as a marketing vehicle. Over the course of the decade, the majors have increased their digital marketing budgets exponentially, creating full-scale departments with expanded capabilities and massive budgets to scale web production and reach audience segments beyond legacy media channels. The increase in websites fueled the growing digital activity among consumers, which, in turn, led to an increase in data generation.
- *Online Real Estate*: Beginning in the late 1990s and early 2000s, and in part spurred by the success of *The Blair Witch Project*, virtually every film and television show (as well as their promotional content on DVDs and Blu-Rays) promoted a website tie-in, creating a cornucopia of web pages, filled with an expanding suite of paratexts, from behind-the-scenes photos, videos, and diaries to trailers and games, as well as interactive technologies, including e-commerce functionality (i.e., *click here to buy the movie*

---

<sup>99</sup> Malinowski, 2015

*soundtrack*).<sup>100</sup> The rise in website presence and features worked to increase consumer engagement with digital film and television content, yielding more data generation.

- *Iterative Functionality*: The website proved a dynamic, iterative tool for the legacy media industry, regularly adding new features that enabled incremental ways to market and promote content. While early websites featured a mix of text, hyperlinks, and Flash-based audio and video, the ongoing modification of markup and scripting languages gave rise to enhanced web functionality, including mobile integration and cross-platform applications, as well as improved interactivity via online content sharing, discussion forums, and high-definition video. The majors accordingly standardized the website as the central nexus of the digital landscape, acting as the gateway to a wide range of digital platforms, the spoke-hub of Hollywood's digital marketing approach. For example, the *Jurassic World* (Universal Pictures, 2015) website linked to social media, mobile apps, games, e-commerce shops, and ticketing sales. Websites were thus positioned as the central organism of digital consumer activity, which set up constant data generation.

The website's increase in digital presence, features, and audience reach marked the foundation of the legacy media industry's generative data pipeline, persistently generating consumer and usage data at increasing scale. As open-access systems, distributed across a range of web servers, websites were fully owned by Hollywood, enabling the majors to generate and own first-party data. As such, film studios and television networks have maintained full ownership over their website data, which has given them direct access to audience information, such as consumer data

---

<sup>100</sup> See, Heritage, Stuart, "You've Got Mail: the forgotten world of 90s movie websites," *The Guardian*, March 15, 2017, accessed June 5, 2020, <https://www.theguardian.com/film/filmblog/2017/mar/15/official-film-websites-youve-got-mail-jurassic-park-space-jam>

(e.g., email addresses, age, and location) as well as usage statistics (e.g., the number of websites visits and clicks, visit time, retention rate).

As the prototypical digital channel, the website further illustrates a universal layer of data access, inherent in all audience feedback mechanisms across the digital landscape. Digital channels and platforms generally incorporate two data access layers. As such, they make data available in two specific ways. The first layer, the *frontend*, contains visible data that is publicly available, such as website comments, Facebook page likes, or YouTube video views. The second layer, the *backend*, is composed of hidden data, information that is not publicly visible and available only to the data owner, including the recorded time users spent on a website, the number of page likes one user generated across their account, or the time spent watching a YouTube video. The frontend-backend dynamic represents a core principle of the data logic, reflecting the modularity and duality of data access: data is not a universally available entity, except for select data available on the frontend, but a corporate-industrial phenomenon governed by ownership structures reflected in the backend.<sup>101</sup>

In the case of the website for *The Blair Witch Project*, the frontend includes data that was visible on the web pages of the site, including comments on the message board and trailer views displayed in the site's web-based video player. The backend includes data that was not publicly available, literally captured on the technological platform backend, including the demographics of visitors, email addresses, and the number of visits, clicks, and comments.

Frontend data is limited and specific whereas backend data is expansive and broad. For example, frontend information includes the number of likes on a Facebook page, largely visible to platform users and external visitors. Meanwhile, the backend data of the Facebook page

---

<sup>101</sup> The terms *frontend* and *backend* are borrowed from the field of computer science. Frontend refers to the visual interactive layer of computers (user interface), while the backend describes the technological layer.

includes information on the users (e.g., age, location) as well as additional usage details (e.g., time spent, other pages liked). Furthermore, frontend data is generally curated, a selection of available information, while backend data is a messy mass of information, incorporating a variety of unstructured data points (i.e., the tip vs. the bottom of a funnel). Backend data can also represent an extension of frontend data, adding additional information to the information that is publicly available. For example, *The Blair Witch Project* website featured a visible visit counter on the main site, recording over 100,000 visits (frontend). The data backend of the website, meanwhile, contained information on the demographics of the visitors, how much time they spent on the site, and when they visited. Similarly, a web-based video player showed the number of times a video has been viewed via a view counter (frontend data), while the backend indicated the actual length of a view, the demographics of the viewer, and when the views occurred.

- *Frontend Data*: The frontend is the visual layer of the digital channel/platform and the visible data layer of the audience feedback mechanism. The frontend visually captures, records, and communicates data through data access mechanisms, including the bulletin board system (BBS), the visit counter / web ticker, forums, and communities.
- *Backend Data*: The computational backend captures, stores, and records data that is not publicly available, including audience demographics and expanded usage data. Access is limited to the data owner through a platform technological infrastructure and associated access mechanisms.

Data access on the frontend makes data available and usable within legacy media organizations, though usability is largely manual as frontend data is anchored in the digital layer of the

channel/platform (i.e., the user interface), and thus cannot be easily accessed and manipulated.<sup>102</sup> Data access on the backend, by contrast, has increased usability as the data is available outside of the digital interface and can be manipulated on the technological backend (i.e., it can be cleaned, structured, and organized in databases).

The frontend-backend framework further extends to the dynamics of Hollywood's data pipeline. Websites generate data through a roster of data access mechanisms that operate on the frontend and backend of the digital landscape. Data access mechanisms are interactive features of digital channels and platforms that translate online consumer activity into data. For instance, on the frontend, the *like button*, a staple of communications software popularized through the social media platform Facebook, generates data output by incentivizing consumers to engage with content through a click. Once a user clicks the like button, the frontend layer visually records the click, making data visible and publicly available. On the backend, a helpful example is the *web cookie*, an invisible piece of code embedded within the pages of a website, which tracks consumer activity across the Internet, following the user journey beyond a single website, generating data output for not one, but all the websites the user visits.

Hollywood's investment in websites<sup>103</sup> generated exponential amounts of data on both the frontend and backend. Yet, while data was widely available, flowing through the backend of the digital infrastructure, it has remained highly fragmented, complicating Hollywood's data access, for several reasons. First, while the majors each owned access to first-party data across their own websites, they did not have full access to the broader network of websites in the digital

---

<sup>102</sup> One way to capture frontend data involves the use of web crawling technology, which automatically crawls the web and collects frontend web data by indexing websites. See, Google, "How Search Organizes Information," accessed June 6, 2020, <https://www.google.com/search/howsearchworks/crawling-indexing/>

<sup>103</sup> See, Drake, Philip, "Distribution and Marketing in Contemporary Hollywood." In: *The Contemporary Hollywood Film Industry*, edited by Paul McDonald and Janet Wasko (Hoboken: Wiley-Blackwell, 2008).

landscape, including the websites of competitors as well as a burgeoning suite of cultural websites about film and television (e.g., the popular film and television site *Ain't It Cool News*). While they could theoretically access limited frontend data on those sites, manually searching the sites to gather information, the backend data was owned by other industry players. Second, while the majors generated data on the backend, they lacked the IT infrastructure and software capabilities to access the data, instead gradually building up data access capabilities over time, which proved a long-winding and complex process. And, third, the majors' primary access to website data was built on the frontend, requiring manual access and lacking scale, which made it difficult to implement a scalable approach to data access.<sup>104</sup>

### Platforms and Pipeline Expansion

The legacy media industry's investment in digital marketing initiatives became increasingly centered on platforms, for several reasons. First, platforms provided increasing audience reach through a centralized infrastructure, offering a direct connection to millions of users on a global scale. Websites, by contrast, were effectively de-centralized and thus lacked the same scale. Second, platforms enabled more functionality for the majors to interact with consumers, allowing for new ways to market and promote content. And, third, given their scale and functionality, platforms came to act as the epicenter of digital activity and the gateway to reach online users.

Platforms represent digital ecosystems that offer products and services to users and leverage the resulting data to optimize their offering and experience in order to keep users engaged. Indeed, platforms operate an economy that emphasizes the exponential collection, consolidation, and commercialization of data. The goal of platforms is to aggregate as many

---

<sup>104</sup> Some legacy media players effectively pursued to control the early web experience through strategic acquisitions. For example, The Walt Disney Company purchased full control of search engine Infoseek in 1999.

users as possible to create a massive pool of consumer and usage data. To achieve this goal, platforms have consistently pursued interconnected layers of scale by growing their user bases and increasing engagement with their ecosystems. Platforms' main strategy to grow their user bases has involved consolidation through strategic acquisitions of competitors and partners. For example, Facebook acquired photo-sharing platform Instagram and messaging platform WhatsApp, which, at the time of the respective transactions, had 50 million and 450 million monthly active users.<sup>105</sup> Platforms further increased engagement with their ecosystems by embedding their products and services into the larger digital landscape, making their functionality a core element of digital activity and interaction. Specifically, platforms made their products and services available to developers through software developer kits (SDKs) and application programming interfaces (APIs), enabling other companies to integrate with their ecosystems. For example, YouTube enabled websites and other platforms to integrate its video player functionality, making YouTube the de facto video player of the Internet.<sup>106</sup>

As a result of these strategic actions, platforms have consistently increased their access to first-party data, effectively owning the information on consumers and their interactions with content in the digital landscape. While platforms share a unified approach to data access, the platform landscape is marked by a wide spectrum of different platform types, with an evolving set of players embedded within each (table 1.3).

---

<sup>105</sup> Facebook made attempts to acquire other platforms in order to expand its audience base, including the failed acquisition of Snapchat. See, *Wall Street Journal*, "Snapchat Spurned \$3 Billion Acquisition Offer from Facebook," November 13, 2013, accessed June 5, 2020, <https://blogs.wsj.com/digits/2013/11/13/snapchat-spurned-3-billion-acquisition-offer-from-facebook/>

<sup>106</sup> See, Patrick, Vonderau, Pelle Snickars & Jean Burgess, eds., *The YouTube Reader* (Stockholm: National Library of Sweden, 2010).

**Table 1.3**

Platform Types				
Examples	Social	Video	Mobile	Digital/Other
	<i>Friendster</i> (2002)	<i>YouTube</i> (2005)	<i>Apple App Store</i> (2008)	<i>IMDb</i> (1990)
	<i>LinkedIn</i> (2003)	<i>Amazon Video Paid VOD</i> (2006)	<i>Google Play Store</i> (2008)	<i>Rotten Tomatoes</i> (1998)
	<i>MySpace</i> (2003)	<i>Amazon Video SVOD</i> (2011)	<i>Amazon Appstore</i> (2011)	<i>BuzzFeed</i> (2006)
	<i>Facebook</i> (2004)	<i>Apple Video PVOD</i> (2006)		
	<i>Twitter</i> (2006)	<i>Apple Video SVOD</i> (2019)		
	<i>Google+</i> (2007)	<i>Netflix</i> (2007)		
	<i>Tumblr</i> (2007)	<i>Facebook Watch</i> (2017)		
	<i>Instagram</i> (2010)			
	<i>Snapchat</i> (2011)			

Accordingly, platforms have access to different types of data, a rich canvas of information that includes social, video, mobile, and digital publishing data (table 1.4).

**Table 1.4**

Platform Data		
Platform Type	Data Type	Example Information
<i>Social</i>	Social data	Likes, affinities/interests
<i>Video</i>	Video and OTT data	Watch histories, video searches
<i>Mobile</i>	Mobile data	Smartphone usage, Game and app downloads mobile ecommerce sales
<i>Digital/Other</i>	Digital publishing data	Reading histories Film and television data (e.g., IMDb, RT) Vertical interests (film, television, etc.)



Hollywood's legacy media industry expanded its generative data access pipeline by creating and distributing increasing amounts of content across a range of platforms, in two principal ways. One, the majors established a systematic presence across all platforms through content licensing and distribution initiatives, including social media (i.e., Facebook pages, Twitter accounts, Tumblr posts, Snapchat profiles), video and OTT (i.e., distributing trailers on YouTube, creating marketing videos for distribution on Facebook, licensing content to Netflix, Amazon, and others), mobile apps and games (i.e., creating and licensing games and apps, such as second-screen experiences), and digital publishing (i.e., creating sponsored content for digital publishers like BuzzFeed). Two, they allocated increasing amounts of their marketing budget to digital advertising, incrementally expanding beyond legacy media channels (i.e., television, radio, out-of-home), promoting their content across platforms through a wide range of digital ad formats, ranging from Google search ads to Facebook carousel ads and YouTube banner ads.<sup>107</sup>

By investing in digital marketing across platforms, the majors significantly expanded the generative layer of their data access pipeline, generating exponential increases in data volume and variety, effectively creating information on their audiences and the interaction with content across social, video, mobile, and digital platforms. Yet, while the investment in platforms generated an increase in data on the film and television industries, it equally limited the majors' access to the resulting data.

Platforms cultivate an information infrastructure that emphasizes data ownership. As such, platforms own access to first-party data that is generated across their products and services. Some of the information is publicly visible, available on the frontend. Indeed, platforms have

---

<sup>107</sup> Digital advertising spend has grown incrementally year-over-year. As of 2020, digital ad spend accounted for half of all advertising spend in the United States. See, Johnson, Bradley, "In a New Milestone, the Internet Will Account for Half of Ad Spending in 2020," *AdAge*, December 23, 2019, accessed June 5, 2020, <https://adage.com/article/datacenter/new-milestone-internet-will-account-half-ad-spending-2020/2223511>

made available massive amounts of frontend data through a suite of data access mechanisms, such as likes, comments, and views. Yet, platforms effectively control their information infrastructure on the backend, keeping the majority of consumer and usage data proprietary.

Platforms only share limited sets of backend data with select partners. For example, Facebook enables access to data on marketing campaigns on its platform in order to showcase the campaign results and thereby incentivize companies to spend their marketing budgets on Facebook rather than on competitive platforms. To that end, Facebook grants film studios and television networks access to second-party, shared data as part of paid marketing campaigns. Data-sharing is effectively a strategic business decision. Consequently, the degree to which platforms share data differs by business model.

- *Ad-supported*: Ad-supported platforms like Facebook, YouTube, or BuzzFeed share data on the frontend and backend to illustrate the results of marketing campaigns and incentivize companies to spend more advertising dollars. Data sharing effectively serves as a promotional tool for the platform.
- *Transactional*: Transactional platforms like the Apple App Store or the Google Play Store are primarily designed for electronic sell-through (EST), selling and renting film and television programming as digital home entertainment hubs. Thus, they only share limited amounts of frontend data, mainly for promotional purposes. For example, both the App Store and Play Store categorize games and apps by the number of downloads to give users an indication of popularity (i.e., 1 million downloads), thereby incentivizing further usage and downloads.

- *Subscription*: Subscription platforms like Netflix or Amazon Prime Video share almost no data with the majors on the frontend and backend since they have no economic incentive to do so. In fact, they keep the majority of their data proprietary as the data represents a competitive advantage against the majors.<sup>108</sup>

Smith and Telang<sup>109</sup> illustrate the economy of second-party data with a detailed breakdown of data-sharing modules. Their work further illustrates the ownership and control platforms have over data access (table 1.5).

**Table 1.5**

Distributor	Transaction-level data	Customer-level data	Direct promotion to customers
iTunes	Yes	Limited (ID and ZIP code)	No
Amazon	Yes	No	No
Google Play	No	No	No
Netflix	No	No	No

In this sense, platforms have enabled exponential data generation for Hollywood’s data access pipeline, yet equally introduced increasing levels of fragmentation. Effectively, distributing content across platforms has increased the legacy media industry’s data output while complicating its direct data access.<sup>110</sup>

---

<sup>108</sup> Subscription platforms’ secretive data sharing policy has led to a debate about transparency. See, Ng, David, “Netflix plays peekaboo with its ratings. Hollywood isn’t amused.,” *Los Angeles Times*, January 24, 2019, accessed June 5, 2020, <https://www.latimes.com/business/hollywood/la-fi-ct-netflix-ratings-20190124-story.html>

<sup>109</sup> Table quoted from Smith & Telang, 2016

<sup>110</sup> Hollywood has created a stream of frontend data across the digital landscape, yet the majors have largely lacked direct access to the underlying information on the backend, which is owned by platforms. For example, studios can see the number of views on a YouTube trailer, but cannot easily see how many people watched the entire video.

### Emerging Technologies and Pipeline Future-Proofing

The legacy media industry further expanded its data access pipeline by creating and distributing content across emerging technologies, such as virtual reality (i.e., VR), augmented reality (i.e., AR), and voice tech (e.g., Amazon Echo). Emerging technologies represent another digital marketing channel, covering a suite of products that are not yet commercially viable, but indicate potential to become future platforms. While they operate under the same principles as platforms, designed to aggregate users in an effort to capture, consolidate, and commercialize data, they lack the scale and influence of platforms, given their nascent development and small user bases. Still, emerging technologies have offered distinct benefits to the majors to expand their reach across the digital landscape, namely new content formats, new audiences, and new data types.

Hollywood has made ongoing attempts to establish an early presence in emerging technologies. For example, in virtual reality, the majors invested in a suite of new content suppliers<sup>111</sup> to create content for the main VR platform of the time, the Oculus Rift, and built internal departments to understand and commercialize the future opportunity in VR, such as 21st Century Fox's innovation lab (launched in 2013), The Walt Disney Company's startup accelerator (opened in 2014), and Paramount's interactive media division (started in 2015). Effectively, the majors created a foundation to capitalize on the promise of emerging technologies, the future potential for scale, and the related data opportunity.

Indeed, emerging technologies offer unique data types that are not inherent in the platform economy. Virtual reality data tracks immersive digital interaction beyond clicks, swipes, and views; augmented reality tracks the interplay between digital behavior and the real

---

<sup>111</sup> See, Matt Pressberg & Matt Donnelly, "Hollywood's Virtual Reality Push: How All 6 Studios Stack Up," *The Wrap*, July 24, 2017, accessed June 6, 2020, <https://www.thewrap.com/hollywood-virtual-reality-push-how-all-6-major-studios-stack-up/>

world; and voice tech emphasizes language input over haptic digital actions. In this sense, emerging technologies represent opportunities to *land-grab* new data types within the information infrastructure.<sup>112</sup>

At the same time, emerging technologies complicate data access by conflating the frontend and the backend, keeping all data entirely hidden from view, not offering a visual layer that makes data publicly available.<sup>113</sup> VR, AR, and voice tech all incorporate visual user interfaces that quantify information, yet that information is designed to structure the experience of emerging technologies rather than documenting consumer and usage data.

As such, while emerging technologies offer limited data volume, they promise more data variety. Thus, by investing in emerging technologies, the majors may have expanded their data access pipeline only incrementally, but laid a foundation for additional data types and potential future data scale.

--- --- ---

Ultimately, by investing in digital marketing initiatives across audience feedback mechanisms, specifically websites, platforms, and emerging technologies, the legacy media industry has generated a constant stream of data on the film and television industries, making growing amounts of consumer and usage data available at scale. Indeed, by integrating with the digital infrastructure of platforms, which translate content input into information output, the majors have created a generative data access pipeline, yielding information that was previously

---

<sup>112</sup> Netflix's early move into streaming has been characterized as "the greatest land grab in the history of media." See, Kafka, Peter, "How Netflix outsmarted everyone else in TV," *Vox*, August 23, 2018, accessed June 5, 2020, <https://www.vox.com/2018/8/23/17770896/netflix-reed-hastings-ted-sarandos-streaming-tv-media-jason-hirschhorn-redef-peter-kafka-podcast>

<sup>113</sup> Consider the example of virtual reality (VR). Virtual reality headsets like the Oculus Rift do not incorporate an externally accessible visual interface layer to showcase public frontend data (i.e., only the user can see data).

unavailable to the legacy media industry. At the same time, data access is fragmented, with platforms owning and controlling access to the majority of data. For the majors, data access is accordingly limited to publicly available information on the frontend, making the process largely manual and not scalable.

### **Data Licensing: Automation and Access**

In 2005, tech giant Google announced the acquisition of Urchin, a small software company that specialized in tracking, collecting, and accessing website data.<sup>114</sup> From the press release (my emphasis):

“Urchin is a web site [sic] analytics solution used by web site owners and marketers to better understand their users’ experiences, optimize content and track marketing performance. **Urchin tools are available as a hosted service, a software product and through large web hosting providers.** These products are used by thousands of popular sites on the Internet. Google plans to make these tools available to web site owners and marketers to better enable them to increase their advertising return on investment (ROI) and make their web sites more effective. **“We want to provide web site owners and marketers with the information they need to optimize their users’ experience and generate a higher return-on-investment from their advertising spending,”** said Jonathan Rosenberg, vice president of product management, Google. “This technology will be a valuable addition to Google’s suite of advertising and publishing products.”

Google effectively presented the Urchin acquisition as a way to deliver more information to website owners and marketers, providing value through data. In particular, the press release characterizes Urchin as a data access mechanism that enables website owners to collect and understand the digital activity on their websites. Moreover, Urchin’s tools are presented as “a hosted service” and “a software product,” underscoring that the service is easy-to-use by a broad

---

<sup>114</sup> Google, “Google Agrees to Acquire Urchin,” May 28, 2005, accessed June 5, 2020, [http://googlepress.blogspot.com/2005/03/google-agrees-to-acquire-urchin\\_28.html](http://googlepress.blogspot.com/2005/03/google-agrees-to-acquire-urchin_28.html)

set of users. Finally, as the new owner, Google positioned itself as the acting arbiter of access to Urchin's data services. Indeed, following the acquisition, Google rebranded Urchin's product suite as Google Analytics, the company's free-to-use website tracking service that grants users access to web information across the Internet, facilitating advanced advertising capabilities.

In a similar vein, social media platform Facebook launched *Facebook Ads* (later dubbed *Facebook Insights*), the fledgling version of what would become its global advertising network, enabling marketers to promote their products and services directly to Facebook's millions of users by leveraging the company's access to advanced consumer and usage data. As CEO Mark Zuckerberg noted in his keynote address to a cross-section of Fortune 500 brands, including a broad set of legacy media companies (my emphasis):

“Facebook Ads represent [sic] a completely new way of advertising online. For the last hundred years [sic] media has been pushed out to people, but **now marketers are going to be a part of the conversation.**”<sup>115</sup>

Similar to Google's announcement, Zuckerberg's statement positions Facebook as the gatekeeper of data access on its platform. Furthermore, Zuckerberg notes that Facebook Ads enables marketers to have more control over their advertising. Indeed, his statement is an indirect reference to the legacy media industry's missing connection with the audience and the resulting lack of access to audience information. He positioned Facebook Ads as a remedy to this state of disintermediation, enabling a direct connection between media and audience, which would enable better marketing and advertising initiatives through advanced access to consumer and usage data. Effectively, the early version of Facebook Ads provided the company's brand

---

<sup>115</sup> Facebook, “Facebook Unveils Facebook Ads,” November 6, 2007, accessed June 5, 2020, <https://about.fb.com/news/2007/11/facebook-unveils-facebook-ads/>

partners with three core abilities: one, to create branded pages on the social network, thereby building an online presence through digital real estate; two, to boost promotional messages across the platform, directly targeting specific audience segments; and, three, to access data on ad performance and user interaction with content. In this sense, Facebook Ads laid a foundation for companies to create, distribute, and measure digital marketing content.

Google Analytics and Facebook Ads represent data access systems. They are proprietary, cloud-based software technology, designed to give third parties access to data on the backend of platforms. As such, these data access systems effectively enable external companies to plug into the larger information infrastructure of platforms and access select data that is generally kept proprietary and not publicly available. Through Google Analytics and Facebook Ads, Google and Facebook allow advertisers to access information on the marketing and advertising campaigns they ran on the platforms. In this regard, platforms leverage data access systems to license data access to third-party companies.

Data licensing provides third-party companies a specific form of access to the information infrastructure of platforms, defined by a specific set of characteristics. First, data licensing enables data access at a cost. While data access software like Google Analytics and Facebook Ads is billed as free-to-use, companies still pay a price for usage. Specifically, they are required to sign up for business accounts on the platforms, effectively making them part of the platform ecosystem. Additionally, they need to spend money on advertising campaigns in order to access the associated data, thereby directly engaging with the ecosystem. For example, when a brand like Coca-Cola runs an advertising campaign on Facebook, deploying shares of its advertising budget, Facebook Ads in return provides “free” access to data associated with the campaign. In this regard, companies are charged an indirect cost to license data from platforms.



Second, licensed data access is limited to the information platforms make available. Rather than granting third parties universal access to the platform backend, platforms license a curated set of information, tailored specifically to the advertising campaigns companies pay for. For example, Facebook Ads presents a standard set of consumer and usage data to illustrate the performance of an advertising campaign, including user demographics, affinities, and interests (i.e., consumer data) as well as impressions and engagement (i.e., usage data). Platforms thus license data that is contextually relevant, yet only represents a small sample of the larger platform data they own.

Third, licensed access is administered by data access systems through specific software technology, a suite of so-called software-as-a-service (SaaS) tools embedded within platform ecosystems.<sup>116</sup> This software runs on the cloud and makes data directly available and usable to companies through a visual interface. The interface enables companies to share, display, and use the data within their organizations. For example, companies can share the Google Analytics and Facebook Ads/Insights interfaces with teams across organizational structures, enabling them to access different types of data.

At the same time, the software layer of the data access interface is owned and controlled by platforms and thus operates based on a set of rules defined by them. Companies need to log into the software, follow the rules of the software, and are limited by the framework of the software (i.e., what data they can access, how they can access it, and what they can do with it). As such, companies do not have direct access to the backend data of platforms. They do not have access to data spreadsheets or virtualized databases to use and manipulate data in different ways.

---

<sup>116</sup> Software as a Service (SaaS) is a digital business model that licenses access to cloud-based software applications in exchange for a monthly or annual subscription fee.

Their access to data is standardized and defined by the licensing terms set by platforms and encoded in the design of the interface.<sup>117</sup>

The use of software highlights a crucial variable of automation in the data access narrative. Data access systems like Google Analytics and Facebook Ads automatically collect, consolidate, and condense data through a software-powered interface. Automation enables consistent data availability and usability for companies. It helps companies to access data from the backend at scale, supplanting the process of manually searching data on the frontend.

The legacy media industry has invested in data licensing as a way to automate the access to platform data at scale. In particular, the majors have maintained a data licensing network through their investment in marketing initiatives across the digital landscape, covering web, social media, mobile apps, online video, and emerging technologies. Indeed, to increase their access to platform data, film studios and television networks have developed strategic partnerships with a broad set of technology companies to leverage the software capabilities of their data access systems, including platforms, startups, and upstarts. Through these partnerships, the majors expanded their data access infrastructure, plugging into a range of data access systems and thereby access a wide spectrum of data, from first-party to second-party and third-party data. More pointedly, the data licensing partnerships effectively enabled the majors to vertically integrate with the data backend of platforms, gaining increased degrees of data access.

The legacy media industry has increased access to platform data through three distinct data licensing partnerships with data access systems:

---

<sup>117</sup> Some platforms enable downloads of unstructured data spreadsheets for use outside of the platform interface.

- *Platforms and Data Applications*: Platforms license data access via a suite of free-to-use, self-serve *data software applications*, giving the majors access to a curated set of data on their backend. The software applications provide access through an interface that displays a visual layer of information. Specifically, the majors are able to use data applications and gain access to platform data if they spend their marketing budget on the platform. Platforms use data applications as a core element of their business model, showcasing the value of their products and services for digital marketing initiatives and thereby incentivizing further usage of their ecosystem. Additionally, by giving third parties access to a select set of information, platforms equally gain access to the data generated by third parties. By licensing data access to film studios and television networks, requiring the majors to distribute content through their channels, platforms thus built a viable repository of industry information. The majors, meanwhile, derive benefit from the licensing partnership by gaining access to more platform data and expanding their overall data access infrastructure. On the other hand, data applications are marked by distinct restrictions. One, data applications only provide access to limited platform data, which makes it difficult for the majors to gain a more holistic information stream. Two, access to data applications is entirely controlled by platforms, creating a dependency that leaves the majors little room to maneuver and negotiate. And, three, data applications only provide access to the backend of one particular platform, rather than a suite of platforms in the digital landscape, which results in fairly narrow access.
- *Startups and Data Subscriptions*: Startups are small, largely independent software companies that serve as the data access link between platforms and companies. Effectively, startups provide intermediary access to a broader set of data, across multiple

platforms and channels, using two principal methods. One, they collect, curate, and consolidate backend data from platforms, made available through licensing agreements with the major platforms. As such, they can access the limited backend data of various platforms rather than just one. Two, startups search, scrape, and synthesize all publicly available data across the digital landscape, pulling the information from the frontend interfaces of websites, social media accounts, and mobile apps via specific software tools (e.g., web crawling and scraping). In this regard, startups provide access to a suite of third-party data, across the backend and frontend, offering a cross-section of proprietary and public information. They offer broader, more holistic access, capturing, fusing, and synthesizing fragmented platform data with vast amounts of public data. Effectively, the data access administered by startups can be seen as an extension of platform data access. Startups license access to this repository of third-party information via *data subscriptions*, pay-to-play software that charges a monthly or annual fee for data access, made available through custom interfaces as well as automatically generated reports.<sup>118</sup> While all startups offer subscriptions that integrate platform and public data, they differ in the type of data sources they provide. To differentiate their business, startups thus prioritize specific data sources, specializing in web, social, mobile, or digital data to complement the access offered by platforms. In this sense, data subscriptions offer access to a combination of cross-platform and public data. For the majors, data subscriptions thus enable additional data scale, offering more volume and variety of information. At the

---

<sup>118</sup> While data startups operate independently of platforms, their business is highly dependent on them. In effect, they need access to platform data to function as intermediaries between platforms and third parties / external companies. As such, they effectively work as extensions of platforms' data ecosystems.

same time, data subscriptions are administered by startups which are highly dependent upon platforms, limiting the control of the majors.

- *Upstarts and Data Investments*: Upstarts represent a set of ad-supported digital media companies that use data to inform their business operations. Their business is built on using data to create, program, and distribute content for advertisers in the digital landscape, using their owned channels (e.g., a website or mobile app) as well as external platforms (e.g., social media accounts, mobile apps, OTT) to distribute their content. To this end, upstarts built proprietary software that integrates three sets of data. One, upstarts distribute content across major platforms and access the associated data. Two, they access public data on their content across the digital landscape, pulling information from public digital channels like websites, blogs, and forums. Three, they access the data from their owned websites and apps, some of which count millions of users and thus represent highly valuable data repositories. As such, upstarts combine platform, public, and owned first-party data into a unique database of information.<sup>119</sup> Similar to platforms and startups, upstarts license their data access via software to third parties. However, the majors pursued a deeper licensing partnership with upstarts by acquiring small ownership stakes in them, with several film studios and television networks investing in upstarts. While these *data investments* can be read as the legacy media industry's attempt to bolster and update increasingly significant digital production capabilities, they also signal a shift from licensing to owning data access. Data investments gave the majors an opportunity to own vast digital databases, filled with information on digital consumers and their

---

<sup>119</sup> For example, digital upstart BuzzFeed built a system to track content sharing across platforms, using first-party, second-party, and public data. See, Ingram, Mathew, "The Real Secret to BuzzFeed's Success Isn't Cat Gifs, It's Data," *Fortune*, February 16, 2016, accessed June 5, 2020, <https://fortune.com/2016/02/16/buzzfeed-data/>

interaction with content. At the same time, acquiring a stake in the data access infrastructure of upstarts involved complex technological and organizational issues.

These data licensing partnerships enabled the legacy media industry to expand its data access infrastructure. The use of data access systems, from data applications to subscriptions and investments, introduced a software layer that increased the scale of data access. At the same time, it yielded more complex dynamics of control and ownership over data access.

Platforms, startups, and upstarts effectively functioned as data *brokers*, deploying their proprietary data access systems to license different degrees of access to the majors. Indeed, the software-as-a-service approach – the seamless, easy-to-use, and scalable implementation of data access software – made data more available and usable for the majors.<sup>120</sup> Hollywood’s data access pipeline created a system where data is consistently generated. The majors’ turn to data licensing laid a foundation for automated data collection where information is captured, collected, and consolidated by software.

### *Platforms and Self-Serve Data Applications*

Platforms deploy data applications to enable data licensing. Data applications are software programs designed to automatically collect, organize, and present select sets of data from the platform backend. They enable external companies to *self-serve*, accessing data through custom interfaces, which make information both available and usable for legacy media organizations.

---

<sup>120</sup> Specifically, data brokers made access available and usable by way of a visual interface, moving data from the backend and synthesizing it on the frontend. Rather than providing access to data sets that need to be cleaned up, structured, and organized, they process data as readable and relatable information. This heuristic approach created a foundation to transform data from an abstract entity into a tangible resource for the legacy media industry.

Platforms leverage data applications as a way to drive and sustain their business model. Data applications are specifically used by ad-supported platforms, which rely on the business of third-party advertisers. By licensing data access, these platforms showcase the value of their data to third parties and thereby incentivize further platform usage and investment in their advertising business. Accordingly, the majority of ad-supported platforms commercialized data applications to enable data access for external companies (table 1.6 shows a selection of data applications).

**Table 1.6**

<b>Platform Data Applications</b>			
<b>Platform</b>	<b>Application</b>	<b>Launch Year</b>	<b>Data Source</b>
Google	Google Analytics	2005	Web
Google	Google Trends	2006	Search
Facebook	Facebook Ads (later dubbed Insights)	2007	Social
YouTube	YouTube Insight (later dubbed Analytics)	2008	Video
Apple	App Store Analytics	2010	Mobile

Platform-based data applications constitute universal data access systems, using a visual interface to enable access to information across the digital landscape, including web, search, social, video, and mobile data. The data interface moves data from the platform backend to the frontend, organizing, structuring, and presenting it on a digital canvas. As such, data applications render information available and usable, framing data access as a tangible resource, rather than an abstract concept, laying the foundation for the emergence of data literacy within the industry.

For the legacy media industry, data applications have facilitated data access in several ways. First, they served as a mechanism for the majors to access the data generated from their digital marketing initiatives. Effectively, they offered a seamless and convenient way to access

data from a specific platform, allowing film studios and television networks to collect and comprehend consumer and usage data generated by their content across platforms. For example, Google Analytics provides access to website data; Facebook Insights provides information on Facebook pages; Apple's App Store gives insight into mobile apps and games.

Second, data applications curated platform data to reveal a specific set of consumer and usage information, effectively making the information not only available to the majors, but digestible and relatable. Third, they promoted advanced usability for the majors, helping legacy media organizations access custom interfaces at any time to view digital information. Using the cloud storage capabilities of platforms, data applications created a level of data transparency that has made data part of the day-to-day of the legacy media business. Finally, data applications are owned, controlled, and managed by platforms, which provide the technological infrastructure. As such, they did not require the majors to invest at scale in technology or talent to access data.

At the same time, data applications also limited and complicated access for the majors. While their use of the cloud enabled scaled data access and usage across organizations, they did not offer universal data access. For one, they only provided access to data from one platform at a time. For example, Google Analytics makes available website data, Facebook Insights provides data on Facebook pages, and YouTube Analytics creates a connection to data from YouTube videos. Furthermore, they only gave access to a curated selection of platform data, while the majority of data remained hidden on the platform backend, which made the information inherently limited. In this regard, the majors did not have any control over broader, cross-platform data access and thus were unable to establish unified data access across platforms. As a result, data access was fragmented and restricted across platforms, which, in turn, effectively worked to reinforce the control and power platforms hold over the information infrastructure.



Additionally, while the use of custom interfaces automated data collection and visualization to enable seamless access, the interfaces only provided a static perspective on data. Put differently, the majors were unable to customize the interface and thus only had limited control over how data was made available. This limited functionality further framed data access in a particular way. The interfaces of data applications are designed and programmed by platforms, which present the available information in a specific way. As such, the interfaces prompted specific readings and interpretations of data, positioning access as directed rather than open. In this sense, the majors only ever had access to data that was filtered, never raw.<sup>121</sup>

Data applications enabled the majors to automate data access at scale, facilitating the collection and comprehension of the data Hollywood's digital marketing consistently generates. Concurrently, they provided limited and fragmented access to platform data, thereby working to reinforce established power dynamics.

### *Startups and Data Subscriptions*

Startups use a subscription model to license data access across platforms and the broader digital landscape. Data subscriptions charge a direct fee for access to custom interfaces as well as reports that make available a consolidated canvas of platform and public data. As such, they enabled the legacy media industry to gain more holistic data access, synthesizing various strands of information from across platforms.

Startups effectively positioned themselves as intermediaries between legacy media companies and the major platforms. For platforms, startups offered an incremental monetization mechanism. Startups paid a licensing fee to platforms in order to pull data from the platform

---

<sup>121</sup> See, Gitelman, 2013

backend and make it available through their own custom interface. As such, startups effectively served as a promotional tool for platforms, showcasing the value of their data to external companies, thereby incentivizing further usage and potential advertising spend across the respective platforms (table 1.7).

**Table 1.7**

<b>Data Startups</b>		
<b>Company</b>	<b>Launch Year</b>	<b>Ownership</b>
Radian6	2006	Salesforce (acquired in 2011)
Crimson Hexagon	2007	Independent (merger with Brandwatch in 2018)
Fizziology	2009	MarketCast (acquired in 2017)
Listen First	2010	Independent
Sprout Social	2010	Independent

Data subscriptions offered increased levels of automation. Beyond the custom interface that aggregates available information in one virtual place, subscriptions made available custom reports, automatically generating data summaries and snapshots that could be accessed on any digital device.

Furthermore, data subscriptions automated access to data across several platforms through the use of custom protocols, generally referred to as application programming interfaces (APIs). APIs enable startups to integrate the data applications of platforms with their own data access systems, thereby creating a holistic tool that synthesizes information across multiple platforms. For instance, consider the startup Crimson Hexagon, which uses APIs from major platforms, including Twitter, Facebook, and YouTube, to facilitate access to Twitter, Facebook, and YouTube data, along with various strands of publicly available data (e.g., website, forum,

and social media comments), in one place. As such, data subscriptions increased the scale of data access while equally limiting the fragmented nature of access in the digital landscape.

While data subscriptions were built on providing a cross-platform approach to data access, their business has effectively relied on platforms. Startups largely do not own any of the data they provide access to. They thus depend on access to second-party data from platforms and publicly available data sources, such as websites, blogs, and forums. The crucial benefit of data subscriptions is to synthesize and fuse data from multiple platforms into a standalone infrastructure that offers more comprehensive and holistic access. While the data subscription model seemingly offers an alternative to platform data applications, it in fact constitutes an extension of it, offering a holistic data view that incentivizes external companies to increase their marketing and advertising spend on platforms. To secure access to platform data, startups developed licensing partnerships with all platforms, effectively paying for access to select platform data. In this regard, the business model of data subscriptions has worked to reinforce the power dynamics of platforms.

The majors have adopted a variety of data subscriptions to boost data access. In fact, film studios and television networks generally managed multiple subscriptions at a time. This multi-subscription approach is likely attributable to various reasons. One possible reason is strategic in nature. By holding several subscriptions, the majors have maintained access to multiple data sources across the digital landscape, prioritizing subscriptions by data vertical (i.e., web, social, mobile, video) to increase the scale of access. Another potential reason is a growing focus on digital experimentation to ensure studios and networks were up-to-date on new technologies. Startups generally offered demos and trials for their software updates, enabling the majors to test-and-learn during a specific time period, which resulted in several technologies being used

across various departments and teams. At the same time, it is likely that the increase in subscriptions reflected a budgetary issue, as digital budgets are allocated across different departments, which are not always aligned on what technology is sourced.

The consistent implementation of data subscriptions has enabled the majors to scale data access through automation. Additionally, the integration of software and automation has established a core foundation for the majors to build and develop data literacy, a growing familiarity with data as a key lever of the media and entertainment business.

The emergence of data subscriptions established focus, minimized fragmentation, and increased ease of data access by synthesizing information from across platforms. However, by tapping into the established data infrastructure of platforms, data subscriptions have also reinforced platform power dynamics.

Legacy media companies have used both data applications and subscriptions to establish and expand data access. While both have been effective in helping the majors build out their data access infrastructure, giving access to increased consumer and usage data as well as advanced usability through software interfaces that enable access across organizational structures, they also proved inherently limited and incremental, prioritizing pockets of data silos in favor of a holistic data infrastructure. Indeed, both data applications and subscriptions place emphasis on the licensing of information over ownership.

### *Upstarts and Data Investments*

The legacy media industry has consistently maintained licensing partnerships to access data in the digital landscape. At the same time, the majors have made efforts to transition from being a licensing partner to acting as an equity stakeholder in data access, effectively exploring strategies

to own data access and bring it in house. To pressure-test the dynamics of data ownership, the majors thus made strategic financial investments in upstarts, a suite of digital media companies that uniquely combine some of the data access capabilities of platforms and startups. By gaining small ownership stakes in their data economies, the majors partnered with upstarts to expand their own data access infrastructure.

Upstarts presented viable investment targets for the legacy media industry, for several reasons. First, the business of upstarts aligned with the economic focus of the legacy media industry, specifically advertising and original programming. Upstarts are digital media companies that create, program, and distribute content for third-party advertisers in the digital economy. Additionally, they invest in original IP development and monetization across their owned websites and apps as well as third-party channels. In this regard, upstarts aligned organically with the corporate structures of the majors.

Second, upstarts developed proprietary technologies to use data for cultural production, engineering content to create the highest possible rate of digital engagement. This mode of production has become progressively viable for legacy media companies as they compete for viewership in an increasingly fragmented media landscape. Furthermore, upstarts have scaled their content output to build massive digital user bases, boasting millions of users that regularly visit their owned and distributed channels and engage with their content, which provided the majors with substantial volume and variety of information.

Third, upstarts built proprietary software to access consumer and usage data across platforms, public digital channels such as websites and blogs, as well as their owned channels (i.e., their websites and mobile apps). As such, upstarts offer access to a diverse data mix,

combining third-party (i.e., public information), second-party (i.e., platform data), and first-party data (i.e., owned channels) into a holistic data access infrastructure.

Upstarts represent a broad cross-section of digital media companies that combine expertise in data and technology with advanced capabilities in digital content production. They are largely ad-supported, create and distribute content across multiple platforms, and collect, consolidate, and commercialize the associated data to move their business forward. In general, they can be categorized into two distinct categories: digital publishers and multi-channel networks.<sup>122</sup> One of the most prominent digital publishers is BuzzFeed. Founded in 2006, BuzzFeed has undergone multiple iterations over the course of its existence, yet largely maintained the core elements of its business. BuzzFeed built a dedicated website and mobile app, yet has equally distributed content across the major platforms, thereby creating a vast content network that generates massive amounts of data. BuzzFeed uses proprietary software to capture and synthesize the information into a strategic two-fold business model. One, it packages the data for advertisers to enable advanced audience targeting. Two, it uses the data to create custom content that is likely to increase with its audience. BuzzFeed founder Jonah Peretti characterized the company's business model as audience-centric, noting that, "[t]here's an opportunity for a modern media company to be more engaged with the audience than ever before, and have a more intimate connection in people's lives, to respond and be reactive to the things that matter to

---

<sup>122</sup> For digital publishers, see, Pang, Justin, "The Future of News and Publishing," *TechCrunch*, February 17, 2016, accessed June 6, 2020, <https://techcrunch.com/2016/02/17/the-future-of-news-and-publishing/>. For MCNs, see, Christian, Aymar Jean, *Open TV: Innovation Beyond Hollywood and the Rise of Web Television* (New York: New York University, Press, 2018).

people.”<sup>123</sup> As such, BuzzFeed positioned itself as a company that operates through access to audience information.

Multi-channel networks (MCNs), digital distributors of video programming, operated similarly in that they created, collected, and commercialized content and data, with the key distinction that MCNs scaled their business within one central platform, rather than across multiple platforms.<sup>124</sup> MCNs primarily built their business on YouTube, aggregating multiple video channels to create a vast network of content that attracted millions of users and monetize the associated audience information through advertising deals. MCNs still distributed content across other platforms as well as their owned channels, yet their main focus remained on the YouTube ecosystem given its high concentration of talent, content, and data.

In this regard, digital publishers offered a diverse data mix, combining first-party data from their owned channels with platform data and public information. Multi-channel networks, meanwhile, offered a similar information infrastructure, yet placed more emphasis on YouTube video data.

Accordingly, the legacy media industry focused upstarts investments on digital publishing and multi-channel networks. Investing in digital publishers allowed the majors to build a foundation to own data access across platforms, while the stakes in MCNs enabled owned access to YouTube platform data. From the late 1990s to the mid-2010s, legacy media

---

<sup>123</sup> Robischon, Noah, “How BuzzFeed’s Jonah Peretti Is Building A 100-Year Media Company,” *Fast Company*, February 16, 2016, accessed June 5, 2020, <https://www.fastcompany.com/3056057/how-buzzfeeds-jonah-peretti-is-building-a-100-year-media-company>

<sup>124</sup> Eventually, multi-channel networks transitioned to a multi-platform model, producing and distributing programming across multiple platforms. As such, they effectively adopted the distribution model of the legacy media industry. See, Lobato, Ramon, “The cultural logic of digital intermediaries: YouTube multichannel networks,” *Convergence*, Volume 22, Issue 4, 2016 and Vonderau, Patrick, “The video bubble: Multichannel networks and the transformation of YouTube,” *Convergence*, Volume 22, Issue 4, 2016.

companies made strategic investments in several digital publishers and MCNs, including BuzzFeed, Vice Media, Maker Studios, Machinima and Awesomeness TV (table 1.8).

**Table 1.8**

<b>Investments in Digital Publishers</b>				
<b>Company</b>	<b>Business Focus</b>	<b>Launch Year</b>	<b>Ownership</b>	<b>Legacy Media Investors (selection)</b>
Refinery29	Digital publisher	2005	Vice Media (acquired in 2019)	Turner Broadcasting, Scripps Networks
Mashable	Digital publisher	2005	Independent	Turner Broadcasting, Time Warner Investments
BuzzFeed	Digital publisher	2006	Independent	NBC Universal
Vice Media	Digital publisher	1994 (digital in 2006)	Independent	The Walt Disney Company, 21st Century Fox, A+E Networks
Vox Media	Digital publisher	2005 (rebranded in 2011)	Independent	Comcast Ventures, NBC Universal
<b>Investments in Multi-Channel Networks</b>				
<b>Company</b>	<b>Business Focus</b>	<b>Launch Year</b>	<b>Ownership</b>	<b>Legacy Media Investors (selection)</b>
Machinima	Multi-channel network	2000	Warner Bros. (acquired in 2016)	Warner Bros.
Maker Studios	Multi-channel network	2009	Disney (acquired in 2014)	Time Warner Investment, Downey Ventures
FullScreen	Multi-channel network	2011	Otter Media (AT&T/WarnerMedia ; control in 2014)	Comcast Ventures, The Chernin Group
StyleHaul	Multi-channel network	2011	RTL Group (acquired in 2014)	RTL Group, Bertelsmann Digital Media Investments
Awesomeness TV	Multi-channel network	2012	Viacom CBS (acquired in 2018)	Verizon, UTA Ventures, DreamWorks Animation



By investing in upstarts, the legacy media industry effectively acquired ownership stakes in the companies' underlying data access systems. As a result, the majors further expanded their data access infrastructure, with a specific focus on owning data access.

--- --- ---

The legacy media industry has consistently maintained a strategic focus on licensing data to expand its data access infrastructure, negotiating partnerships with a suite of technology, specifically platforms, startups, and upstarts, and leveraging their proprietary software systems to automate data access. In particular, the use of a software-as-a-service (SaaS) approach enabled a mode of access that has made data increasingly available and usable for legacy media organizations, for several reasons. One, software-based data access systems like data applications, subscriptions, and investments facilitated seamless and convenient access, displaying data through custom interfaces that can be used and shared by various departments and teams within the organizational structures of film studios and television networks. Two, the use of cloud-based software did not require the majors to invest in complex and expensive technological on-premise infrastructure (i.e., building physical data centers), instead offering a licensing model that provides data access through virtualized data centers. And, third, partnering with technology companies to license data access via software proved similar to the majors' long-standing research approach, outsourcing surveys, questionnaires, and interviews to external research vendors that deliver research and information reports.

Furthermore, by working with a suite of technology companies on data access, the legacy media industry has consistently expanded its data literacy, effectively creating a foundation to deal with data as a form of capital within the evolving digital media landscape.

Yet, while the focus on data licensing has expanded data access for the majors through software automation, it has further cemented platforms' control and ownership over data access. A large part of the majors' data access, excluding owned channels such as specific film and television websites, is in fact owned and controlled by platforms. Platforms license a limited amount of data from their backend to the majors, directly (via data applications) and indirectly (via data subscriptions, administered by startups). Indeed, platforms also provide the infrastructure for upstarts whose business has been built on the global audience research of platforms. Thus, while the majors were able to navigate the entrenched data disintermediation, their focus on data licensing defined access through the licensing terms set by platforms.

### **Data Acquisitions: From Aggregation to Centralization**

In 2014, The Walt Disney Company announced the acquisition of Maker Studios ("Maker"), one of the digital media industry's leading multi-channel networks, a distributor of programming across YouTube.<sup>125</sup> Founded in 2009, Maker had previously received investments from other legacy media players, including media conglomerate Time Warner, French media company Canal+, and film producer Jon Landau. The company's business model was geared towards the creation, programming, and distribution of digital short-form content, with a primary focus on YouTube. At the time of the acquisition, the Maker portfolio included over 55K YouTube channels in over 100 countries, led by a roster of emerging online talent with substantial fan bases, accounting for over 650 million subscribers with 11 billion video views per month.

---

<sup>125</sup> See, Barnes, Brooks, "Disney Buys Maker Studios, Video Supplier for YouTube," *New York Times*, March 24, 2014, accessed June 5, 2020, <https://www.nytimes.com/2014/03/25/business/media/disney-buys-maker-studios-video-supplier-for-youtube.html>

Additionally, Maker had launched a dedicated website, called Maker.tv, to expand its business operations beyond YouTube and establish a proprietary video advertising platform.

Bob Iger, Chief Executive Officer of Disney at the time, positioned the acquisition as a distribution play that would extend Disney’s digital monetization capabilities, noting, “[w]e see it first and foremost as a distribution platform and a very successful one, one that not only can command more eyeballs, more consumption, but with that more advertising revenue or revenue in general.”<sup>126</sup> Additionally, he acknowledged that Maker enabled Disney to close a digital distribution gap, stating “[w]e did not believe that we had the ability... to distribute as effectively and to sell as effectively” and “[it would have taken] a long time to build [that] kind of technological expertise.”<sup>127</sup>

Jay Rasulo, Disney’s Chief Financial Officer at the time, summarized the acquisition rationale as a way to extend Disney’s key business pillars, the production and distribution of content, into the digital economy:

“Our acquisition strategy has been pretty clear. We are looking for one of two things. Intellectual property that we can distribute through the Disney ecosystem which today is theme parks, theatrical, TV, cruise ships, etc. Or something that we can use to extend that network. Maker is a perfect example of that. They are a content creation vehicle... Maker Studios have clearly established themselves as a short-form creation studio. But they are also masters of distribution. It was a twofer for us.”<sup>128</sup>

---

<sup>126</sup> Graser, Marc, “Bob Iger Explains Why Disney Bought Maker Studios,” *Variety*, May 6, 2014, accessed June 5, 2020, <https://variety.com/2014/biz/news/bob-iger-explains-why-disney-bought-maker-studios-1201173389/>

<sup>127</sup> Ibid

<sup>128</sup> Lunden, Ingrid, “Disney’s Next Acquisition Could Further Its Digital distribution,” *TechCrunch*, January 7, 2015, accessed June 5, 2020, <https://techcrunch.com/2015/01/07/disney-eyes-more-acquisitions-with-a-focus-on-digital-distribution/>

Rasulo further noted that Disney specifically looked to acquire Maker’s “underlying technology,” which facilitated a direct connection to consumers and, as such, represented a viable data access line. “Every company wants more direct relationships with consumers. They want to own the data.”<sup>129</sup>

As such, Maker effectively constitutes a data access acquisition, a way for legacy media companies to absorb the data access infrastructure of digital media companies, specifically their underlying technology and user base. While the majors continued to maintain strategic partnerships with platforms and technology companies to license data access, the acquisition of digital media companies enabled a nascent form of data ownership.

Notably, the acquisition of Maker did not enable Disney to own direct access to data. Maker did not operate as a platform that collects, consolidates, and commercializes data access. Rather, Maker operated a network on top of a platform. As such, Maker did not own the data it generated. Instead, Maker built a system that leveraged YouTube data to create and distribute programming on YouTube. As such, Maker maintained a consistent feedback loop between its content, its audience, and YouTube, all tied together by data. In this regard, Maker’s data access was still disintermediated by a platform intermediary, yet it had a closer connection to digital audience data than legacy media companies, with its business being fueled by the constant infusion of platform data.

Disney acquired Maker for its digital production and distribution capabilities, yet the company’s inherent value was its close relationship with YouTube. Indeed, by bringing Maker in-house, Disney effectively bought a stake in YouTube’s data economy, giving the company enhanced access to YouTube’s information infrastructure. Furthermore, Maker provided Disney

---

<sup>129</sup> Ibid

access to actionable data, filtered through Maker's short-form programming and millennial audience, which proved highly relevant to Disney's business. Finally, Maker had built proprietary software systems to make data available and usable, which proved another value add for Disney whose legacy media organization lacked the respective software tools.

The Disney-Maker union marked the apex of Hollywood's engagement with MCNs. In 2014, the industry recorded three MCN acquisitions, which included DreamWorks Animation's purchase of AwesomenessTV (\$33 million), Disney's Maker deal (\$500+ million), and Warner Bros.' majority stake in Machinima (\$18 million).<sup>130</sup> Previously, Hollywood had made a series of strategic investments in MCNs, laying a foundation for expanded data access. The concentrated mergers & acquisitions activity around MCNs can be attributed to a key set of factors:

- *YouTube*: YouTube constitutes a juggernaut within media and entertainment, drawing hundreds of millions of viewers on a daily business. The rise of YouTube has significantly impacted the business of legacy media companies, giving audiences an alternative, and largely free, channel to engage with content. MCNs represented an opportunity for legacy media companies to succeed in the YouTube ecosystem.
- *Integration*: The majors regarded MCNs as suitable partners that would enable seamless integration with the legacy media business model. In effect, the MCN business model aligned with the majors' focus to bundle and distribute content across platforms.
- *Data Ownership*: MCNs enabled the majors to expand their data access infrastructure, for several reasons. One, MCNs offered more data volume and variety on YouTube's

---

<sup>130</sup> Warner Bros. subsequently acquired Machinima in 2016. See, Spangler, Todd, "Warner Bros. Acquires Full Control of Machinima," *Variety*, November 17, 2016, accessed June 5, 2020, <https://variety.com/2016/digital/news/warner-bros-acquires-machinima-1201920793/>

information repository. Two, they increased the majors' connection to consumers, thereby minimizing the degree of disintermediation. Three, MCNs' proprietary technology and talent helped the majors increase their data literacy: "MCNs are data-native in ways that Hollywood, which remains focused on Nielsen ratings, programming blocks, and veteran instinct, simply isn't."<sup>131</sup> And, four, MCNs enabled the majors to exert more control over data access by owning a component of YouTube's business.

Hollywood's intense, though temporary, activity around MCNs reflected the legacy media industry's growing interest in establishing more control and agency over data access. In particular, by way of data investments and acquisitions, the legacy media industry has increasingly worked to centralize data access, moving from the model of licensing multiple data sources to prioritizing data from one platform. With Maker, Disney expanded its data access in YouTube, leveraging the platform's video and social data to add volume and variety to its developing information infrastructure. Furthermore, by bringing the company's assets in-house, specifically the underlying technology and talent, the majors have worked to aggregate data capabilities, adding expertise and skill not common in the legacy media business. For Disney, Maker delivered advanced capabilities that could be scaled across the organization, including software and talent, thereby increasing data access on a structural level.<sup>132</sup> Through centralization and aggregation, the legacy media industry has taken steps to vertically integrate with the

---

<sup>131</sup> Ball, Matthew, "How YouTube MCNs are Conquering Hollywood," *Redef*, December 15, 2014, accessed October 5, 2019, <https://redef.com/original/how-youtube-mcns-are-conquering-hollywood>

<sup>132</sup> See, Patel, Sahil, "Inside Disney's troubled \$675 mil. Maker Studios acquisition," *Digiday*, February 22, 2017, accessed June 5, 2020, <https://digiday.com/media/disney-maker-studios/>

platform economy, establishing a centralized data access point (i.e., one core platform) that enabled the majors to aggregate new technological assets and skills.

Multi-channel networks represent but one example in a long lineage of Hollywood efforts to centralize data access and aggregate capabilities and expertise in-house. Film studios and television networks have long pursued acquisitions to streamline the complexity of the digital information landscape, with the end goal of assuming more ownership and control over data access. The majors positioned acquisitions as a strategic business move to enhance their digital marketing and distribution footprint, which would help them reach new audiences and unlock incremental revenue streams. At the same time, the acquisition targets all held massive data repositories, offering advanced access to platform data.

It is instructive to review a list of data deals that transpired over the course of the 21st century (table 1.9). In 1999, The Walt Disney Company agreed to acquire Internet search portal Infoseek for \$770 million.<sup>133</sup> Infoseek provided a digital portal that enabled quick discovery of content in the largely fragmented Internet landscape, effectively curating a suite of online front pages, or portals, to maximize audience reach and advertising revenue. Per Disney's logic, owning Infoseek would enable the company to centralize access to web data from millions of users and aggregate capabilities to commercialize the data into digital advertising revenue.

In similar pay-to-play fashion, News Corp. bought social media pioneer MySpace for \$580 million, reportedly beating out Viacom with a last-minute offer. For News Corp., MySpace offered an exclusive source of data access. At the time, the platform was particularly attractive given its audience of 22 million young adults, which represented an unprecedented concentration of digital users. As such, News Corp. likely envisioned MySpace's data repository as a way to

---

<sup>133</sup> See, Bicknell, Craig, "Disney Buys Into Infoseek," *Wired*, June 18, 1998, accessed June 5, 2020, <https://www.wired.com/1998/06/disney-buys-into-infoseek/>

maximize digital advertising revenue. Additionally, MySpace’s focus on audio and video content made it a viable distribution platform for News Corp. to build a dedicated direct-to-consumer platform and digital content destination down the line.

**Table 1.9**

<b>Data Acquisitions</b>			
<b>Legacy Media Company</b>	<b>Digital Media Company</b>	<b>Year</b>	<b>Data Access Rationale</b>
The Walt Disney Company	Infoseek	1999	Web and search data on films and television shows
News Corp.	MySpace	2005	Web and social data on film and television audiences
Viacom	YouTube (failed)	2004-05	Video data on film and television use
CBS Interactive	CNET Networks	2008	Web data (specifically from CNET’s TV.com destination, a database of user-generated TV show descriptions and trivia information)
Warner Bros.	Rotten Tomatoes	2011	Web and mobile data on film and television reviews from critics and users
Viacom	VidCon	2018	Web, social, and mobile data on fans of Vidcon’s YouTube influencers and talent

Not all data acquisitions came to pass. Viacom reportedly sought to acquire YouTube in an attempt to control digital video distribution and expand access to video and social data.

Similarly, Disney considered acquisitions of digital publishers BuzzFeed<sup>134</sup> and social media network Twitter.<sup>135</sup> Furthermore, while data acquisitions followed a conceptual industry logic,

<sup>134</sup> Levy, Karyne, “Disney Tried To Buy BuzzFeed, But BuzzFeed Wanted \$1 Billion,” April 28, 2014, accessed June 5, 2020, <https://www.businessinsider.com/disney-reportedly-wanted-to-buy-buzzfeed-but-the-deal-fell-apart-2014-4>

<sup>135</sup> Clifford, Tyler, “Bob Iger on Not Buying Twitter,” *CNBC*, September 24, 2019, accessed June 5, 2020, <https://www.cnn.com/2019/09/24/disney-bob-iger-on-not-buying-twitter-i-got-cold-fee.html>



they highlighted several challenges for legacy media companies that hindered integration. Most notably, digital media and technology companies operated corporate cultures that were highly differentiated from legacy media companies, complicating any consolidated and collaborative efforts. Furthermore, while the majors owned the underlying technology of their acquisitions, they had to integrate it with their existing IT infrastructure to ensure that data access could work across the newly combined organization, rather than remaining siloed, which proved highly complex and difficult. Finally, studios and networks had to find ways to sync their long-standing business model of multi-platform distribution with a centralized focus on a specific platform.

Overall, the legacy media industry's ongoing mergers & acquisitions focus on digital media and technology companies indicated a growing interest in establishing more focus and control over data access, though the industry struggled with adjusting to the differentiated industry dynamics of the digital media economy.

--- --- ---

Data acquisitions have added scale and complexity to Hollywood's data access infrastructure. In addition to licensing data from multiple platforms through strategic partnerships, the majors have increasingly focused on centralizing and aggregating data access in-house, enabling more control and agency. Indeed, the majors have pursued a model of vertical integration with the data economy, which has expanded data access in several ways. One, the majors have worked to centralize access by prioritizing data from one central audience feedback mechanism, establishing a specific data focus. Two, the majors' commitment to a central data focus has made information more actionable, highlighting information that is specifically focused on media and entertainment, leaving out any irrelevant data. Three, the majors have

focused on aggregating data technology and expertise through enhanced access, laying a foundation to access more volume and variety through one focused platform.

At the same time, data acquisitions introduced new challenges for Hollywood's data access infrastructure. One, while the focus on one platform has given the majors more control and agency over access, the data is still owned by platforms. Two, while the majors have expanded their access to data talent, they faced challenges in integrating talent within their existing organizational structures. And, three, the acquisition of digital companies required the majors to update their existing IT infrastructures, which proved costly and time-consuming.

Fundamentally, the turn toward data acquisitions signaled the legacy media industry's willingness to experiment and engage with the dynamics of the digital media economy in an effort to manage and control rather than simply license data access.

### **Direct-to-Consumer: Toward Data Incubation**

In 2014, the legacy media industry laid the foundation to reconfigure its wholesale business model by launching a preliminary platform infrastructure across the film, broadcast, and cable television industries. As previously noted, the majors had long cultivated initiatives to control distribution and develop deeper connections with consumers by acquiring digital media companies, from Internet portals to multi-channel networks and digital publishers. Most of these experiments, however, proved short-lived, temporary, and tangential, weighed down by the legacy media industry's long-standing reliance on intermediaries to facilitate distribution. Yet, due to fundamental shifts in the media landscape, driven by the rising dominance of platforms (the success of Netflix in particular), shifting consumer viewing habits, and pressured economics

for legacy media channels, the industry increasingly integrated with the logic of the digital economy to chart a direct-to-consumer future.

In the mid-2010s, legacy media companies introduced a suite of over-the-top services, circumventing the traditional distribution infrastructure and delivering content directly to consumers. Warner Bros., one of Hollywood's most iconic film studios, initiated the sea change with the launch of Warner Archive Instant, a direct-to-consumer platform that made thousands of the studio's films available online. Meanwhile, premium cable channel HBO released the beta version of HBO Now, a subscription video-on-demand (SVOD) service, which offered HBO's original and licensed programming outside of the legacy cable bundle for the first time.<sup>136</sup> Finally, CBS, one of the original Big Three television networks, entered the digital realm with CBS All Access, its own dedicated video-on-demand service, featuring a selection of old television shows and movies, live television programming, as well as new original content.

With the launch of their own platforms, the major studios and networks ventured into the direct-to-consumer economy, distributing films and television shows directly to consumers, effectively sidestepping intermediaries. While the majors continued to distribute the majority of their programming across legacy channels, keeping the traditional business model intact, they also took a first step to controlling distribution, the relationship with the audience, and the underlying information infrastructure.

The nascent adoption of platforms reflects Hollywood's strategic pivot to adapt its long-standing business model of multi-platform distribution with a direct-to-consumer approach, expanding the practice of partnering with intermediaries to managing relationships with consumers, which brought about several strategic adjustments:

---

<sup>136</sup> For a comprehensive take on non-linear, digital distribution bundles, see Lotz, Amanda D., *The Television Will Revolutionized*, 2nd Edition (New York: New York University Press, 2014).

- *Distribution*: The legacy media industry updated its distribution model, extending beyond multi-platform distribution and windowing to exclusive releases across owned and operated platforms. As a result, the majors created a foundation to generate, capture, and own data from their content, establishing “a proximity to customers that doesn’t have intermediaries.”<sup>137</sup>
  
- *Economics*: The launch of direct-to-consumer platforms constituted a new, experimental way to monetize content for the majors, moving from a hitherto exclusive focus on advertising and transactional licensing revenue to an on-demand business model that generates revenue through subscriptions.<sup>138</sup> In this sense, the majors expanded their economic focus to subscription revenue and associated subscriber numbers, in addition to the aggregate statistics associated with television’s advertising model. As such, the economic model prioritized data as a key component of the legacy media business.
  
- *Technology*: With the launch of platforms, the majors effectively operated as digital media and technology companies. Platforms introduced a new technological infrastructure that required maintenance, supervision, and investment. In this regard, the majors developed a set of approaches to manage the technological platform infrastructure, from managing it themselves (e.g., CBS All Access had a dedicated internal technology team) to partnering with a specialized technology company (e.g., HBO partnering with the streaming technology company BAMTech) to acquiring

---

<sup>137</sup> Belloni, Matthew, “Bob Iger Talks Disney’s Streaming Service,” *Hollywood Reporter*, September 20, 2018, accessed June 5, 2020, <https://www.hollywoodreporter.com/news/bob-iger-disneys-streaming-service-james-gunn-star-wars-slowdown-1145493>

<sup>138</sup> The subscription model took on two forms, an ad-free and an ad-supported model. See, Lotz, 2017

technology players (e.g., Disney’s 2017 acquisition of BAMTech, Paramount’s 2019 acquisition of Pluto TV, Comcast/NBC Universal’s 2020 acquisition of Xumo).

With the launch of direct-to-consumer platforms, the major studios and networks effectively adapted the business model of platforms. This new business focus enabled the majors to remove the long-standing reliance on traditional intermediaries (e.g., distributors, exhibitors, retailers), begin developing a direct relationship with consumers, and, as a result, own the underlying access to data. Indeed, by launching their own platforms, the majors laid a foundation to build a data access infrastructure from the ground up. Previously, film studios and television networks licensed select data on their content from platforms. With their own platforms, the majors essentially cut out the intermediaries to incubate their own data.

The direct-to-consumer shift revised the dynamics of data access. While the majors have continued to distribute content across multiple platforms, generating and licensing data through a suite of data access systems, they equally began distributing content across their own platforms, incubating first-party data and expanding overall data access. Effectively, the majors moved from managing a multitude of data licensing partnerships and data access systems to owning the underlying data infrastructure. In this regard, the investment in direct-to-consumer platforms enabled the majors to exert more control and agency over data access.

The legacy media industry’s shift to the direct-to-consumer business was not an abrupt transition, yet the result of a gradual, long-gestating development. Given the differentiated dynamics of distribution, economics, and technology, the majors carefully negotiated the shifting digital media landscape, carefully weighing any potential impacts on the legacy media business. As such, the industry underwent a suite of interrelated phases of direct-to-consumer shifts: *Exploration, Experimentation, and Institutionalization.*

While this type of periodization is helpful, it is important to note that the three phases do not constitute distinct, self-contained categories. Rather, they represent fluid concepts that carry and communicate guiding principles of the legacy media industry's maturation of direct-to-consumer platforms, namely a growing tendency to build direct relationships with consumers, the gradual integration of technological platform infrastructure, and increasing degrees of ownership and control<sup>139</sup> (table 1.10).

- *Exploration*: The majors explored early direct-to-consumer models during the dot.com boom. Nascent streaming technology enabled the transmission of online video files through websites, even though the Internet lacked the scale and capacity to reach a mainstream audience prior to the rise of cloud computing. During this time, Hollywood's most pressing competitor was Microsoft, which, driven by CEO and founder Bill Gates' axiom that content is king, was heavily engaged in the *Televisionspace Race*<sup>140</sup> with the patented release of set-top boxes to stream content directly into living rooms. While streaming technology proved premature, unable to deliver on the ambitious vision of the time, the legacy media industry focused on digital downloads as an alternative option, making movies available for home viewing and retaining control over the experience. Similar to modern digital transactions via electronic sell-through (EST), consumers could purchase access to content in a hosted virtual environment, for a specific time period.

---

<sup>139</sup> For more information the history of media industries, see, Alisa Perren & Jennifer Holt, eds., *Media Industries Studies: History, Theory, and Method* (Hoboken: John Wiley & Sons, 2011) and Albarran, Alan B., ed., *The Social Media Industries* (New York: Routledge, 2013).

<sup>140</sup> See, Rose, Frank, "The Televisionspace Race," *Wired*, April 1, 1998, accessed June 5, 2020, <https://www.wired.com/1998/04/mstv/>

With the dominance of DVDs and Blu-Rays, direct-to-download became a prominent digital feature that was eventually overtaken by cloud-based streaming technology.

- *Experimentation:* The legacy media industry initiated platform experiments as cloud computing entered its mature commercial stage and platforms like Facebook, YouTube, and Netflix rose to prominence. During this era, a select group of film and television companies, notably NBCUniversal, NewsCorp. (long-time owner of 20th Century Fox and Fox Television prior to their 2019 acquisition by Disney), and The Walt Disney Company formed a streaming consortium called Hulu, an initially ad-supported, later subscription-based video-on-demand platform. Hulu's nascent product enabled the next-day online viewing of traditional television programming and allowed the majors to test out the technology and business of the developing platform economy. Over time, the majors took lessons from Hulu to build niche platforms, small-scale video-on-demand services that focused on select programming, made available through owned platforms. Niche platforms represented experimentation at scale, a way to roll out platform technology to avid audiences with an affinity for genre programming. Early niche platforms prioritized popular genre content like anime and horror.
- *Institutionalization:* The majors explored a shift to a dedicated direct-to-consumer platform model by creating proprietary services with exclusive access to their content libraries. While the phases of exploration and experimentation represented add-on business models, designed to complement the traditional legacy media business, the pivot to platform institutionalization signaled an emerging attempt to adapt the business logic

of the digital economy, develop direct relationships with consumers rather than intermediaries, and unlock data access as a form of capital.

**Table 1.10**

<b>Data Incubation</b>				
<b>Phase</b>	<b>Service</b>	<b>Year</b>	<b>Legacy Media Involved</b>	<b>Description</b>
<b>Exploration</b>	CinemaNow	1999	Lionsgate (investor)	Streaming website
	MovieBeam	2002	Buena Vista Datacasting (Disney subsidiary)	OTT streaming service that required custom set-top box and enabled rentals and electronic sell-through (EST)
	Movielink	2002	Joint venture of Paramount, Sony, MGM, Universal, and Warner Bros	Video-on-demand service that made movies available for download
	NBC Direct	2008	NBC	Video-on-demand (VOD), direct-to-download service
<b>Experimentation</b>	Crunchyroll	2006	Otter Media (joint venture of Chernin and AT&T)	OTT streaming service
	Hulu	2008	NewsCorp., NBCUniversal, Disney, Time Warner	Ad-supported video-on-demand service; initially post-airing online catch-up, then dedicated streaming service
	Shudder	2015	AMC Networks	OTT streaming service
	Seeso	2016	NBCUniversal	OTT streaming service
<b>Institutionalization</b>	Warner Instant	2014	Warner Bros.	OTT streaming service
	HBO Now	2014	HBO	OTT streaming service
	CBS All Access	2014	CBS	OTT streaming service
	Showtime	2015	Viacom	OTT streaming service
	Disney+	2019	Disney	OTT streaming service



	HBO Max	2020	WarnerMedia (subsidiary of AT&T)	OTT streaming service
	Peacock	2020	NBCUniversal	OTT streaming service

The launch of direct-to-consumer platforms marked an expansion of Hollywood’s data access infrastructure from generating and licensing to owning access. Indeed, the majors have consistently managed a spectrum of strategies to integrate with the information infrastructure. On the one hand, they have negotiated access through multiple licensing partnerships, leveraging access systems from platforms, startups, and upstarts. On the other hand, they have increasingly captured their own first-party data, owning access through the direct-to-consumer infrastructure of their platforms. Yet, while the industry has maintained a balance of these two approaches, they effectively represent differentiated ways to access data, with a suite of implications.

Licensing partnerships enabled data access in a specific way. First, they allowed the majors to access a volume of data from across platforms, ensuring a scale of information. Furthermore, they gave the majors exposure to a variety of information, blending select second-party data from platforms with third-party data from startups. As such, the majors gained access to a spectrum of information, covering multiple data types (e.g., web, social, mobile, video). For example, by partnering with Facebook, the majors can access social, mobile, and video data on their audience’s interaction with their content. Additionally, by leveraging the services of startups, they could consolidate select data across multiple platforms, including publicly available information, in a cohesive way. Finally, licensing partnerships provided convenient and seamless access that did not require any significant technological investment on part of the majors. By contrast, licensed access is controlled by platforms and thus invariably limited to the

information that is made available. In this sense, the information is at once limited in availability, usability, and actionability.

Direct-to-consumer platforms enabled the majors to incubate, control, and own data access directly. As such, they were able to access all the information available and, by consequence, gain a more holistic picture of the data at their disposal. As a result, owned first-party data proved more actionable for the majors, enabling them to make use of the data in ways that are not possible with licensed data, where the majority of the information is owned by platforms. Yet, direct access to data on the backend also raises new challenges for the majors. In particular, the majors no longer have access to the volume and variety of information provided by platforms with hundreds of millions of users. Indeed, legacy media platforms effectively started with comparatively small user bases and require constant investment to generate scalable data output. At the same time, the data generated by platform users is highly focused and narrow. While platforms provided access to various data types, legacy media platforms exclusively focused on video data. Furthermore, even though the majors owned access through their own platforms, intermediaries still remained involved. To increase the scale of their audience, the majors have maintained distribution partnerships with platforms, effectively sharing the associated revenue and data. Amazon's Channels platform provides an interesting use case in this context. Amazon Channels allows consumers to sign up for direct-to-consumer platforms from legacy media companies, such as HBO Now and Showtime. Consumers can watch HBO programming directly through Amazon, which provides legacy media companies with a scalable promotional marketing channel and enables Amazon to take a share of the revenue while capturing the entirety of the underlying consumer and usage data.

Overall, while minimizing the volume and variety of data available from licensing partnerships, the launch of platforms has effectively enabled the majors to further centralize and aggregate access, increasing their level of control, agency, and focus over data access. Owning data access is a way for the majors to reverse the disintermediation of the legacy media industry's established distribution infrastructure, to build close relationships with consumers that facilitate a new business model where studios and networks capture, consolidate, and commercialize data on top of content.

Ownership of data access has become a central undercurrent of the shifting media landscape, which, according to former Disney CEO Bob Iger, "is increasingly defined by direct relationships between content creators and consumers."<sup>141</sup> Over the course of the 21st century, all the major studios and networks have launched proprietary direct-to-consumer platforms in an effort to own a direct relationship with their audience and control access to the associated data and underlying information infrastructure. This direct-to-consumer platform investment has fundamentally revised the data access dynamics in media and entertainment. Don Buckley, the former Chief Marketing Officer of Showtime, aptly summarized the new industry environment (my emphasis):

**"For a long time, we had very little access to data. It belonged to the distributors. Since we launched our streaming product in 2015, we're much better equipped with data and have built out the systems required to analyze and activate any conclusions we reach based on the data. It's a great base. It's many, many millions of people. Their viewing behavior becomes a prototype for the rest of the network for whom we don't have data."**<sup>142</sup>

---

<sup>141</sup> Proença Santos, Amanda, "Disney to Start Own Streaming Services, Remove Content From Netflix," *NBC News*, August 8, 2017, accessed June 5, 2020, <https://www.nbcnews.com/pop-culture/tv/disney-start-own-streaming-services-remove-content-netflix-n791001>

The statement illustrates the legacy media industry’s increasingly actionable access to data through platform ownership, an ability for the majors to not only reach a substantial volume of consumers with their programming, but leverage the resulting data across their organizations. At the same time, this newly-enabled data access has created an industrial environment where the majors operate as something akin to platform-based data companies, not merely accessing information, but increasingly making use of it.

--- --- ---

Over the course of the 21st century, Hollywood has developed, implemented, and managed a dynamic infrastructure to access a growing volume and variety of data in the digital landscape. Effectively, the legacy media industry has moved from lacking any data access to establishing a system where the majors effectively generate, license, and own data access.

- *Data Generation:* The majors distributed films, television shows, and related digital content across a growing set of platforms, generating consumer and usage data at scale. While digital distribution generated a high volume and variety of data, the information was fragmented across the digital landscape, with limited access, controlled by platforms, and a manual way of accessing information via the frontend of the digital sphere.
- *Data Licensing:* The majors developed partnerships with technology companies (i.e., platforms, startups, upstarts) to license data access. This partnership model enabled the automation of access, with technology companies using cloud computing capabilities to

---

<sup>142</sup> Otterson, Joe, “TV’s Top Marketing Execs Break Down the Power of Data at Variety’s Massive Summit,” *Variety*, March 21, 2018, accessed June 5, 2020, <https://variety.com/2018/tv/news/tv-marketing-variety-massive-summit-1202733153/>

automatically and conveniently deliver data to the majors via easy-to-use software interfaces that required little technological infrastructure on the part of the legacy media industry. Furthermore, the majors gained access to select platform data (via platforms) as well as a cross-section of data from across multiple platforms (via startups and upstarts). At the same time, the majors had little control and agency over access, relying entirely on a suite of intermediaries to tap into the information infrastructure.

- *Data Ownership*: The majors, through a string of strategic digital media investments and acquisitions, laid a foundation toward data ownership, effectively incubating the technological infrastructure needed to access their own data. This economic activity, however, presented several challenges of integration that affected access. The majors further launched their own direct-to-consumer platforms, which enabled data ownership and access to actionable first-party data. While the majors established a central focus on their own audience and content information, the new platform focus did not have the same volume and variety of data offered through licensing partnerships.

Through the interrelated strategies of data generation, licensing, and ownership, the majors have developed a system that enables a spectrum of access, from negotiating partnerships with platforms, startups, and upstarts to license select data at scale to cultivating proprietary access within their own platform infrastructure. Throughout, film studios and television networks have increasingly adopted and adapted the business logic of the direct-to-consumer economy, expanding the legacy media industry's long-standing focus of content output with data input. Indeed, the majors have established a suite of mechanisms to access an increasing volume and variety of data, ranging from third-party to second-party and first-party information. While the

resulting information infrastructure is complex and complicated, at once comprehensive and limited, broad and narrow, the legacy media industry has effectively enacted a systematic data access strategy. As such, the industry has not only set an approach to effectively negotiate data access, but increasingly take control of the process. In this regard, the majors are no longer disintermediated from the direct line to consumers, cut off from the flow of digital information. Instead, they have integrated with the access pipeline of the digital landscape, thereby unlocking a new information infrastructure.

The development of Hollywood's new information infrastructure is not an isolated incident. It forms part of a larger value chain where data has come to function as a new form of capital in the media and entertainment industry. Data access is a core element of the value chain, yet only represents an incremental piece. With the traditional distribution setup moving into a two-way feedback loop between the majors and consumers, consumer and usage data is generated, captured, and consolidated consistently, creating a wealth of information. Yet, in this media environment, "[...]the primary issue isn't how you get all this data; it's what you do with it when you get it."<sup>143</sup>

As the legacy media industry worked to establish access to data, the majors simultaneously had to consider the implications of access on legacy media organizations. With data as an everyday industrial reality that factors into organizational decision-making, rather than an abstract concept separated from the business, studios and networks had to increasingly negotiate the industrial dynamics of *making sense* of data.

---

<sup>143</sup> Wallenstein, 2013

## Part II – Data Analysis

## Context

### Data and Decision-Making

“I find the big fundamental difference between Silicon Valley and Hollywood is quant and qual.”<sup>144</sup>

Ted Sarandos, Chief Content Officer at Netflix

In an interview with GQ magazine, published shortly after the release of *House of Cards* (Netflix, 2013-18), Netflix’s Chief Content Officer Ted Sarandos crystallized the streaming platform’s long-term business plan, to operate as a dedicated media company that produces distinct original programming and competes with the media incumbents. He declared succinctly, “The goal is to become HBO before HBO can become us.”<sup>145</sup> Drawing a parallel to marquee cable programmer HBO, Sarandos signaled Netflix’s ambition to transition from content licensing and distribution into production, build its offering into a distinctive media brand, and establish creative clout in Hollywood. Journalists, analysts, and pundits have frequently referenced this statement in charting Netflix’s development, framing the company’s trajectory as a convergence with Hollywood’s legacy media industry.

There is another conceptual way to look at the proposed Netflix-HBO constellation, which has received considerably less attention. While Sarandos put the spotlight on Netflix’s aspiration to become a media company, he also intimated an alternative vision, sketching the

---

<sup>144</sup> a16z Podcast, “The Internet of Taste, Streaming Content to Culture,” February 18, 2018, accessed October 5, 2019, <https://a16z.com/2018/02/18/content-culture-digital-sarandos-summit/>

<sup>145</sup> Hass, Nancy, “Is Netflix the Next HBO?,” *GQ*, January 29, 2013, accessed August 20, 2019, <https://www.gq.com/story/netflix-founder-reed-hastings-house-of-cards-arrested-development>



scenario of HBO turning into Netflix to become a streaming platform in the direct-to-consumer business.<sup>146</sup> Effectively, Sarandos outlined two industrial shifts, not one. The first takes an outside-in view, highlighting a new path for Netflix as it makes inroads into media and entertainment, building a vision of a new entrant taking on the incumbents. The second view considers a wide-ranging industrial pivot for the entire legacy media industry, with HBO (read: all of Hollywood) extending into Netflix's digital, over-the-top territory. In this regard, Sarandos set up an interdependent dynamic of convergence, where platforms aspire to move into the media business, while media companies, in turn, adopt the business logic of platforms.

Following Sarandos' shrewd declaration, Netflix made this concept of industrial convergence the central element of its PR strategy, modifying, iterating, and challenging the story as the business evolved. At the Aspen Institute, Sarandos further clarified the dynamic between Netflix and HBO, emphasizing that the majors were unlikely to replicate Netflix's platform model due to the divergence in how both industries operate. He argued that Netflix's core business focus, rather than managing production, is to manage consumer relationships and data (my emphasis):

“Our **DNA** is basically processing and **managing the relationships** with 60M+ people in 67 countries around the world, **direct relationships with all those consumers**. HBO doesn't have any of those relationships.”<sup>147</sup>

---

<sup>146</sup> HBO moved into the D2C business with the launch of HBO Now in 2014 and HBO Max in 2020. See also, Schatz, Thomas, “HBO and Netflix: Getting Back to the Future,” *Flow*, January 20, 2014, accessed August 5, 2019, <https://www.flowjournal.org/2014/01/hbo-and-netflix-%E2%80%93-getting-back-to-the-future/>

<sup>147</sup> Aspen Institute Podcast, “The New Golden Age of Television,” September 10, 2015, accessed October 5, 2019, <https://www.aspeninstitute.org/podcasts/aspen-ideas-to-go-podcast-katie-couric-netflix-new-age-television/>

In effect, Sarandos constructed a narrative of differentiation, arguing that HBO, and Hollywood by proxy, lacked the capabilities and direct consumer relationships that define the business logic of platforms. More pointedly, he identified the ability to manage data as a core difference between platforms and legacy media companies.

Sarandos grounded this differentiation in the way both industries operate. In an interview with venture capitalist Marc Andreessen, he said, “I find the big fundamental difference between Silicon Valley and Hollywood is quant and qual. The whole efficiency-driven thing is very Silicon Valley. And the whole quality-driven thing is very Hollywood. And rarely do those things meet [...]. Hollywood is all about relationships. There is no quant.”<sup>148</sup> For Sarandos, Netflix operates with a *quant* logic, drawing on data and analytical thinking to inform corporate decision-making. Hollywood, by contrast, operates with a *qual* logic, which emphasizes “impressionistic modes of decision-making,”<sup>149</sup> relying on creative instinct, talent, and relationships to make decisions. The quants think in data (i.e., rational evidence and reasoning), while the quals are driven by intuition (i.e., implicit feeling and honed instinct)<sup>150</sup> (table 2.1)

**Table 2.1**

The Spectrum of Industrial Decision-Making	
<i>Hollywood Impressionism</i>	<i>Silicon Valley Analytics</i>
Qualitative model driven by gut instinct, institutional knowledge, talent relationships, and creative intuition	Quantitative model driven by scientific rationalization, technical analysis, and empirical evidence

---

<sup>148</sup> a16z Podcast

<sup>149</sup> Napoli, 2014

<sup>150</sup> The quant/qual dichotomy is conceptual and dynamic rather than practical and absolute. Digital platforms like Netflix emphasize data in decision-making, yet equally employ human capital to make adjustments and changes. Hollywood has long relied on institutional knowledge and instinct, yet equally draws on audience research.

Sarandos' perspective offers a useful foundation to examine the role of data analysis in media and entertainment. Sarandos frames Netflix's approach to analyzing data as the company's core business differentiator in that it translates its information access into business insights that inform corporate decision-making. Sarandos evokes an organizational culture that thinks and operates analytically, relying on data in a media environment historically characterized by alternative, more subjective and impressionistic forms of decision-making.

Indeed, Netflix has built its corporate brand around making data-centric decisions. The company's leadership team has touted its data capabilities in countless interviews, dating back to its IPO roadshow in the early 2000s, inspiring coverage that discursively links Netflix and data into a conceptual entity. Netflix has consistently emphasized data-driven decision-making as core to its corporate DNA, promoting an organizational culture grounded in the analysis of data. Consider the introduction to Netflix's corporate research website, which outlines the company's corporate cultural manifesto, framed as a commitment to the process of *analytics* (my emphasis):

**“Netflix has been a data-driven company since its inception. Our analytic work arms decision-makers around the company with useful metrics, insights, predictions, and analytic tools so that everyone can be stellar in their function. Partnering closely with business teams in product, content, studio, marketing, and business operations, we perform context-rich analysis to provide insight into every aspect of our business, our partners, and of course our members' experience with Netflix. Our portfolio of work involves diving into large, complex data to answer ambiguous business questions. We work cross-functionally across business domains to discover and assess new opportunities, create new business metrics to measure success, and inform prioritization. We also strive to make analytic tools self-service to make data and insights even more accessible. We provide analytic thought leadership that provides data-grounded perspective to help Netflix think critically and differently about its business to improve our service.”<sup>151</sup>**

---

<sup>151</sup> Netflix Research, “Analytics: Driving Insights from Data,” *Netflix*, accessed October 5, 2019, <https://research.netflix.com/research-area/analytics>

The text offers a self-reflexive perspective on how Netflix conceptualizes data-driven decision-making. First, Netflix represents a data-driven company that derives its corporate identity from data analysis. Furthermore, Netflix analyzes data to create, share, and measure insights in support of decision-makers. Finally, Netflix's business logic is grounded in data analysis, with the entire company basing decisions on data-driven insights. Netflix displays a *culture of analytics*,<sup>152</sup> a corporate-industrial mindset that prioritizes data analysis as a way to find patterns, create insights, and measure impact in order to inform decision-making in the media and entertainment business.<sup>153</sup>

Netflix's focus on analytics and rationalization is emblematic of Silicon Valley's archetypal business logic, which prizes data analysis as the core lever of the digital economy. Platforms collect, store, and access ever-increasing amounts of data on a daily basis, yet they create business value through the constant analysis of that data. Value is derived from a procedural feedback loop that mines data via analytical practices and procedural layers. First, platforms conduct technical analysis of data, employing a range of scientific methodologies, from statistics to mathematics and computer science, to distill data into insight, creating information that is relevant to their business. This practical approach requires substantial scientific effort and quantitative logic. For example, a subscription-based video platform service like Netflix considers behavioral patterns that drive viewership across its content portfolio and video service. For an ad-supported platform like Facebook, it involves decoding the audience

---

<sup>152</sup> The concept of *data analytics* comes from management sciences and draws on statistics, mathematics, and computer science to identify meaningful patterns in data, thereby providing a rationale for decision-making.

<sup>153</sup> For a brief description of data analytics in the legacy media industry, see Schlesinger, Scott, "Using Analytics to Predict Hollywood Blockbusters," *Harvard Business Review*, October 11, 2012, accessed June 5, 2020, <https://hbr.org/2012/10/using-analytics-to-predict-hollywood-blockbusters>, quote: "It's all about identifying patterns in past data, melding them with current data points that are readily available, and then taking action to improve business performance."

segments and preferences on its service to develop offerings for advertisers. In the case of Amazon, analysis focuses on what drives product transactions in order to maximize sales.

As part of this technical analysis, platforms deploy technology to analyze large sets of data. In particular, they leverage software applications to translate access to massive amounts of data into specific relevant insights through what Netflix labels “analytic tools.” The use of analytical software further enables the visual packaging and communication of insights across the organizational structure, ranging from Excel spreadsheets to charts and visual data interfaces, ensuring that the company can execute against data-driven insights by “arming [sic] decision-makers around the company with useful metrics, insights, predictions, and analytic tools [...]”<sup>154</sup>

Finally, platforms draw on scientific methodology and software technology to foster, promote, and maintain an organizational culture that makes decisions based on data-driven insights. In this cultural environment, data analysis drives decision-making, with insights being shared across the organization. Analytics thus describes a layered techno-cultural process of creating, packaging, and communicating data-driven insights (table 2.2).

**Table 2.2**

Data Analysis Model		
<i>Technical-Practical</i>	<i>Technological</i>	<i>Cultural-Organizational</i>

Hollywood, by contrast, is generally seen as lacking an analytical organizational culture and said to draw on decades of institutional knowledge, talent relationships, and creative expertise to drive decision-making. In an environment where “nobody knows anything,”<sup>155</sup> the

---

<sup>154</sup> Netflix Research

<sup>155</sup> Goldman, William, *Adventures in the Screen Trade: A Personal View of Hollywood and Screenwriting* (New York: Warner Books/Grand Central Publishing, 1983).

ability to dream up the winning formula was highly sought-after capital. As the Wall Street Journal's Ben Fritz aptly summarized, "Hollywood has ignored data for so long, it's amazing. It has really been a business driven by the guts and the tastes of the people who run studios."<sup>156</sup>

Indeed, the corporate-organizational culture of studios and networks is generally seen as overly impressionistic and gut-driven. While shorthand and narrow, this characterization still reflects fundamental truths of industrial reality, though it requires some nuance. First, Hollywood's business is not solely grounded in the idea of creative genius, though heavily weighted in that direction from a PR perspective.<sup>157</sup> Rather, the industry has constructed a system of multiple levers to drive value creation, including institutional knowledge, talent relationships, the marketing of creative individuals, controlled distribution, a model of scarcity, and cross-platform monetization.<sup>158</sup> Second, the majors have never operated in a hermetically-sealed creative environment, entirely shut off from external information, but have cultivated partnerships with market research firms to gather data on content reception, tracking, and performance, using results from focus groups, polls, and surveys to adjust decision-making.<sup>159</sup> And, third, the industry has a track record of change-makers attempting to shift organizational behavior and change the way media companies operate, to varying degrees of success.<sup>160</sup>

---

<sup>156</sup> Knowledge@Wharton Podcast, "Can Hollywood Survive Streaming Services?," *Wharton School of Business*, May 4, 2018, accessed June 4, 2020, <https://knowledge.wharton.upenn.edu/article/can-hollywood-survive-streaming-services/>

<sup>157</sup> See, McDonald, Paul, *The Star System: Hollywood's Production of Popular Identities* (London: British Film Institute, 2005).

<sup>158</sup> See, Wasko, 2003

<sup>159</sup> See, Handel, Leo A., "Hollywood Market Research," *The Quarterly of Film Radio and Television*, 7 (3), 1953, 304-310.

<sup>160</sup> See, Balio, 1985

Yet, over the course of the digital era, the tides have shifted in the media and entertainment industry, giving analytics increasing prominence over established practices of decision-making. In effect, data's role in shaping and driving corporate decision-making has gained increasing momentum in the 21st century, with data analysis becoming a "core competency in all facets of the entertainment biz,"<sup>161</sup> and a critical factor in creating, managing, and extending digital business value. The rise of Netflix and data-driven platforms has generally been singled out as the cause celebre of this phenomenon, yet in fact represents but one symptom of a larger industrial shift towards rationalization, marked by a complex set of cultural, economic, and technological developments:

- *Data Scale*: New digital technologies, cloud computing in particular, led to an ever-increasing scale of information, creating a pressing need for companies to find, identify, and extract relevant insights.
- *Data Analysis Companies*: Technology companies invested in analytics products and services designed to translate information into insight, deploying various methodologies of analysis to make sense of data.
- *Data Analysis Technologies*: The development of software applications increasingly enabled the automatic analysis of data, automating the process of distilling general information into specific insight.

---

<sup>161</sup> Wallenstein, 2013

- *Data Analysis Skills*: Data literacy, the ability to read, translate, and understand data as a form of insight, emerged as a key skill of the direct-to-consumer economy.<sup>162</sup> Indeed, economist Hal Varian deemed the category of data analyst the sexiest job of the 21st century.<sup>163</sup> The rise in data created significant corporate investments in data analysis as well as a growing demand for data analysis education in universities.
  
- *Business Valuation*: Platforms like Amazon, Facebook, and Google have increased their stock value through the use of data, seeing exponential returns in the market at running valuations. Additionally, specialized analytics companies saw growing demand among a steady increase in data, leading to substantial venture funding, ongoing mergers and acquisitions, rising valuations on Wall Street, and successful startup IPOs.

In an era where shareholder returns, organizational culture, and cultural production have become increasingly driven by data analysis, Hollywood’s reliance on intangible creativity appeared somewhat outmoded in the eyes of commentators, professionals, and competitors.<sup>164</sup> The result was a media environment that became more and more defined by, as Napoli notes, “a steady although contentious process of ‘rationalization,’ in which impressionistic modes of decision-making are replaced by more data-driven, analytical approaches.”<sup>165</sup>

---

<sup>162</sup> See, LinkedIn Workforce Report, August 10, 2018 and World Economic Forum Report, “Data Science in the New Economy,” *World Economic Forum*, July 2019.

<sup>163</sup> See, Davenport, Thomas H. & Patil, D.J., “Data Scientist: The Sexiest Job of the 21st Century,” *Harvard Business Review*, October 2012, accessed October 5, 2019, <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

<sup>164</sup> See, Bilton, 2017

<sup>165</sup> Napoli, 2014



Yet, the industrial incursion of data analysis has neither replaced nor suppressed the legacy media industry's traditional modes of operation. The majors have continued to base corporate decisions on creative ideation, institutional knowledge, and talent relationships, though the industry's analytical undercurrent, always present in its long-standing engagement with audience research, has increasingly come to the fore.<sup>166</sup> In this regard, studios and networks effectively managed a balancing act of decision-making variables in an expanding toolkit. As one chief executive observed, “[d]ata helps you make more intelligent decisions [...]. Having great creative instincts remains at the core of what we're doing. Data is just another tool that we're going to use.”<sup>167</sup> Similarly, another executive highlighted the primacy of instinct and intuition in considering quantitative data input (my emphasis):

“Our gut knows better than any other part of us what's good and what's right. There's obviously a business side to it and you need to be focused and analytical, to pull from as many inputs and to have as many insights as you can. **But then you have to forget all of that and hope it synthesizes in the back of your brain somewhere.**”<sup>168</sup>

Given the industry's long-standing prioritization of creativity and institutional knowledge, decision-makers have shown a discreet ambivalence towards data analysis, carefully considering the implications for Hollywood's established organizational cultures, characterized by a lack of familiarity and comfort with the data-driven, analytical practices of by platforms. As such, the

---

<sup>166</sup> Per one analyst: “[n]one of these companies have the ostrich syndrome - none of them are putting their heads in the sand. They're forward-thinking and they're aggressively pushing their teams to make sense of this world.” See, Lang, Brent, “Big Media Gambling on Future Digital Riches (Study),” *The Wrap*, January 28, 2014, accessed June 6, 2020, <https://www.thewrap.com/media-entertainment-companies-will-lose-money-set-digital-future-study/>

<sup>167</sup> Littleton, Cynthia, “Viacom's Philippe Dauman Talks New Metrics, Creative Tools and Wall Street's ‘Short-Term-ism’,” *Variety*, November 4, 2015, accessed June 5, 2020, <https://variety.com/2015/biz/news/viacom-philippe-dauman-big-data-summit-1201633402/>

<sup>168</sup> Marketplace Podcast, “The Business of TV in 2020,” *Marketplace*, January 22, 2020, accessed June 12, 2020, <https://www.marketplace.org/shows/marketplace/the-business-of-tv-in-2020/>

industry has largely framed the emergence of data analytics as “just another tool,”<sup>169</sup> another variable to consider, an additional layer to mine in thinking about cultural production.

Similarly, Netflix and other tech-turned-media platforms such as Amazon, Google, and Apple have shown more consideration of Hollywood’s impressionistic approach, at times downplaying the impact of analytics in favor of gut feeling in an effort to organically integrate into the media and entertainment marketplace. Indeed, Ted Sarandos has repeatedly adjusted Netflix’s approach to decision-making, initially noting that “[i]t’s 70 percent gut and 30 percent data, with the data on top”<sup>170</sup> before updating the formula to simply 70/30 and later 80/20 in favor of instinct over data. The result is a convergence of operational modes of decision-making, expressed through dynamically shifting attitudes and perspectives on data analysis, embedded in changing organizational cultures.

In this sense, data analysis has manifested as another factor of consideration in decision-making for Hollywood, taking on the form of another lever and value creation driver in the legacy media industry. Indeed, the industry has neither fully embraced nor rejected the rise of data analysis. Instead, the majors have negotiated the integration of different forms of decision-making, insisting on institutional knowledge and creative intuition, while actively adopting analytical practices from platforms and adapting to a confluence of developments driven by data analysis. Thus, it is vital to frame the integration of data analysis as a complex, complicated, and all-consuming process, marked by cultural, economic, and technological layers. Julie Rieger, a former data executive at 20th Century Fox, effectively distilled this phenomenon in reflecting on

---

<sup>169</sup> Hazelton, 2016

<sup>170</sup> Wu, Tim, “Netflix’s Secret Special Algorithm is a Human,” *The New Yorker*, January 27, 2015, accessed June 5, 2020, <https://www.newyorker.com/business/currency/hollywoods-big-data-big-deal>

her decade-spanning work in applying data analysis in Hollywood's legacy media industry. After working across film and television and partnering with platforms and technology companies on a variety of projects, she noted, "[w]hile data has revolutionized our business, it's not enough to just *have* data. On its own, data doesn't tell you how to solve problems and change the trajectory of your business. You need a vision for how to use that data."<sup>171</sup>

Part I examined how the formation of an expansive and layered information infrastructure enabled the majors to access data in the evolving digital economy. Part II aims to build on this development by exploring the usage and impact of data on corporate decision-making in the legacy media industry. Specifically, it traces the manifestation of data analysis practices across the organizational structures of film studios and television networks in form of an emerging culture of analytics. The focus lies on a careful negotiation of two decision-making models, the analytical and the impressionistic, the quant and the qual, new and old media. As such, it proposes that data analysis neither replaced nor supplanted, but integrated established modes of decision-making with data-driven rationales, making data analysis an incremental factor in thinking about the business of media and entertainment. The result is a conceptual model that frames the culture of data analytics as grounded in the convergence of quant and qual, a complex integration of analytical and impressionistic modes of decision-making, administered by corporate actions and reactions across the organizational structures of legacy media companies.

In this regard, Part II constitutes a continuation of Part I's data access narrative, tracing a set of industrial strategies and developments that leveraged data access into business insight by way of data analysis. However, data access and analysis are not linear processes, but part of a non-linear, interdependent data value chain.

---

<sup>171</sup> Think with Google, 2018

To this end, Part II aims to address the following questions on Hollywood's integration of data analysis as an industrial mode of operation:

- How did the legacy media industry establish the expertise and capabilities to translate data into insights?
- How did the majors integrate data analysis into established modes of industrial operation?
- How did data analytics manifest as an incremental driver of decision-making in the organizational cultures of film studios and television networks?

The answers to these questions require looking at data analysis as a layered process of cultural, economic, and technological considerations rather than a linear set of practical applications.

### **A Framework for Data Analysis**

Data analysis is generally understood as output-focused, generating specific contextual insights through the application of scientific methodologies. For example, Netflix's programming decisions are shaped by the output of data analysis, a compilation of insights expressed through charts, graphs, decks, and other corporate texts, indicating viewing trends among audience segments that signal the viability of particular programming for Netflix. The industrial logic behind this form of decision-making is of particular consideration here, while the underlying details of technical analysis are not. To this end, for the purpose of this project, data analysis is conceptualized as procedural, rather than technical, involving a variety of corporate strategies and industrial developments that shape Hollywood's approach to decision-making.

Accordingly, Part II does not focus on the technical dimension of data analysis. It does not examine statistical or mathematical methods of analysis and associated outputs. Instead, it

traces the industrial factors that have enabled and, to some degree, institutionalized analytical forms of decision-making in the legacy media industry:

- *Tech Talent*: Hollywood has seen the emergence of a new workforce segment, with professionals versed in data analysis entering the industry. These professionals tend to have a background in quantitative fields such as economics, statistics, and data science and their roles ask for analytical thinking, proficiency with data analysis software, and experience in translating data into insights. The majors further created new departments focused on data analysis within the broader org charts of media organizations.
- *Data Analysis Tools*: The legacy media industry has built an infrastructure of software applications that automate the analysis of data, facilitate the classification of data as insight, and enable the communication of data insights. The majors have further developed partnerships with data analysis vendors, from startups to upstarts and established players, to implement data analysis within organizational cultures.
- *Data Metrics*: Film studios and television networks had to negotiate the rise of new data metrics, conceptual models that describe data as insight. Historically, the legacy media industry has focused on a small set of metrics, mainly box office and ratings, yet the digital era led to a substantial increase in metrics, which required the majors to understand, value, and commercialize new sets of information.

The formation of data-driven decision-making has not replaced impressionistic modes of operation. The majors have continued to rely on creative tuition, talent relationships, and institutional knowledge, as well as internally commissioned, externally compiled audience

research. Rather, data analysis has added a new variable to how media organizations work, at once complementing and challenging the legacy media approach to making decisions.

As such, part II traces three interrelated stages in which data and decision-making have manifested in Hollywood's legacy media industry. First, the industry has negotiated the influx of new talent and technologies that enable and automate data analysis, both the practical-technical methodology as well as a new organizational structure to rationalize decisions (*Data as a Service*). Second, the majors have negotiated techniques and processes to translate data into insight by managing an expansive set of metrics (*Data as Intelligence*). And, third, they have negotiated ways to communicate data insights across organizational structures and establish data-driven insights as an element of corporate culture (*Data as Storytelling*).

- *Data as a Service*: The majors have worked to establish capabilities that enable and automate data analysis. Specifically, they have implemented new technologies and human resources into the legacy media system, licensing and acquiring data analysis software, and managing external relationships with data analysis vendors as well as building in-house data analysis departments and teams.
- *Data as Intelligence*: The digital economy has incorporated analytical tools that translate data into insights by way of standardized metrics. These metrics function as curated and filtered representations of data, designed to convey value within the digital economy.<sup>172</sup> For example, the industry has identified watch time (i.e., the time someone watches a film or a television show) as crucial data insight since it is seen as an indicator of viewer interest, content preference, and overall media activity. The majors have incorporated

---

<sup>172</sup> See, Gitelman, 2013. All data is, by default, curated and filtered by the platforms where the data is generated. Technological analysis, conducted by software applications, presents another form of processing.

digital metrics into their legacy media business, providing new ways to measure the digital interaction between consumers and content to inform decision-making.

- *Data as Storytelling*: Film studios and television networks have developed partnerships with analytics companies to report and communicate data insights across organizational structures. Analytics companies specialize in software applications that perform data analysis by way of visualization, using a visual language to extract, package, and communicate data insights, thereby facilitating data sharing and understanding across media organizations.

Together, these three stages form a developing culture of analytics where data is conceptualized, communicated, and capitalized as insight to inform corporate decision-making. The majors did not simply follow the business model of platforms, yet incorporated distinct practices of data-driven decision-making, responding to the cultural, economic, and technological developments posed by data analysis. As such, the legacy media industry is neither quant nor qual. It is a unique hybrid that has integrated impressionistic and analytical forms of decision-making.

## Chapter 2

### Analytical Actions: Establishing a Culture of Analytics

“Do you want to do *Moneyball* for Hollywood?”<sup>173</sup>

Thomas Tull, Founder of Legendary Entertainment

In 2003, startup entrepreneur-turned-private equity investor Thomas Tull founded Legendary Entertainment, a fledgling film production house, backed by \$500 million in investment from corporate investors, including AIG, Bank of America, and Falcon.<sup>174</sup> Tull, a walk-on player rather than a Hollywood insider, was among a coterie of financiers following the siren call of the media and entertainment business, which had increasingly campaigned for external funding to cover the growing production costs associated with a global market expansion. In most cases, corporate funding was fleeting and sporadic given the legacy media industry’s notorious unpredictability and high-risk environment marked by fickle returns. As the *New York Times* observed, “Hollywood has a history of sending outsiders packing with little to show for their investments but an empty wallet and a handful of premiere-party passes.”<sup>175</sup> Tull’s venture, however, was built on a differentiated logic. Taking an entrepreneurial approach, he envisioned Legendary as a portfolio company of various film investments where scale would outweigh individual risks, likening each production to a startup. He reportedly hired an outside accounting

---

<sup>173</sup> Knowledge Project Podcast, “Following Intellectual Curiosity with Thomas Tull,” *The Knowledge Project with Shane Parrish*, May 28, 2019, accessed June 6, 2020, <https://fs.blog/knowledge-project/thomas-tull/>

<sup>174</sup> See, Holson, Laura M., “Warner Venture with Investors,” *New York Times*, June 22, 2005, accessed June 6, 2020, <https://www.nytimes.com/2005/06/22/business/media/warner-venture-with-investors.html>

<sup>175</sup> Ibid



firm to crunch studio box office returns from the past ten years in an effort to identify what made film productions successful.<sup>176</sup> The close financial reading revealed that scale (i.e., numerous films with substantial budgets) played a key role in ensuring that movie investments could generate positive returns. Accordingly, Tull devised a corporate strategy to invest in a slate of movies, rather than select titles, with the goal of scaling box office success over a continuous period of time.

In 2005, Legendary secured a slate financing deal with Warner Bros., agreeing to co-produce and co-finance a slate of 25 movies over the next several years.<sup>177</sup> Under the agreement, Legendary would put up half of a film's production budget in exchange for shared revenue streams, thereby drawing downstream profits from theatrical returns, home entertainment, and merchandise sales. Legendary's funding enabled Warner Bros. to defray a substantial portion of production costs and mitigate risk in an increasingly complex media environment that saw a steady increase of blockbuster releases with \$100+ million price tags.<sup>178</sup> At the same time, the partnership with Warner Bros. gave Legendary instant credibility in the industry, access to intellectual property (IP) and creative talent, as well as an opportunity to generate long-term returns for its investor base.

While the first set of films generated moderate returns, Legendary eventually produced a roster of successful, high-return titles, including *300* (Warner Bros., 2006), *Batman Begins*

---

<sup>176</sup> See, Baker, Chris, "Q&A: Movie Exec Thomas Tull's Journey From Wall Street to Hollywood," *Wired*, October 19, 2009, accessed June 6, 2020, <https://www.wired.com/2009/10/ff-qa-tull/>

<sup>177</sup> Kelly, Kate, "Defying the Odds, Hedge Funds Bet Billions on Movies," *Wall Street Journal*, April 29, 2006, accessed June 6, 2020, <https://www.wsj.com/articles/SB114627404745739525>

<sup>178</sup> See, Elberse, Anita, *Blockbusters: Hit-making, Risk-taking, and the Big Business of Entertainment* (New York: Henry Holt, 2013).

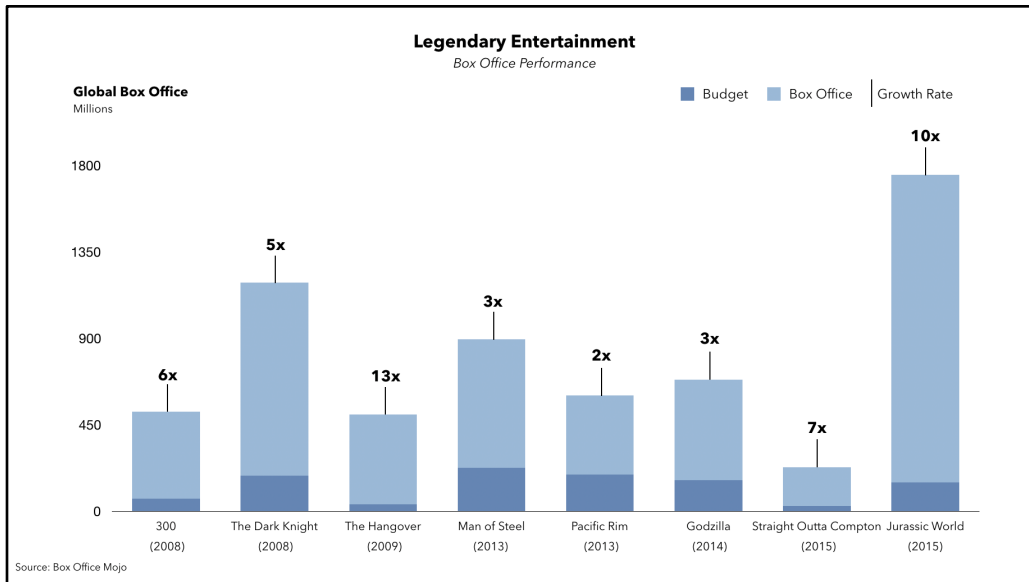
(Warner Bros., 2006), and *The Hangover* (Warner Bros., 2009), all of which spawned franchises with sequels and cross-platform media tie-ins, including video games, comic books, and toys.

After its slate deal with Warner Bros expired, the company negotiated a new partnership with Universal Pictures, delivering another string of hits, such as *Straight Outta Compton* (Universal Pictures, 2015) and *Jurassic World* (Universal Pictures, 2015) (figure 2.1). The box office success of these films notwithstanding, Legendary also had a flurry of misses, notably *Superman Returns* (Warner Bros., 2006), *Lady in the Water* (Warner Bros., 2006), *Watchmen* (Warner Bros., 2009), *Blackhat* (Universal Pictures, 2015), and *Warcraft* (Universal Pictures, 2016), all of which underperformed at the domestic box office.<sup>179</sup> Legendary's portfolio approach balanced the losses from these productions through the multiples and ancillary revenue streams from its hit movies as well as a growing media ecosystem that incorporated television production, comic book publishing, and digital content distribution, domestically and internationally.<sup>180</sup>

---

<sup>179</sup> Movies like *Warcraft* (Universal Pictures, 2016) actually turned a profit in China. See, Vincent, James, "China has turned Warcraft into the highest-grossing video game film ever," *The Verge*, June 21, 2016, accessed June 6, 2020, <https://www.theverge.com/2016/6/21/11988990/warcraft-highest-grossing-video-game-movie>

<sup>180</sup> Legendary expanded into the digital media business with acquisitions of Nerdist Industries (digital video and podcasting); Geek & Sundry (digital marketing agency); and Amy Poehler's Smart Girls (digital community).



**Figure 2.1:** Legendary’s select track record of films includes a range of box office returns.

On the surface, Legendary’s investment strategy appears fairly conventional, reflecting the legacy media industry’s standard of the time. The company prioritized deals involving recognizable IP with an in-built audience, drawn from a variety of pop culture artifacts.<sup>181</sup> Tull, a self-declared *fanboy*, pursued his creative penchant in backing projects based on comic books and video games, leading a new wave of blockbuster releases that were increasingly au courant in Hollywood.<sup>182</sup> Next to studio franchise output, such as Paramount’s *Mission Impossible* series (1996 - Present), 20th Century Fox’s comic book adaptations (2000-2020), and Marvel Studios’ emerging cinematic universe (2008 - Present), the majority of Legendary’s film slate effectively reflected mainstream industry formula. Still, Tull did not consider himself an industry insider and actively sought to position Legendary as a company with an outside-in view, bringing an

<sup>181</sup> See, Garrahan, Matthew, “Producer Follows His Own Script,” *Fortune*, December 5, 2010, accessed June 6, 2020, <https://www.ft.com/content/96d90b5c-00aa-11e0-aa29-00144feab49a>

<sup>182</sup> On comic book movies, see James N. Gilmore & Matthias Stork, *Superhero Synergies: Comic Book Characters Go Digital* (Lanham, MD: Scarecrow Press, 2014). On video game adaptations, see Brookey, Robert Alan, *Hollywood Gamers: Digital Convergence in the Film and Video Game Industries* (Bloomington: Indiana University Press, 2010).

external perspective to Hollywood's industrial practice. As he noted in an interview, "[w]hat I've been good at throughout my career is looking at an industry or business model and finding the points where I could improve it and make it more efficient."<sup>183</sup>

Accordingly, while Legendary's IP-driven film slate was effectively de rigueur, its corporate strategy ran counter to the normative dynamics of the legacy media industry, following an analytical approach that blended creative sensibilities with meticulous information-gathering, pattern recognition, and data crunching. In an interview with *Wired*, Tull outlined how his professional background and investment experience informed his approach to film financing and production. On his job in private equity, he noted (my emphasis):

**"My job was to identify opportunities and then work with those companies to make them grow. How are you going to get attention? Do your financials make sense? How will you market it? Will it work internationally? What sort of afterlife will it have? So we view each movie almost like that."**<sup>184</sup>

When looking at new film investments, Tull thus analyzed a variety of underlying information sets rather than exclusively focusing on the creative or the IP. Indeed, Legendary's business strategy signaled a rising focus on data analysis to rationalize media investments. Tull's portfolio approach to the business did not only reflect a shift in the economics of cultural production (i.e., how much capital is spent in Hollywood), but an emerging reconfiguration of Hollywood's corporate decision-making dynamics (i.e., how capital is spent and why).

Tull's analytical approach presented a contrast to the legacy media industry's established order of operation. Hollywood has long operated under impressionistic modes of decision-

---

<sup>183</sup> Knowledge Project Podcast, 2019

<sup>184</sup> Ibid

making, relying on creative instinct, talent relationships, and institutional knowledge to manage the business. As such, the majors have persistently drawn on tried-and-true mechanics, informed by the lived experience of executives and senior leadership, investing in familiar faces and known formulas in order to ensure film and television programming appeals to the largest share of an increasingly global audience. This investment in scale, manufacturing tentpoles and blockbusters targeted at mainstream viewers, had emerged as the modus operandi of the business by the early 2000s.<sup>185</sup>

Tull embraced the industry's focus on blending IP with high production costs, yet considered the mass audience approach to market films on a broad scale archaic and inefficient in a digital media environment increasingly defined by a variety of competing choices provided by the Internet, from websites to social media, mobile apps, and streaming video. In particular, he took issue with the indiscriminate spend of marketing budgets, which were spread across all available media channels (e.g., television, newspapers, out-of-home, digital), irrespective of the target audience's behavior or interests. He saw an opportunity to create more efficiencies in how Hollywood made strategic business decisions, specifically around capital expenditures. In particular, he posited that Legendary could effectively minimize overall production costs by retooling its marketing approach, selectively targeting specific audience segments rather than trying to reach a mass audience.<sup>186</sup> His vision involved identifying the audience members with

---

<sup>185</sup> See, Elberse, 2013

<sup>186</sup> Former industry executive and USC lecturer David Weitzner commented on the perceived inefficiencies of Hollywood's established marketing approach: "There is a huge amount of money just being thrown away." See, Rainey, James, "The Perils of Promotion: Pricey TV Campaigns, Fear of Change Shackles Movie Spending," *Variety*, March 8, 2016, accessed June 6, 2020, <https://variety.com/2016/film/features/movie-marketing-advertising-tv-campaigns-1201724468/>

the highest likelihood of seeing a Legendary movie, based on clusters of digital information. As he elaborated in an interview on his business approach in Hollywood:

“If it’s clear to me [sic], psychographically, that you’re not interested in a movie, I don’t want to bother you. And, conversely, if you’re someone who is persuadable, and we could make interested, that would make a material impact on our business. We’d be able to cut out a large chunk of advertising dollars and be much more efficient.”<sup>187</sup>

Tull’s rational way of thinking can be seen as a logical extension of Wall Street’s growing presence in the legacy media industry.<sup>188</sup> Legendary was built on an approach that utilizes economic analysis and provable scenarios to mitigate the risk inherent in the unpredictable and fickle media and entertainment business, and extended this analytical framework beyond production planning into marketing execution. Tull colloquially referred to this holistic strategy as *Moneyball for Hollywood*,<sup>189</sup> a direct reference to Michael Lewis’ book, *Moneyball: The Art of Winning an Unfair Game* (2003), which examines the use of statistics and empirical decision-making to change the corporate-organizational culture of professional baseball, countering the sports’ inherent reliance on institutional knowledge with a pivot toward data analysis. The *Moneyball* reference illustrates Tull’s ambition to structure Legendary as an analytical and evidence-based media organization that utilizes data to manage the balance sheet.

To institutionalize the *Moneyball* ethos within Legendary, Tull re-envisioned the company’s organizational structure and operating approach. He built up Legendary’s analytical

---

<sup>187</sup> Knowledge Project Podcast, 2019

<sup>188</sup> See, Epstein, Edward Jay, *The Hollywood Economist: The Hidden Financial Reality Behind The Movies* (New York: Melville House, 2012).

<sup>189</sup> See, Kirsner, Scott, “Making movies the ‘Moneyball’ way,” *Boston Globe*, March 31, 2016, accessed June 6, 2020, <https://www.bostonglobe.com/business/technology/2016/03/31/making-movies-moneyball-way/Uzgwh2cdGthA1N3nZHqz0N/story.html>

capabilities with the acquisition of StratBridge, LLC., a small boutique data analytics firm that specialized in the sports industry, using data analysis to efficiently determine and manage the performance and associated costs of players.<sup>190</sup> StratBridge brought technical expertise and technological resources in data analysis that proved key for Legendary. Under the guidance of Eric Schmidt, Chairman of Google and a Legendary board member at the time, Tull envisioned StratBridge as the centerpiece of Legendary’s internal decision-making engine, dubbed *Legendary Analytics*, installing the company’s founder Matt Marolda as Chief Analytics Officer. Effectively, Tull acquired and integrated a data analysis department within Legendary, creating a foundation for the company to operate analytically, grounded in data.

Legendary assembled a team of diverse practitioners to put Tull’s vision of an analytical organization into practice,<sup>191</sup> establishing and implementing the following core functional groups within Legendary Analytics:<sup>192</sup>

- *Analysts*: Legendary’s analysts made up a multi-disciplinary team of data scientists, led by a Ph.D. in astrophysics, with experts from sports science and dynamic pricing, tasked

---

<sup>190</sup> See, Gaudiosi, John, “Legendary Pictures CEO talks tech that gave ‘Godzilla’ its box-office roar,” *Fortune*, May 22, 2014, accessed June 6, 2020, <https://fortune.com/2014/05/22/legendary-pictures-ceo-talks-tech-that-gave-godzilla-its-box-office-roar/>

<sup>191</sup> The analytics team consisted of 60 professionals, or, per Legendary’s Chief Analytics Officer at the time, “59 more than anyone else in Hollywood.” See, Hayes, Dade, “Legendary Entertainment Selling Majority Stake In Its Applied Analytics Unit,” *Deadline*, November 13, 2017, accessed June 6, 2020, <https://deadline.com/2017/11/legendary-entertainment-selling-majority-stake-in-its-applied-analytics-unit-1202206956/>

<sup>192</sup> See, Krigsman, Michael, “Moneyball for movies: Data science and AI in Hollywood,” *ZDNet*, February 20, 2018, accessed June 6, 2020, <https://www.zdnet.com/article/moneyball-for-movies-data-science-and-ai-in-hollywood> and John A. Deighton & Leora Kornfeld, “Legendary Entertainment: Moneyball for Motion Pictures (Case Study),” *Harvard Business School*, May 2016.

to analyze large data sets across the digital landscape in an effort to identify patterns and relations that enable the company to generate new and useful information on the business.

- *Developers*: Legendary hired a team of developers and engineers to build proprietary data analysis software tools, designed to test and automate experiments identified by analysts.<sup>193</sup> The developer team effectively delivered scalable analysis, enabling the company to validate and expand the work of analysts.
- *Communicators*: Legendary added a suite of strategy and marketing professionals with experience in advocating and implementing decisions based on the analytical work of analysts and developers, with a specific focus on executing data-driven campaigns. The team further worked to share insights with teams across the company as well as studio partners Warner Bros. and Universal Pictures.

The group of analysts, developers, and communicators worked to analyze, interpret, and communicate data to inform corporate decision-making across the organization. Tull envisioned a top-down organizational structure that would consistently translate the growing amounts of data across the digital media landscape to inform the company's business path forward. The output of Marolda's team would inform decision-making across all of Legendary. Analysts would search for patterns in data; developers would test analytical theories and hunches with software technology; and communicators would craft strategies to communicate results across the company, all with the goal of minimizing costs and risk while maximizing revenue and

---

<sup>193</sup> See, Moore, Malcolm, "Film studios select movie-fan data for starring role to drive ticket sales," *Financial Times*, September 6, 2016, accessed June 6, 2020, <https://www.ft.com/content/8489fd08-15bc-11e6-b197-a4af20d5575e>



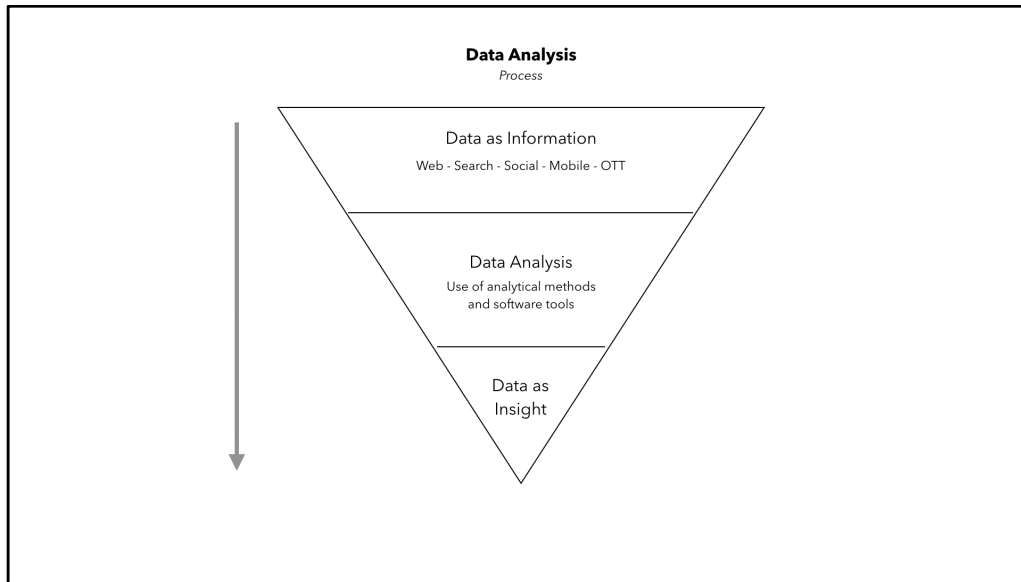
success. As such, the analytics division was tasked to help Legendary make decisions that optimize business decisions in film production and marketing, including where to invest, how much, and how to save costs.<sup>194</sup>

Legendary Analytics effectively approached data not as a mass of information, but a kernel of insights. This conceptual shift from *information* to *insight* constitutes the underlying crux of data analysis. Information represents a constantly scaling volume of data points within the evolving digital media landscape. Insight constitutes an interpretation of data, a translation based on a company's business goals, an end product of analysis that showcases the value of information in a specific context. In this regard, data analysis can be conceptualized as the process of turning information into insight, distilling massive amounts of data from across the digital landscape into a specific set of data that enables a company to inform business decisions. As such, it operates as a generative and interpretative process designed to extract meaning and value from data.

Conceptually, data analysis as a process is based on a functional methodology. Companies access data from an established information infrastructure, mine the data for specific sets of information that align with their business goals, and extract the matching patterns. It effectively represents an analytical exercise that identifies information relevant to a company's business strategy, thereby generating valuable insight. In other words, as information quantity is distilled, the quality of insight increases (figure 2.2).

---

<sup>194</sup> See, Knowledge Project Podcast, 2019



**Figure 2.2:** As data is processed through analysis, information transforms into insight.

Legendary’s analytics division effectively cultivated an operational process that deployed analysis to translate data into business insights, using a key set of steps. The team leveraged access to a database of consumer and usage data, including first-party data via thousands of email addresses as well as second- and third-party data compiled from social media profiles, public blogs and forums, as well as movie box office performance statistics and television ratings.<sup>195</sup> Then, they defined the parameters of the analysis, which meant identifying those audience segments with the highest likelihood of seeing a new Legendary movie. In particular, Marolda noted that Legendary’s focus centered on audiences “we consider to be given, meaning they’re going to watch the movie no matter what.” Among this group of dedicated fans, Legendary segmented out viewers indicating a potential interest in watching the movie in the theater. Tull and Marolda dubbed these segments “persuadables.”<sup>196</sup> Using this type of

<sup>195</sup> See, Marolda, Matthew, “Changing Hollywood Paradigms with Analytics,” *Innovation Enterprise On Demand*, 2014, accessed June 6, 2020, <https://ieondemand.com/presentations/changing-hollywood-paradigms-with-analytics>

<sup>196</sup> See, Krigsman, 2018

segmentation, Legendary effectively worked to define a movie's audience from the bottom-up, starting with the most likely viewers, rather than taking a top-down mass audience view as commonly practiced in Hollywood. Furthermore, the team then proceeded to the technical data analysis, utilizing manual classifications and automated software tools, with the specific goal of discovering insights about persuadable audiences. They drilled down massive amounts of information, identifying insights that matched Legendary's key audience segments, defining the consumer and behavioral patterns of viewers that might be persuaded to see a Legendary movie if presented with the right marketing message. Communicators shared the insights across the organization to inform the implementation of Legendary's marketing approach. In this regard, the analytics division leveraged data access to generate insights that drove business actions.

Legendary presents a new type of media organization in which data analysis shapes the way the company operates. The company's internal analytics division was set up to mine a database of consumer background and behavioral information, discover patterns that indicate an interest in Legendary's movie slate, and effectively determine an audience from the bottom-up. Tull and his leadership team prioritized the resulting insights in the planning and execution of production and marketing strategies. As such, data analysis acted as a core lever of Legendary's approach to corporate decision-making, which differs from key tenets of the legacy media industry. One, while Legendary embraced genre formulas and IP, the company further deployed data analysis to test and iterate how it would invest in the media and entertainment business. Two, contrary to the traditional approach of legacy media companies, which generally contract with market research vendors to execute focus groups and tracking surveys shortly before a new movie release, Legendary used data analysis holistically throughout all business stages, from the script stage through production and marketing. Three, Legendary did not choose institutional

knowledge over data insights or vice versa, but synthesized both approaches. As such, Legendary represents an emerging culture of analytics, a corporate-industrial mindset that prioritizes data analysis as a key variable of corporate decision-making in the legacy media industry.

The conceptual dimension of a *culture of analytics* is drawn from a set of methodologies. On a high level, it adapts the idea from computer science which positions analytics as a process to “handle the challenges of large data sets”<sup>197</sup> and thereby extract meaning from data. Furthermore, it draws on management and business studies which examines the way enterprises and organizations make sense of data through analytical methods. In particular, the process to generate meaning is characterized as a discovery of insights across the digital media landscape, sifting through massive amounts of data to reveal specific data points that prove relevant to a company’s strategic plans.<sup>198</sup> This process of rationalization incorporates industrial and technological components. Companies with an analytics focus follow a strategic corporate mandate to prioritize data and insights, which manifests in the setup of their organizational structure (i.e., data analytics departments, executives, and professionals) as well as their technological infrastructure (i.e., the use of data analysis software tools to manage the technical features of data analytics, drawing on experienced professionals, advanced technologies, and proven methodologies to analyze information). In this sense, they practice a corporate culture that systematically deploys computational expertise and technology to translate data into insights. This focus on generating meaning and value from data through analysis is particularly evident in the technology industry. Platform-based companies, in effect, have built global brands

---

<sup>197</sup> Duan, Lian & Xiong, Ye, “Big data analytics and business analytics,” *Journal of Management Analytics*, Volume 2, Issue 1, 2015.

<sup>198</sup> See, Schlesinger, 2012

on their engagement with data and information. For example, Google’s mission is to “organize the world’s information and make it universally accessible and useful,” a corporate positioning statement that synthesizes culture and data analysis into an organic entity.<sup>199</sup> Furthermore, from Google to Facebook and Netflix, platforms have emphasized the corporate focus on analytics as a competitive differentiator, a strategic prioritization of insights that enables business innovation and strategic advantage. Indeed, various studies on corporate performance argue for a connection between business success and data-driven decision-making.<sup>200</sup> As such, the digital marketplace and direct-to-consumer economy is effectively grounded in a culture of analytics. It represents a corporate-industrial environment driven by data analysis and insights.

Legendary’s culture of analytics can thus be seen as a reconfiguration of the technology industry’s industrial logic and its extension into the realm of media and entertainment. Indeed, Legendary’s analytics division illustrates a core set of industrial strategies that have increasingly worked to integrate data analysis into the legacy media industry, gradually adapting Hollywood’s industrial DNA, from a creative-driven business run by an entrenched rank of institutional powers to a corporate environment considering data analysis as a decision-making factor.

- *Capabilities*: Tull built a range of internal capabilities to integrate data analysis within Legendary’s organizational structure, acquiring an external data analysis company and bringing the talent and software tools in-house. Additionally, he made data analysis a top-down priority for the company, installing a Chief Analytics Officer and imbuing the

---

<sup>199</sup> See, Google, “Our Approach to Search,” accessed June 6, 2020, <https://www.google.com/search/howsearchworks/mission/>

<sup>200</sup> See, LaValle, Steve, Lesser, Eric, Shockley, Rebecca, Hopkins, Michael S & Kruschwitz, Nina, “Big Data, Analytics and the Path From Insights to Value,” *MIT Sloan Management Review*, December 21, 2010, accessed June 6, 2020, <https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/>

analytics division with leadership support to play a central role within Legendary's corporate decision-making process.

- *Framework:* Legendary's analytics division developed a framework to translate massive amounts of information into relevant business insights. By focusing on dedicated fans and "persuadables," the team set parameters that enabled focus in data analysis, a way to look for and discover data that indicates interest among a specific set of viewers.
- *Delivery:* Legendary emphasized the communication of data insights, operating an organizational structure in which data was shared and distributed across the company. This holistic approach ensured that the key levers of the organization incorporated data insights into decision-making dynamics. Furthermore, Tull and Marolda emerged as thought leaders in Hollywood, advocating their analytical approach on the conference circuit. Overall, Legendary paved a way for an industrial environment where data analysis became a more tangible and operational element of the business.

Legendary's culture of analytics promoted the use of data analysis to shape corporate decision-making in the legacy media industry, synthesizing an analytical mindset with technical-computational practices, embedded within a highly creative environment. This industrial approach ostensibly appears as an anomaly within Hollywood. The legacy media industry has long operated as a creative-first business, driven by the intuition and institutional knowledge of inside gatekeepers, impervious, and arguably resistant, to more analytical practices. As Smith and Telang found in their expansive research studies on media organizations, "most

entertainment firms have little institutional strength or political capital with data analytics.”<sup>201</sup>

Yet, as access to information became increasingly widespread within the shifting media environment, driven by the proliferation of digital platforms, shifting consumer behavior, and growing competition, data analysis emerged as a central skill in the legacy media industry to manage the growing amounts of data shaping the dynamics of the legacy media business.<sup>202</sup>

Effectively, the majors have adapted an emerging analytical approach, building up capabilities for data analysis, implementing frameworks to conceptualize data as insight, and developing strategies to communicate data as a relevant business driver. In this sense, rather than a deviation from the norm, the Legendary model has come to signal Hollywood’s shift toward a data-driven operating logic.

Accordingly, this chapter examines a set of interrelated industrial developments that have given rise to a formative culture of analytics in Hollywood over the course of the 21st century.

- *Data as a Service*: Hollywood studios and networks have long maintained an external network of market research vendors to generate insights on new film and television releases by way of focus groups, surveys, and interviews. The majors outsource research work to external vendors and operate internal research departments to consolidate and communicate insights within their organizational structures. Over the course of the new millennium, the majors have taken steps to expand their research capabilities to respond to the growing availability of digital information by building up data analysis capabilities (i.e., talent and technologies). In particular, they have set up data analytics departments,

---

<sup>201</sup> Smith & Telang, 2016

<sup>202</sup> See, Dickey, Josh L., “Bizzers: Execs must master data analytics,” *Variety*, November 29, 2012, accessed June 6, 2020, <https://variety.com/2012/film/news/bizzers-execs-must-master-data-analytics-1118062878/>

filled with data professionals versed in data analysis software tools, thereby laying a foundation for a more holistic approach to analytical decision-making within media organizations. Effectively, the majors have borrowed corporate-structural elements from technology companies and platforms to adapt their organizational structures, establishing a suite of analytics service teams that analyze, interpret, and communicate data.

- *Data as Intelligence:* The majors have adapted new frameworks to make sense of data and approach it as a viable form of intelligence. Specifically, they have implemented a rising set of digital metrics, an expansive set of key performance indicators (KPIs) that frame information as insight. While studios and networks have long prioritized a small set of KPIs to measure success, namely box office and ratings, the shifting digital media environment has yielded an explosion of new metrics that provide additional information on the media audience and content. These metrics structure, classify, and segment information as insight by assigning data specific relevance, meaning, and value within the media and entertainment business, enabling the majors to track and analyze consumer background and usage data. As a result, the majors have worked to interpret, inform, and incorporate a wave of metrics as part of their decision-making dynamics.
- *Data as Storytelling:* The majors have widely integrated software tools designed to analyze, interpret, and communicate data insights. Data analysis software uses digital interfaces that visually consolidate, coordinate, and conform data to help organizations make sense of the growing amounts of digital information. By implementing data analysis software, the majors have created an organizational structure in which data is shared and distributed via software interfaces, creating a level of transparency that has



made information more available and accessible. In particular, the use of data analysis and visualization software reflects a growing trend to tell stories with data by presenting information visually, thereby enabling the communication of actionable insights to inform the decision-making process.

Collectively, these three interrelated conceptual stages illustrate an evolving industrial infrastructure that has increasingly positioned data analytics as a variable in Hollywood's decision-making process.

### **Data as a Service: The Formation of the Analytical Enterprise**

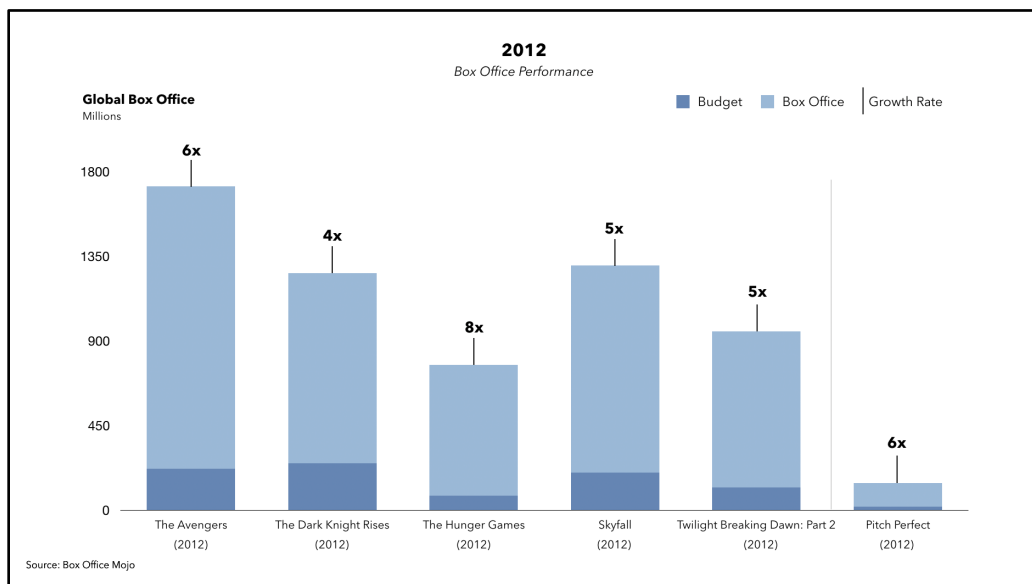
2012 was a year dominated by franchise filmmaking. The list of the year's highest grossing films represented a cross-section of comic book movies, YA novel adaptations, and popular IP, led by the Walt Disney Studios' *The Avengers*, Warner Bros. Pictures' *The Dark Knight Rises*, Lionsgate's *The Hunger Games*, Sony Pictures' *Skyfall*, and Summit Entertainment's *The Twilight Saga: Breaking Dawn Part 2* (figure 2.3). The top 10 films of the year grossed over \$500 million in global box office respectively, with four titles crossing the billion-dollar mark. Notably, all major studios, including mini-majors like Lionsgate and Summit Entertainment, capitalized on franchise hits during this time; all, except Universal Pictures.

While Universal owned popular franchises like *Fast & Furious* and *Jurassic Park*, the studio did not release any sequels in 2012, instead focusing on the launch of *new IP*, drawing from board games and folklore myth with titles such as *Battleship* and *Snow White & The Huntsman*, neither of which proved successful.<sup>203</sup> Nevertheless, 2012 proved a pivotal year for

---

<sup>203</sup> See, Masters, Kim, "'Battleship' Fallout," *Hollywood Reporter*, May 23, 2012, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/battleship-universal-box-office-taylor-kitsch-327972>

Universal, with a string of successful mid-budget original releases that broadened the studio’s slate and carved out a niche opportunity in the marketplace, including the actioner *Contraband*, the R-rated comedy *Ted*, and *Pitch Perfect*, a college-set musical comedy, filled with catchy cover songs and starring up-and-coming actors with a growing digital fan base. *Pitch Perfect* grossed \$115 million against a \$17 million budget (a 6x increase) and spawned two sequels, which nearly grossed half a billion dollars worldwide. It emerged as one of the industry’s surprise hits of the year, a crossover venture, an unexpected breakout in Hollywood’s dense IP inventory.<sup>204</sup> For Universal, however, the film’s success was not an unexpected surprise, but an analytical anticipation.



**Figure 2.3:** While *Pitch Perfect*’s box office appears small compared against major box office hits, its return rate puts it within the upper echelon of the year’s most financially viable titles.

Set in the competitive world of collegiate acapella performances, *Pitch Perfect* was conceived as a vehicle for a young female audience, effectively targeting the college crowd

<sup>204</sup> See, Cunningham, Todd, “Universal Bets Perky ‘Pitch Perfect’ Will Keep Box Office Beat,” *The Wrap*, October 4, 2012, accessed June 6, 2020, <https://www.thewrap.com/universal-bets-perky-pitch-perfect-keeps-box-office-beat-58856/>

under 25 in one of the industry's core quadrants.<sup>205</sup> In the lead-up to the marketing campaign, Universal followed the conventional test-and-learn playbook, contracting market research firms to conduct focus groups, polls, and test screening exit surveys to track the film's initial reactions and fine-tune promotional messaging in the marketing campaign accordingly. Additionally, the studio expanded beyond the traditional research routine to prioritize new methods of digital information-gathering, seeking to capture audiences that consume, discuss, and engage with media online. At the time, the Internet had a 79% penetration rate in the U.S., with frequent usage over-indexing on young adults aged 18-29, more likely to have a college education and higher income.<sup>206</sup> In particular, marketing executives showed a keen interest in reaching this audience demographic on social media, which had increasingly become the portal to online activity and, by extension, a central node of film and television discourse, overwhelmingly driven by young adults.<sup>207</sup> *Pitch Perfect*'s blend of music covers, performative display, and visual choreography appeared tailor-made for the interactive tune-in culture of social media.<sup>208</sup> The majors focused their attention on Twitter, an epicenter for user-generated content (UGC) about films and television shows.<sup>209</sup> Indeed, the micro-blogging platform had experienced

---

<sup>205</sup> For an overview of the audience quadrant model, see, Napoli, 2010.

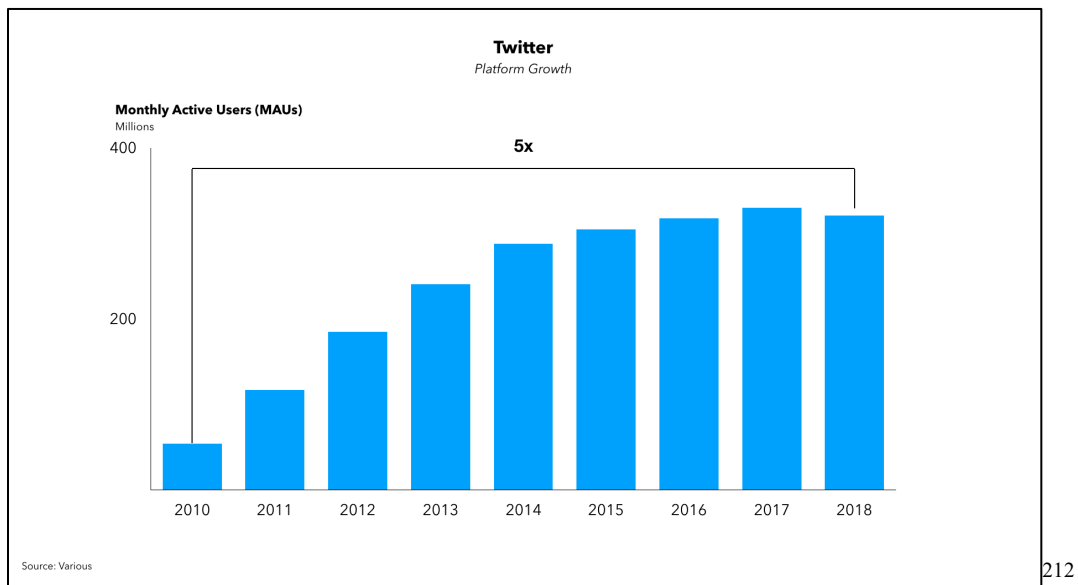
<sup>206</sup> See, Pew Research, "Internet/Broadband Fact Sheet," June 12, 2019, accessed June 6, 2020, <https://www.pewresearch.org/internet/fact-sheet/internet-broadband/>

<sup>207</sup> See, Pew Research, "Social Media Usage: 2005-2015," October 8, 2015, accessed June 6, 2020, <https://www.pewresearch.org/internet/2015/10/08/social-networking-usage-2005-2015/>

<sup>208</sup> See, Jenkins, Henry, Ford, Sam & Green, Joshua, eds., *Spreadable Media: Creating Value and Meaning in a Networked Culture* (New York: New York University Press, 2013).

<sup>209</sup> See, Godley, Chris, "THR's Social Media Poll: How Facebook and Twitter Impact the Entertainment Industry," *Hollywood Reporter*, March 21, 2012, accessed June 6, 2020, <https://www.hollywoodreporter.com/gallery/facebook-twitter-social-media-study-302273/1-social-media-as-entertainment>

exponential user growth over the previous two years, primarily driven by users aged 18-34, *Pitch Perfect*'s core demographic<sup>210</sup> (figure 2.4). In an effort to understand the film's early resonance with the target audience, Universal thus complemented its traditional research approach with a practice called social media analysis, setting out to track the online conversation on *Pitch Perfect* by surveying the audience on Twitter.<sup>211</sup>



**Figure 2.4:** Twitter grew its audience base exponentially from 2010-15, followed by incremental growth.

To this end, Universal partnered with Fizziology, a small analytics firm that specialized in social media analysis. Specifically, the company provided a service called *social listening*, an analytical practice that has effectively become synonymous with social media analysis. Social

<sup>210</sup> See Twitter Investor Relations for overview of user demographics, accessed June 6, 2020, <https://investor.twitterinc.com/home/default.aspx>

<sup>211</sup> For a critical perspective on the use of data analytics tools in the digital media industry, see, Andrejevic, Mark, *Infoglut: How Too Much Information Is Changing the Way We Think and Know* (New York, Routledge, 2013).

<sup>212</sup> Twitter Investor Relations, accessed June 6, 2020, <https://investor.twitterinc.com/home/default.aspx> and Ingram, Matthew, "Facebook vs. Twitter: An Infographic," *Gigaom*, December 20, 2010, accessed June 6, 2020, <https://gigaom.com/2010/12/20/facebook-vs-twitter-an-infographic/>

listening refers to the process of determining the popularity of a brand or product by analyzing data across social media platforms as well as blogs, forums, and wikis, using data analysis software tools to parse online conversation.<sup>213</sup> In particular, this type of data analysis deploys software automation to gather unstructured data in form of text, images, and videos across social media platforms, identifies common patterns among them, and extracts relevant meaning. In practice, social listening deconstructs what has been termed the *social media conversation*, the interactive fabric of messages, posts, comments, replies, and reactions that link users and create a two-way dialogue. Social listening utilizes software tools to highlight keywords, themes, and mentions that provide an estimate of a subject's online popularity, capturing both the volume (i.e., the number of conversations relative to the number of a platform's users) as well as the quality of conversation (i.e., the key topics of conversation and the indicated sentiment associated with them), rendering a historical snapshot of a topic's resonance across social media.

While analytics companies deploy social listening across the entire digital landscape, from websites to search and video, they specifically focus on social media platforms due to the exponential volume of online conversations and associated data. Analytics companies collect, capture, and analyze this mass of social media data in two ways. One, they use *crawling* software technology to automatically index information from thousands of public websites (e.g., blogs, forums, wikis) and download web pages to consolidate, classify, and conform unstructured text and visuals into data. Two, they leverage application programming interfaces (APIs) to tap into data repositories of major social media platforms (e.g., Facebook, Twitter, YouTube), effectively gaining access to a statistical information sample from millions of users. Specifically, analytics

---

<sup>213</sup> See, Hollander, J.B., Graves, E., Renski, H., Foster-Karim, C., Wiley, A., Das, D., "A (Short) History of Social Media Sentiment Analysis." In: *Urban Social Listening* (London: Palgrave Macmillan, 2016).

companies focused on Twitter as the cornerstone of social media analysis, leveraging the platform as the “world’s largest focus group,”<sup>214</sup> given its *scale* (i.e., Twitter had over 180 million monthly active users worldwide by the end of 2012), its *association with media and entertainment* (i.e., previous studies showed a high affinity to discuss film and television programming among Twitter users compared to other social media platforms<sup>215</sup>), and its *high-functioning API* (i.e., Twitter’s API was labelled *Firehose* for providing comprehensive data access, referring to a higher statistical sample of real-time information).

For the *Pitch Perfect* analysis, Fizziology thus trained the focus of its software on Twitter, searching for textual references to the film by combing through thousands of comments, mentions, and posts. Social listening requires adjusting semantic search patterns to exclude irrelevant data that is either too general (i.e., search results that highlight information about *general pitches deemed perfect*) or too specific (i.e., search results that emphasize the adjective *pitch-perfect*), ensuring that the analysis incorporates and converts the most relevant information. Warner Bros. reportedly experienced this issue in examining the online conversation around the film *Creed* (2015), a sequel to the original *Rocky* franchise. The social listening analysis of the media title “Creed” pulled in references to the rock band *Creed* as well as the video game *Assassin’s Creed* (Ubisoft, 2007-2018), which proved challenging for the semantic analysis. Warner Bros. solved the problem with a classification algorithm, focusing the analysis on content classified for the film *Creed* and its associated information.<sup>216</sup> Universal effectively took

---

<sup>214</sup> Littleton, Cynthia, “10 Things We Learned at Variety’s Big Data Summit,” *Variety*, November 4, 2015, accessed June 6, 2020, <https://variety.com/2015/digital/news/10-things-we-learned-at-variety-s-big-data-summit-1201634065/>

<sup>215</sup> See, Nielsen, “Nielsen and Twitter Establish Social TV Rating,” December 18, 2012, accessed June 6, 2020, <https://www.nielsen.com/us/en/press-releases/2012/nielsen-and-twitter-establish-social-tv-rating/>

<sup>216</sup> See, Ciamprone, Danny, “Warner Bros.’s Brian Kursar is the Data Knight,” *Sync*, March 15, 2017, accessed June 6, 2020, <https://sync-magazine.com/2017/wb/>

the same approach for *Pitch Perfect*. Based on the studio's input on the film's story, cast, and themes, and information from early test screenings, Fizziology's software filtered contextually relevant information to highlight audience feedback.

Fizziology's analysis of the social conversation around *Pitch Perfect* reportedly revealed a key set of insights for Universal.<sup>217</sup> First, a large part of the online conversation was driven by young men, in addition to young women, suggesting a new audience opportunity for the studio. Second, male viewers seemingly came to see the film by accompanying women to early screenings, indicating that the final film resonated with the male audience segment based on their post-viewing activity on social media. And, third, both men and women discussed key parts of the film rather than commenting on the general story, highlighting specific musical sequences and breakout characters, which provided Universal with direction on the film's standout elements that resonated with audiences. Effectively, the analysis gave Universal an overview of the core audience and their perspective on the film, prompting the studio to release the film a week early to capitalize on the college-centric buzz during the back-to-school period, focus the marketing outreach on audiences' favorite scenes and characters, and target incremental male viewers on top of the core female audience. Universal may not have known that *Pitch Perfect* would be a hit, but it had a sense that it would resonate with a custom audience segment.

Conceptually, the insights generated from Fizziology's social listening analysis are similar to the reports traditional market research firms have customarily delivered to studios for decades. Both capture an audience's demographic background and viewing reaction, enabling studios to estimate a film's popularity and adjust their marketing approach accordingly. Yet,

---

<sup>217</sup> See, Ungerleider, Neil, "'Pitch Perfect' And How Analytics Are Transforming Movie Marketing," January 10, 2014, accessed June 6, 2020, <https://www.fastcompany.com/3024655/pitch-perfect-and-how-analytics-are-transforming-movie-marketing>

beyond that, social media analysis represents a highly differentiated practice, for several reasons. First, it ups the scale of the data involved in the analysis (i.e., millions of users on social media vs. a small representative sample of movie-goers polled in traditional research). Second, it follows an open and unfiltered nature of gathering information (i.e., the focus on everyday comments on social media vs. directed questions in focus groups and surveys). And, third, it employs data analysis software to gather information, identify patterns, and extract meaning (i.e., the use of analytical automation and interpretation vs. the manual tabulation of data).

Additionally, analytics companies like Fizziology differed in how they deliver insights to studios. Traditional market research firms generally provided custom reports to the majors, highlighting the key results from focus groups, polls, and surveys. Analytics companies equally shared custom reports in order to showcase the results of their analysis, yet further incorporated a digital solution. Beyond the physical insights report, they provided virtual access to data insights, enabling studios to log into an online portal that hosts the analysis.<sup>218</sup> Effectively, analytics companies designed custom interactive interfaces that showcase, or *report*, data insights. Dubbed *reporting dashboards*, these digital interfaces are dynamic and up-to-date, compared to the static and set nature of traditional reports, consistently gathering and analyzing data across the digital landscape to update the analysis and generate insights at regular intervals. While interfaces are equally coded to an extent that presents information in a certain way, they also allow for a user's interaction with the data on display. As such, they make insights available in an online environment, which allows the major studios and networks to view up-to-date insights at any time as well as adjust the reporting format to gain additional information.

---

<sup>218</sup> Traditional research vendors equally adopted digital delivery formats. See, Wyatt, Justin, "Market Research in the Media Industries: On the Strategic Relationship between Client and Supplier." In: *Making Media Work: Cultures of Management in the Entertainment Industries*, edited by Derek Johnson, Derek Kompare & Avi Santo (New York: New York University Press, 2014).



Analytics companies accordingly operate a software-as-a-service (SaaS) business model. Using the SaaS model, they license access to insights through software-based digital interfaces. Their software analyzes data, translates digital information into insights, and hosts the results to make them virtually accessible. In exchange, the majors pay a licensing fee. By partnering with analytics companies, the majors receive access to data insights via custom reports and usable interfaces that automatically mine and analyze audience conversation across the digital sphere.

Data analysis software, as administered by analytics companies, does not only function as a static visual display case for insights. It effectively operates as a self-serve tool that enables the majors to access different types of insights. Dashboards are designed to show different representations of data analysis, also known as different *cuts of data*. Essentially, dashboards highlight different views of data, displaying various pieces of information, generally classified as metrics, dynamically. For example, the online conversation of a film can be cut by social media comments (i.e., the volume of conversation), the age of social media users (i.e., the demographics of the audience), or a specific timeline, to list but a few options.

Additionally, this interactive functionality, in some cases, offers the capability to adjust the visual parameters of the analysis, reset the interface, and build custom dashboards from scratch. This plug-and-play approach effectively enables the majors to run their own analytical studies. Rather than the analytics companies' default dashboard, this custom approach allows the majors to create analytical views that are highly customized to their business. While more flexible and dynamic than the wired dashboard provided by analytics companies, the approach is also more high-touch, requiring the majors to manage the dashboard design through internal teams. To this end, Hollywood has increasingly brought in data professionals to manage analytical processes in-house, building, maintaining, and analyzing data insights dashboards. For

example, for the release of *Star Wars: The Force Awakens* (Disney, 2015), the Walt Disney Company deployed an internal data analytics team to analyze the massive online conversation in the lead-up to the film release, working with a variety of analytics vendors to capture insights on the film’s buzz across social media.<sup>219</sup> Hollywood’s expanding partnerships with analytics companies led to a growing adoption of data analysis software among the majors, which emerged as a central element of the industry’s research methodology.

Companies like Fizziology are part of a cadre of analytics companies that offer data analysis software to the media and entertainment industry. Analytics companies license software-as-a-service tools that at once conduct end-to-end data analysis, while further enabling analytics professionals to perform data analysis themselves. Effectively, they operate as technology companies, using proprietary software to deliver data analysis services via digital dashboards. Dashboards deliver data insights via default reporting settings, enable custom data analysis through an interactive visual display, and consistently refresh the analysis through access to APIs from digital platforms. As such, analytics companies have come to represent an extension of Hollywood’s research repertoire, an emerging cottage industry that has largely been institutionalized.<sup>220</sup> Hollywood has long cultivated relationships with market research companies to gather and analyze information, dating back to the use of opinion polling during the industry’s classical period in the 1930s.<sup>221</sup> Thus, analytics companies do not constitute a new development, but rather a reconfiguration and expansion of an existing operational research infrastructure.

---

<sup>219</sup> As Disney’s former VP of analytic insights noted, “It’s such a massive amount of data — it’s energizing to my team.” See, Littleton, 2015, “10 Things We Learned at Variety’s Big Data Summit”

<sup>220</sup> See, Laporte, Nicole, “Pitch Perfect: How Universal’s Digital Marketing Helped It Have The Best Year Ever,” *Fast Company*, September 11, 2015, accessed June 6, 2020, <https://www.fastcompany.com/3050984/pitch-perfect-how-universals-digital-marketing-helped-it-have-the-best-year-ever>

<sup>221</sup> See, Ohmer, Susan, *George Gallup in Hollywood* (New York: Columbia University Press, 2006).

Over the course of the 21st century, the evolving digital landscape has created a veritable sea of information, with increasing volumes of data being generated on a daily basis. This continuous influx in data generation was driven by a confluence of cultural, economic, and technological factors, including the commercialization of the Internet, and platforms specifically, the growing availability of consumer hardware such as smartphones and mobile computers, rising Internet access across demographics, as well as a continuous shift from legacy media to digital media usage. The result is a media environment that consistently generates information at exponential scale, propelled by waves of consumer background and usage data, “a shift from an era of scarcity of audience data to an era of overabundance.”<sup>222</sup> To put this in perspective, it is instructive to consider the scale of the data boom over time.

- *Internet Usage*: From 2000 to 2019, the number of Internet users in the U.S. grew from 50% to over 90%, with time spent online growing from a daily average of 37 minutes to over 6 and a half hours.<sup>223</sup>
- *Expanded Access*: Broadband technology, mobile cellular networks, and cloud computing have expanded Internet access across the U.S. Additionally, the rise of smartphones has mainstreamed mobile Internet access. In 2010, about 60 million people owned a smartphone. In 2019, it was over 260 million people.<sup>224</sup>

---

<sup>222</sup> Havens, Timothy, “Media Programming in an Era of Big Data,” *Media Industries Journal*, Volume 1, Issue 2, 2014.

<sup>223</sup> See, Dan Packer & Lee Rainie, “More Online, Doing More,” Pew Research, February 18, 2001, accessed June 6, 2020, <https://www.pewresearch.org/internet/2001/02/18/more-online-doing-more> and Dolliver, Mark, “US Time Spent with Media 2019,” eMarketer, May 30, 2019, accessed June 6, 2020, <https://www.emarketer.com/content/us-time-spent-with-media-2019>

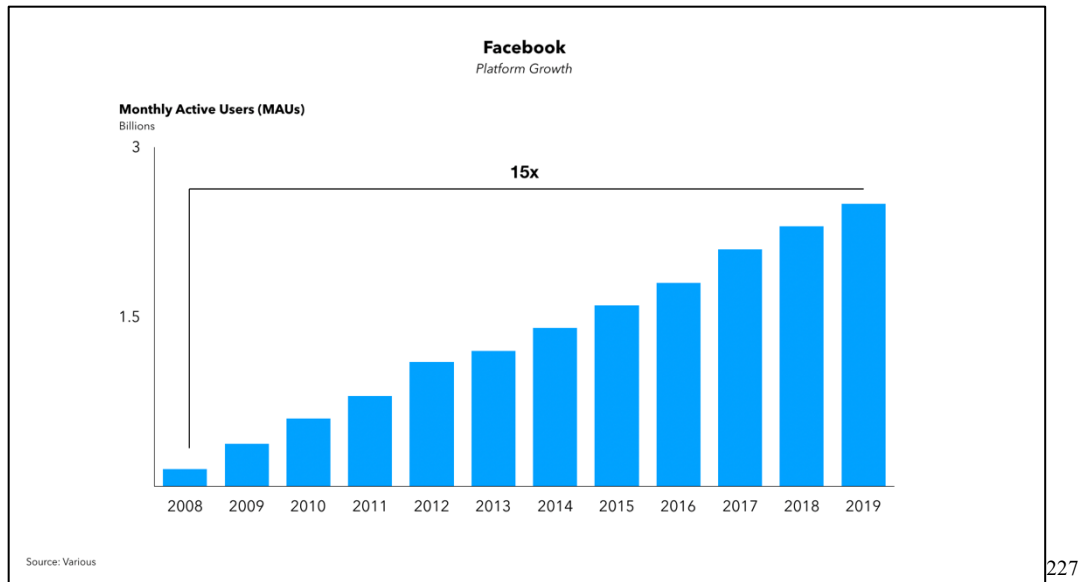
<sup>224</sup> Statista, “Number of smartphone users in the United States from 2018 to 2024,” April 21, 2020, accessed June 6, 2020, <https://www.statista.com/statistics/201182/forecast-of-smartphone-users-in-the-us/>

- *Mobile Usage:* Digital activity on mobile devices has grown exponentially from 54 million users in 2010 to over 132 million users in 2017, making it one of the fastest growing behavioral developments in the digital era.<sup>225</sup> The shift to mobile was driven by the growing availability of mobile devices (e.g., smartphones, tablets), the growing functionality of mobile apps and games, and the rise of mobile video viewing.
  
- *Video Streaming:* Digital video streaming has increased with the commercial development of video-on-demand platforms. Subscription video-on-demand (SVOD) platforms and ad-supported video-on-demand (VOD) have established themselves as alternative viewing models to broadcast and cable television, commanding an increasing share of viewers. For example, Netflix grew its paid SVOD subscribers from 21.5 million at the end of 2011 to 167 million at the end of 2019, a 7x increase.<sup>226</sup>
  
- *Social Media:* Social media platforms have aggregated the largest share of online activity, consolidating billions of users who generate billions of posts, thereby flooding the networks with information. As an example, Facebook grew its number of monthly active users from 100 million in 2008 to over 2.4 billion at the end of 2019 (figure 2.5).

---

<sup>225</sup> For a detailed data overview of the United States, see, OECD, “Mobile broadband subscriptions per 100 inhabitants Q4 2009 – Q4 2019 (Indicator),” accessed June 6, 2020, <https://data.oecd.org/broadband/mobile-broadband-subscriptions.htm>

<sup>226</sup> Netflix, accessed June 6, 2020, <https://www.netflixinvestor.com/ir-overview/profile/default.aspx> and Lee, Edmund, “Netflix Reports a Subscriber Bump,” *New York Times*, January 21, 2020, accessed June 6, 2020, <https://www.nytimes.com/2020/01/21/business/media/netflix-q4-2019-earnings-nflx.html>



**Figure 2.5:** Facebook’s global audience base has grown exponentially over the past decade.

Hollywood consequently faced ever-growing amounts of digital information on its audience and content. While the surplus in data was generally characterized as an opportunity to optimize business operations and create economic value, as exemplified by the success of platforms, analog industries like the legacy media business, with little to no existing technological infrastructure to manage the information increase, were presented with the unprecedented challenge of making sense of all the information. As one industry executive wryly noted, “no data, more problems.”<sup>228</sup> Indeed, the majors faced dynamics of unknown scale and complexity. While studios and networks previously sampled aggregate statistics from small groups of audiences, from the lower thousands downwards, the digital landscape flooded them with millions of data points, carrying much more granular and complex information, thereby

<sup>227</sup> Facebook Investor Relations, accessed June 6, 2020, <https://investor.fb.com/home/default.aspx> and Carlson, Nicholas, “Facebook Reaches 500 Million Users,” *Business Insider*, May 17, 2020, accessed June 6, 2020, <https://www.businessinsider.com/facebook-reaches-500-million-users-2010-5>

<sup>228</sup> Variety Staff, “Big Data Summit: Metrics, Analytics, ‘Wild West’ Opportunities Parsed by Industry Leaders,” *Variety*, November 10, 2016, accessed June 6, 2020, <https://variety.com/2016/tv/news/big-data-summit-facebook-vimeo-twitter-1201915456/>

fundamentally revising traditional forms of research and analysis. As another industry executive pointedly summarized, “Trying to extract actionable intelligence from all that data is like trying to find a very tiny needle in the world’s largest haystack.”<sup>229</sup>

Accordingly, the majors increasingly turned to a differentiated portfolio of analytics companies to make sense of the growing amounts of digital information in the media landscape. Following the established approach of contracting with market research firms, studios and networks *outsourced* the analysis of data to specialized vendors. At the same time, the partnerships with analytics companies expanded the legacy media industry’s traditional research operations. The majors incorporated data analysis software and hired data professionals to manage the technology and tools in-house, increasingly adopting an automated and scalable data-as-a-service approach. They effectively *insourced* software technology and talent, reconfiguring the industrial approach to collecting, capturing, and analyzing information from a relationship model to an integrated infrastructure of automated (i.e., technological tools and expertise) capabilities.

By adapting the SaaS logic, the legacy media industry shifted from an external to an internal model of licensing information, which was marked by a distinct set of shifts:

- *From outsourcing research work to insourcing data analysis capabilities:* the majors partnered with analytics companies to license software tools and established teams of data analysts versed in managing the respective vendors and tools.
- *From manual information-gathering (i.e., focus groups, phone/online polls, surveys) to software automation and computational analysis:* the majors expanded their strategies to

---

<sup>229</sup> IBM, “The Race to Probe the Twittersphere,” *The Atlantic*, accessed June 6, 2020, <https://www.theatlantic.com/sponsored/ibm-transformation-of-business/the-race-to-probe-the-twittersphere/280/>

collect, gather, and analyze data, deploying software to manage the growing amounts of digital information and generating insights with advanced computational methodologies.

- *From operating within an information monopoly to a rhizomatic network of solutions:* the majors historically worked with a small set of market research firms that maintained holistic control over the process (e.g., National Research Group and Nielsen), but the accelerating pace and drive for innovation in the digital media economy created a cottage industry of analytics startups.

The majors increasingly worked to establish an in-house infrastructure of data analysis capabilities, setting up a corporate-organizational environment that bore increasing resemblance to the operational dynamics of the technology industry, in which “decisions will increasingly be based on data and analysis rather than on experience and intuition.”<sup>230</sup> As such, studios and networks gradually adapted the industrial setup of the tech enterprise, expanding the traditional research approach to prioritize data analysis as a core lever in how the business operates. Effectively, the majors set the foundational stage for a data-driven enterprise, an organizational model that utilizes data analysis across all areas of the business. The legacy media industry’s iteration incorporated two layers.

- *Talent:* The majors worked to establish internal technology expertise to enable data analysis as a corporate-industrial process across the entire enterprise. To this end, they took a two-pronged approach, implementing a senior leadership structure with a strategic

---

<sup>230</sup> See, Lohr, 2012

vision for data at the organizational level (i.e., top-down) as well as creating teams of data analysts versed in analytical practices (i.e., bottom-up).

- *Top-down*: The majors established senior analytics positions and hired seasoned professionals to set a data analysis strategy, manage partnerships with analytics companies, and lead teams in analytical execution. The top-down approach set an organizational structure for data analysis within the legacy structure of film and television companies, while equally granting the majors credibility in a digital media landscape increasingly flooded with data.
- *Bottom-up*: The majors established data analytics departments and hired a new workforce of analysts at varying levels, from entry-level to advanced, in order to manage the output of data analysis software.
- *Technology*: The majors worked to establish data analysis software as a key element of Hollywood's research system, from partnering with analytics companies to license virtual software access to acquiring analytics companies and even building data analysis tools from the ground up. For studios and networks, data analysis software represented an investment in achieving "data literacy,"<sup>231</sup> a media enterprise capable of translating information into insight to inform strategic decisions. Accordingly, the majors managed a spectrum of relationships with analytics companies, gradually moving capabilities in-house to increase their control and lessen reliance on the technology industry.

---

<sup>231</sup> Dickey, 2012



- *Licensing Partnerships*: The majors partnered with platforms and startups to test a wide range of data analysis software tools, which reflected both a surplus of choice as well as shifting analytical strategies within studios and networks.
- *Acquisitions and Acqui-Hires*: The majors made a small number of acquisitions to expand their data analysis capabilities and install seasoned executives and teams within their organizational structures in order to elevate the role and value of data analysis within the legacy media industry.
- *Technology Developments*: The majors deployed their IT resources to build new data analysis software and create a competitive differentiator against technology companies and platforms.

Collectively, these strategic developments reconfigured the legacy media industry's established research system, established the industry's data analysis capabilities, and created a foundation for an analytical approach to corporate decision-making.

### Data Analysts in Hollywood

Hollywood pursued a two-pronged strategy in building up its data analysis workforce and infusing data expertise and sensibilities within their corporate organizational structures. From the top down, the majors focused on installing senior leadership to set a vision for data analysis across the enterprise, implement a departmental structure to execute against the strategic vision, advocate for data analysis across the organization, and build credibility with internal teams (e.g., research executives, production leaders, heads of strategy) and external partners (e.g., technology companies). From the bottom up, the majors hired a growing number of data analysts to manage

relationships with analytics companies, maintain software tools, and perform data analysis. As such, the legacy media industry followed an integrated holistic approach to establish a data analysis talent pipeline.

Over the course of the previous decade, the majors had established executive ranks focused on data analysis. Based on an analysis of publicly available data from LinkedIn for the time period of 2012-2016, the leading film studios and television networks employed at least one high-level executive with a focus on data analysis.<sup>232</sup> Within their job descriptions, several executives highlighted thought leadership as a key responsibility, indicating a remit to advocate for data analysis as a key lever within the organizational structures of legacy media companies. The majority of executives described leading teams as one of their core objectives, signaling the majors’ growing commitment to data analysis as an analytical team practice rather than the work of siloed individuals. Finally, many executives had a professional background in data analysis, either through their education (i.e., degree in a quantitative field like statistics, economics, mathematics, data science, or computer science) or previous work experience (i.e., work at technology companies or platforms) (table 2.3).

**Table 2.3**

<b>Data Analysis Leadership</b>		
<b>Key Responsibilities</b>	<b>Professional Experience</b>	<b>Education</b>
<ul style="list-style-type: none"> <li>● Thought leadership</li> <li>● Partnership development</li> <li>● Technology management</li> <li>● Financial management</li> <li>● Team building</li> </ul>	<ul style="list-style-type: none"> <li>● Experience at technology companies (i.e., platforms, data analytics companies)</li> <li>● Experience at digital retailers</li> </ul>	<ul style="list-style-type: none"> <li>● Graduate degree (i.e., M.B.A., M.S.)</li> <li>● Technical, quantitative field (i.e., business, engineering, computer science)</li> </ul>

<sup>232</sup> High-level executive is defined as Vice President or above. I used the terms *data analysis*, *data analytics*, and *data science* to filter search results in LinkedIn. Studios and networks included in the search analysis were Disney, Fox, Warner Bros., Universal Studios, ABC, CBS, NBC, HBO, and Showtime.

The presence of senior leadership did not instantly establish data analysis as a corporate-organizational priority. The new brass had to negotiate for cultural capital within the legacy media industry, effectively assimilating into Hollywood's culture of impressionism. As one data executive remarked on her experience in the legacy media business: "You can't come off as 'I'm this math person, you can't know what I know.' If you don't speak a language they can understand, they just don't listen to you."<sup>233</sup> As such, data executives had to adjust to the normative dynamics of legacy media organizations while consistently advocating for the use of data analysis at the enterprise level. As seen in the initial Legendary example, the integration of data analysis as a supporting mechanism for corporate decision-making in Hollywood is dependent upon the investment of company leadership at the C-suite level. Thomas Tull, founder and Chief Executive of Legendary at the time, declared data analysis a priority for the company, giving the internal analytics team a mandate to expand the organization's legacy approach through data analysis. The lack of top-down direction, by contrast, could stymie data analysis efforts, leaving teams to operate in isolation, siloed, and thus disconnected from the business. In describing her experience of running a large-scale data analysis department at a major studio, Julie Rieger, the former head data scientist at 20th Century Fox, emphasized the criticality of leadership support (my emphasis):

**"We have a CEO, Stacey Snider, who gets it. If you don't have that, then you can kick and scream about data as much as you want, but it'll never be a company-wide priority. We've been lucky with Stacey, because she wanted to bring Silicon Valley data sensibilities to the film industry. In fact, Stacey spends quite a bit of time in the data science lab with us, where we'll make decisions together based on the data."**<sup>234</sup>

---

<sup>233</sup> Variety Staff, 2016

<sup>234</sup> Think with Google, 2018

While the creation of executive presence in data analysis did not ensure the holistic integration of an analytical corporate-industrial mindset, the majors still established a working infrastructure to discuss data analysis at the level of major decision makers within legacy media organizations.

On the other side of the corporate spectrum, the majors increased their hiring quota of data analysts to manage the growing amounts of data and establish an analytical engine to translate digital information into insights. The data deluge had led to an exponential demand for data professionals across all industries, and Hollywood proved no exception.<sup>235</sup> Indeed, the legacy media industry acknowledged a need for data analysts to help make sense of the information waves hitting the media business from all sides. As Variety put it (my emphasis):

**“The volume of data available about consumer behavior emerging from online sources, set-top boxes, smartphones and old-fashioned retail sales has become a tidal wave that threatens to overwhelm even the most sophisticated analysts. The need for expertise in sorting through the deluge for kernels of insight is turning data scientists into the A-listers of media and entertainment.”**<sup>236</sup>

Data analysts were characterized as A-listers and “rock stars of your business,” highlighting the growing capital ascribed to data analysis in the legacy media business.<sup>237</sup>

While data leaders are focused on setting the strategic vision for an analytical organization that draws on data analysis to guide decision-making, data analysts are part of teams dedicated to performing the technical analysis. As such, their mandate focuses on

---

<sup>235</sup> See, Knowledge@Wharton, “What’s Driving the Demand for Data Scientists?,” *Wharton School of Business*, March 8, 2019, accessed June 6, 2020, <https://knowledge.wharton.upenn.edu/article/whats-driving-demand-data-scientist/>

<sup>236</sup> Littleton, 2015, “10 Things We Learned at Variety’s Big Data Summit”

<sup>237</sup> Ibid

managing relationships with analytics companies, maintaining data analysis software, performing technical data analysis (i.e., statistical analysis of digital information on media programming and audiences). To contextualize the role of the data analyst in Hollywood's legacy media industry, the following presents a set of key skills and responsibilities, accumulated from official job descriptions in the legacy media industry as compiled on professional employment-based online services during the period from 2014-2016.<sup>238</sup>

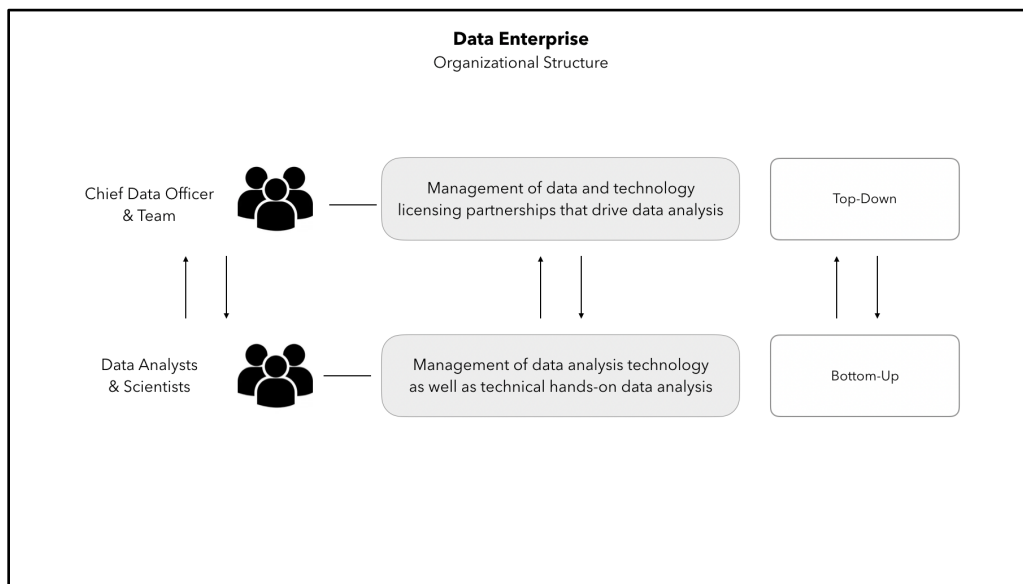
- *Analytics Experience*: Data analysts should have experience working on data analysis projects, with hands-on practice in collecting, gathering, and analyzing large sets of data.
- *Degree in Quantitative Field*: Data analysts should have a BA/BS or advanced degree in mathematics, statistics, economics, finance, computer science, or other quantitative discipline. The data analyst role is built on expertise in highly analytical fields.
- *Familiarity with Analytics Software*: Data analysts need to be familiar with a software tools, including best-of-breed solutions (i.e., tools that analyze data across all platforms) and point solutions (i.e., tools that analyze specific platform data, such as social media).
- *Knowledge of Programming Languages*: Data analysts need experience with programming languages, including SQL (Structured Query Language) to pull data from large-scale databases and scripting languages like Python or R for statistical modelling.

The professional profile of a data analyst thus places emphasis on technical skills, with a specific focus on analytical methodologies and tools. Hollywood built a workforce focused on collecting,

---

<sup>238</sup> Data analyst as a category is defined as including a range of job titles, including *Business Intelligence Analyst*, *Digital Analyst*, *Data Analytics Analyst*, and *Social Media Analyst*. For key job characteristics, I looked at official job descriptions from film studios and television networks on LinkedIn and jobs sites, such as Glassdoor and Indeed.

gathering, and translating information into insights. From the top down, executives worked to advocate a strategic vision for data analysis, developing relationships and capital to promote an analytical approach to decision-making within legacy media organizations. From the bottom up, teams of data analysts managed the ongoing influx of information, generating insights to help sell in the vision of a new analytical organization (figure 2.6).



**Figure 2.6:** Data leadership (top-down) and data analysis (bottom-up) structures form the interdependent professional layer of the emerging analytical enterprise organization in the legacy media industry.

While Hollywood established a working infrastructure to incubate data talent, the legacy media industry’s quest for data analysts proved challenging and complex, for several reasons. One, the legacy media industry was effectively developing a high-level strategic approach to data analysis while hiring professionals to manage the growing amounts of digital information effectively. As such, the majors maintained a test-and-learn model that likely required frequent adjustments and changes to the organizational structure. Two, data professionals come with a professional background that is distinctly different from Hollywood’s general workforce, which requires ongoing negotiation of differing corporate approaches and mindsets. And, three, the

majors had to actively compete with technology companies for data professionals, which required new sets of HR incentives and outreach strategies.<sup>239</sup>

### Data Analysis Software in Hollywood

Hollywood systematically invested in the development, cultivation, and integration of data analysis software in order to support a more analytical corporate focus. This process proved complex and non-linear, driven by a key set of strategies: strategic partnerships, acquisitions and acqui-hires, and technology development.

### *Strategic Partnerships*

The legacy media industry built partnerships with a wide range of analytics companies to license software tools that enabled the analysis of data across the digital landscape. The partnership model gave the majors a broad set of capabilities to analyze multiple data sources, from websites to social media, mobile apps, and streaming devices, such as OTT services and set-top boxes. As such, the majors established a system to generate insights on the interplay between their content and audience across the entire digital media economy. Furthermore, partnering with analytics companies ensured the automatic delivery of insights reports as well as direct virtual access via insights dashboards, minimizing the workload for the majors. At the same time, the majors had the ability to conduct their own analysis, using a variety of data analysis software tools licensed by analytics companies. While granting flexibility and multi-functionality, this approach equally

---

<sup>239</sup> For example, Disney launched an internal data analytics training program in 2001 and later expanded it into an industry event, the annual Data & Analytics Conference. See, Disney Institute Blog, “A Closer Look at the Disney Data & Analytics Conference,” *Disney Institute*, January 24, 2020, accessed June 6, 2020, <https://www.disneyinstitute.com/blog/a-closer-look-at-the-disney-data-analytics-conference/>

proved costly and complex, requiring the majors to navigate multiple relationships and manage a multitude of software tools.

The legacy media industry partnered with a range of analytics vendors that can be classified into three categories.

- *Data Management*: Large B2B enterprise companies that enable the collection, organization, and analysis of large sets of internal data, which is generated and owned by the majors. Internal data financials, production data (e.g., casting, footage, logistics, asset management), distribution data (e.g., distribution costs, theatrical rentals) and advertising (e.g., advertising revenue, ad performance). Data management companies provide software tools that consolidate and analyze internal data, enabling the majors to get a holistic view of their business. Representative companies include Teradata, Google (via its Big Query software platform), and Microsoft (via its Azure software platform).
- *Platform Data*: Platforms provide data analysis software and customized dashboards to large enterprise customers, including the major studios and networks. Platform analytics dashboards translate the majors' content and audience information into platform-specific insights. For example, Facebook's Insights dashboard enables the majors to gain insight into their audience's demographics, video viewing behavior, and social media activity. Representative companies and their software tools operate across the digital landscape and can be grouped by category, including social media (e.g., Facebook Insights, YouTube Analytics) and mobile (e.g., Apple's App Analytics, Google's Play Console).
- *Analytics Startups*: Analytics startups specialize in data analysis software with a vertical focus. Companies like Crimson Hexagon, Fizziology, ListenFirst Media, and



Socialbakers focus on social media analysis. While platforms only provide analytical tools focused on their own platform data, analytics startups operate cross-functionally, utilizing platform APIs to conduct more expansive analyses across multiple platforms. For example, Crimson Hexagon’s social media software analyzes data from Facebook, Twitter, and YouTube, as well as public websites, blogs, and forums. Parrot Analytics, meanwhile, focuses on video analytics across platforms, analyzing viewing behavior. As such, analytics startups cast a wide analytical net across the digital landscape.

By partnering with multiple analytics companies across the digital landscape, the majors created an expansive analytical focus, which increased the output of insights, yet equally resulted in an “alphabet soup of competing analyses” for the majors that required ongoing evaluation of the software vendors and tools.<sup>240</sup>

### *Acquisitions and Acqui-Hires*

Hollywood worked to bring analytics software technology in-house with the acquisition of digital media companies specializing in analytics solutions. While costly in the short term, acquisitions were rationalized as a long-term investment in analytical capabilities. Specifically, acquisitions gave the majors access to proprietary software technology, which would lessen their reliance on analytics vendors and provide them with more pricing power and negotiation leverage in future partnership development. Furthermore, they would be able to implement the company’s workforce of data professionals in order to expand their growing analytics divisions.

---

<sup>240</sup> Barnes, Brooks, “Hollywood Tracks Social Media Chatter to Target Hit Films,” *New York Times*, December 7, 2014, accessed June 6, 2020, <https://www.nytimes.com/2014/12/08/business/media/hollywood-tracks-social-media-chatter-to-target-hit-films.html>

As such, the legacy media industry focused on acqui-hires, looking to harvest data analysis skills and expertise in addition to technological capabilities.

The majors acquired a distinct set of companies that expanded the legacy media industries capabilities in data analysis, retaining both the underlying technology and talent workforce. In particular, they focused on companies that provide data analysis services to the media and entertainment industry.

- *Multi-Channel Networks*: Multi-channel networks (MCNs) produce, program, and distribute digital content across platforms, with a particular focus on advertising-supported programming. As such, MCNs consistently work to illustrate the performance of their content to advertisers as a way to validate their investment and incentivize ongoing ad spend. To this end, MCNs developed proprietary analytics dashboards that provide insight into audience demographics, viewing behavior, and content performance across platforms. At the same time, MCNs employed a workforce of media professionals versed in analytical thinking, data, and insights. Notable acquisitions include Dreamworks' acquisition of AwesomenessTV (2014), the Walt Disney Company's purchase of Maker Studios (2014), and Warner Bros.' integration of Machinima (2016).
- *Digital Marketing Startups*: Digital marketing startups specialize in concepting, producing, and running digital marketing campaigns across platforms in order to promote new entertainment releases and talent. They operate across the digital landscape and, consequently, built proprietary analytics dashboards to translate online campaign and talent information into insights. Notable acquisitions include the Walt Disney Company's

merger with DigiSynd (2008),<sup>241</sup> Warner Bros.' integration of Flixster (2011),<sup>242</sup> and Viacom's buy-out of WhoSay (2018).<sup>243</sup>

- *Analytics Startups*: Analytics startups are a more recent acquisition target as the majors have increasingly invested their own technology development, moving from licensing to incubating analytics software solutions in an effort to scale data analysis across their organizational structures and establish it as a corporate priority. While Hollywood has not yet made major acquisitions, the majors have signed long-term deals and expanded their partnership structure, signaling a growing interest in analytics.<sup>244</sup>

By bringing data analysis capabilities in-house, the majors increasingly signaled their commitment to data analysis and analytical decision-making, though acquisitions frequently did not yield the desired effects, causing issues of integration due to differences in strategic vision, work culture, and operational capabilities.<sup>245</sup>

---

<sup>241</sup> See, Rafat, Ali, "Disney Buying Storytelling Social Net FanLib; DigiSynd in Process," *CBS News*, June 3, 2008, accessed June 6, 2020, <https://www.cbsnews.com/news/disney-buying-storytelling-social-net-fanlib-digisynd-in-process>

<sup>242</sup> See, Szalai, Georg, "Warner Bros. to Acquire Rotten Tomatoes Owner Flixster," *Hollywood Reporter*, May 4, 2011, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/warner-bros-acquire-rotten-tomatoes-185237>

<sup>243</sup> See, Bruell, Alexandra, "Viacom Acquires Whosay to Help Advertisers Create More Branded Content," *Wall Street Journal*, January 8, 2018, accessed June 6, 2020, <https://www.wsj.com/articles/viacom-acquires-whosay-to-help-advertisers-create-more-branded-content-1515420000>

<sup>244</sup> See, Siegel, Tatiana, "Warner Bros. Signs Deal for AI-Driven Film Management System," *Hollywood Reporter*, January 8, 2020, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/warner-bros-signs-deal-ai-driven-film-management-system-1268036>

<sup>245</sup> For an overview of the complex dynamics of digital media industries, see, Mann, Denise, "Welcome to the Unregulated Wild, Wild, Digital West," *Media Industries Journal*, Volume 1, Issue 2, 2014.

## *Technology Development*

As Hollywood's IT infrastructure and analytics capabilities matured, the legacy media industry incrementally built its own technology to collect, gather, and analyze data, gradually expanding from a licensing approach to an internal incubation of technology. Internal technology development gave the majors an opportunity to fine-tune software tools to specific business goals, create efficiencies within the organization, and secure buy-in and adoption. At the same time, this investment required substantial investment, including the development of software and data engineering teams, which proved more costly, complex, and challenging for the majors.

A prominent example of internal technology development comes from film production company Legendary. Following the acquisition of data analytics firm StratBridge, Legendary built a proprietary software tool called *Eddington*, designed to segment, identify, and target movie-goers with a high affinity for Legendary's films. The software enabled Legendary to optimize its marketing approach, at times reducing costs by up to 30%, while maximizing ticket sales and box office in the process. The company transformed the software into a software-as-a-service product and licensed the technology to other studios, networks, and brands.<sup>246</sup>

Another example comes from the world of television broadcasting. Television conglomerate Turner developed an internal data analysis software tool, dubbed *Launchpad*.<sup>247</sup> As part of an internal strategy to maximize advertising sales across Turner's expansive portfolio of social media accounts, Launchpad was presented as an integrated sales and analytics solution

---

<sup>246</sup> See, Jarvey, Natalie, "Legendary Hires Bankers to Sell Analytics Business," *Hollywood Reporter*, November 13, 2017, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/legendary-hires-bankers-sell-analytics-business-1057371>

<sup>247</sup> See, Patel, Sahil, "How Turner trained 500 employees to sell brand social videos globally," *Digiday*, June 15, 2017, accessed June 6, 2020, <https://digiday.com/media/how-turner-trained-500-employees-to-sell-branded-social-videos>

that enables advertisers to see the affinity of Turner's audience for their brand. In effect, Turner built a software product that leveraged multiple platform APIs to pull psychographic data from across social media platforms. The company used the data to create a segmentation of their social media audience, designed to help advertisers discover those fans with the highest affinity for their brand. In this regard, Turner utilized data analysis software to create audience insights that delivered added value to advertisers in their social media campaigns. As part of the technology development, Turner further trained its internal salesforce on how to use and present the software, effectively adopting a more analytical sales approach across the company.

Effectively, the legacy media industry moved from licensing to acquiring and building technology in-house, signaling a growing effort to control the data value chain and vertically integrate analytical capabilities. By investing in a data-as-a-service infrastructure and automating data analysis across their organizational structures, the majors gradually developed a foundation to expand their operational focus, from a purely impressionistic creative business to an increasingly analytical and data-driven enterprise.

Jesse Redniss, former Executive Vice President of Data Strategy at WarnerMedia, experienced the process of making data analysis an enterprise-wide initiative across media organizations first-hand. Following AT&T's 2018 acquisition of Time Warner, the company consolidated multiple media businesses into one unit, merging film production company Warner Bros., news network CNN, cable channel HBO, digital arm Otter Media, and television network Turner into the supra-structure of WarnerMedia. As part of the process, Redniss and other data executives partnered on a holistic data strategy to deploy analytical capabilities across the enterprise. He summarized the approach succinctly (my emphasis):

“We’re starting to look at each other’s roadmaps and platforms that we use and the way that our different data scientists are looking at methodology in identifying people and identifying households. Across WarnerMedia, **we’re all coming together.**”<sup>248</sup>

--- --- ---

While the adoption of a data-as-a-service approach has enabled the majors to increasingly leverage data analysis as a lever in corporate decision-making, the overall process has proven complex, winding, and difficult given the legacy media industry’s entrenched focus on established, more impressionistic parameters, such as creative decisioning and talent relationships. Indeed, data analysis has largely remained incremental, rather than central, to Hollywood’s business strategy. With the exception of Legendary, none of the major studios and networks had a chief data executive guiding the corporate strategy of the majors. Furthermore, the majors lacked a holistic data strategy for the entire organization, deploying multiple analytical approaches across individual teams and departments, which resulted in data silos rather than integrated, enterprise-wide efforts. As a result, it appears that “most entertainment firms have little institutional strength or political capital with data analytics.”<sup>249</sup>

Overall, Hollywood’s investment in data-as-a-service infrastructure effectively expanded the majors’ analytical capabilities, promoted an industrial approach to using data analysis in corporate decision-making, and thereby laid the foundation for an emerging culture of analytics. In effect, the majors built on and extended the legacy media industry’s existing approach to market research, deploying a suite of resources and capabilities to collect, gather, and analyze data, and extract valuable insights on audience behavior and content performance. With the

---

<sup>248</sup> Variety, “Turner’s Jesse Redniss on How AT&T Changes the Data Game,” *Strictly Business Podcast*, October 9, 2018, accessed June 5, 2020, <https://variety.com/2018/digital/news/listen-turners-jesse-redniss-on-how-att-changes-the-data-game-1202973028/>

<sup>249</sup> Smith & Telang, 2016

ongoing proliferation of data across the digital landscape, the majors increasingly had to navigate, manage, and make sense of a media environment defined by data deluge. This corporate-industrial process fundamentally expanded and revised how Hollywood classified, processed, and evaluated data.

### **Data as Intelligence: The Rise of Data Metrics**

Hollywood's 1997 summer movie season was a financial windfall for the industry. Filled with a string of blockbuster hits and successful mid-budget releases, the period generated successful returns, led by tentpole titles such as *The Lost World: Jurassic Park* (Universal Pictures), *Men in Black* (Sony Pictures), and *Air Force One* (Sony Pictures), popular mid-size ventures like the Julia Roberts vehicle *My Best Friend's Wedding* (Tristar Pictures), and the low-budget indie smash *The Full Monty* (20th Century Fox). At the time, many industry executives and analysts ascribed the summer's box office success not only to the quality of the films, but a critical scheduling shift. In May, Paramount Pictures announced that its summer movie epic, James Cameron's *Titanic*, would be delayed to the end of the year, reportedly causing relief among executive ranks worried about an overcrowded summer schedule. As the *New York Times* noted: "Studio executives all over town, alarmed about the glut of big-budget movies being rolled out virtually every weekend, awoke this morning with the happiest news of spring. The opening of "Titanic," the James Cameron epic that has had technical problems, has been moved from July 2 to Dec. 19."<sup>250</sup>

---

<sup>250</sup> Weinraub, Bernard, "As Problems Delay 'Titanic,' Hollywood Sighs in Relief," *New York Times*, May 29, 1997, accessed June 6, 2020, <https://www.nytimes.com/1997/05/29/movies/as-problems-delay-titanic-hollywood-sighs-in-relief.html>

*Titanic* was not only seen as a risk. Some perceived it as a critical miscalculation, a budding disaster of a disaster movie that could reverse the upward trajectory of the movie business. Drawing comparisons to infamous box office bombs like *Cleopatra* (20th Century Fox, 1963), *Heaven's Gate* (United Artists, 1980), and *Waterworld* (Universal Pictures, 1995), the film was rumored to be an excessive overreach marked by creative hubris and financial gluttony. Early test screenings reportedly led to a severe cutting exercise to adjust the film's length as well as a new ending, raising doubts about its mainstream audience appeal. The trades reported on production delays, corporate in-fighting, and a vision gone awry. As the Washington Post summarized, "[e]ven before anyone has actually seen it, the movie has engendered its own legend of ego and excess, both the personal and financial kind."<sup>251</sup>

Yet, contrary to the industry insider narrative, *Titanic* proved an unprecedented hit, defying initial speculation. It became the most successful film of 1997, emerged as the most successful film of the decade, and, at the time, gained the title of the world's most successful film, grossing over \$2.2 billion at the box office. To this day, it remains the third most successful film of all time, based on theatrical revenue adjusted for inflation, outranked only by David O. Selznick's production of *Gone with the Wind* (Loew's Inc., 1939) and James Cameron's own *Avatar* (20th Century Fox, 2009). Additionally, *Titanic* received an A+ CinemaScore, a statistical rendering of the film's audience reception during its theatrical run, which put it in the upper 1% of the most popular films within that framework.<sup>252</sup> At the 1998 Academy Awards, *Titanic* received 14 nominations, tying the previous record holder *All About Eve* (20th Century

---

<sup>251</sup> Sharon Waxman & Paul Farhi, "Going Down With The Ship?," *Washington Post*, May 25, 1997, accessed June 6, 2020, <https://www.washingtonpost.com/archive/business/1997/05/25/going-down-with-the-ship/5c0ed22f-9c2a-4420-8c4a-b2b983eea681/>

<sup>252</sup> For more information on CinemaScore, see company website: <https://www.cinemascore.com/cinemascores/>



Fox, 1950) and won 11 Oscars, tying *Ben Hur* (Loew's Inc., 1959.) for the highest amount ever awarded.<sup>253</sup> In effect, *Titanic* became the most critically acclaimed and financially successful film of its time.

While the legacy media industry did not foresee the phenomenon that was *Titanic*, Hollywood arguably did not miscalculate its assumptions about the film's projected performance. Indeed, the industry considered a set of existing legacy benchmarks in prognosticating the film's supposed trajectory from expensive special effects vision to inflated, over-budget experiment. Industry executives were skeptical about *Titanic* because their analytical lens focused on a simple, tried-and-true formula to manage the dynamics of the legacy media business:

*Historical setting x troubled production x delays and re-shoots x an overdrawn budget = box office failure*

Based on this model and associated inputs, *Titanic* projected a risk of financial disappointment. Indeed, based on Hollywood's standard operating formula, the film fulfilled many criteria of a pressured, high-risk investment. It lacked a modern setting for contemporary audiences to relate to; it lacked established stars; it had a troubled production; it did not meet its original release date; it was too long; test audiences disliked a pivotal scene; it went over budget, which was already far too expensive by any past standard. As such, *Titanic*'s unexpected success made it an anomaly within Hollywood, an outlier that defied the logical parameters of the business, a deviance from the industry's operating formula and decision-making guide.

---

<sup>253</sup> In 2004, *The Lord of the Rings: The Return of the King* (Warner Bros., 2003) became the third film to win 11 Academy Awards and the first to perform a clean sweep, winning within all of its nominated categories.

Accordingly, *Titanic* offers a useful starting point into Hollywood's long-standing approach to information-gathering and corporate decision-making. The legacy media industry has long cultivated a system to track and inform the state of the business, with a focus on financial performance as well as reception/audience feedback. In the film industry, studios have actively tracked box office results since the inception of the studio system in the 1920s. Trade publications like *Variety* began reporting estimates of film rentals early on and later published financial reports on a weekly basis. *Variety's* box office reporting morphed into a staple of industry coverage, with *Entertainment Weekly* subsequently publicizing results to a broader audience. At the same time, market research firms, notably Entertainment Data, Inc. or EDI (previously known as National Gross Service), started to tabulate domestic gross receipts in the 1980s, eventually expanding to worldwide reporting, delivering a unifying benchmark for the industry. Studios continue to collect, calculate, and communicate box office on a weekly basis.<sup>254</sup>

Furthermore, studios have used a suite of research practices to track viewer feedback on their content output, such as early test screenings and focus groups, telephone and online opinion polls, and various survey formats, creating connective tissue between financial performance and audience reception. A notable example is CinemaScore, a film measurement company founded in 1979, which surveys a small set of audiences in major metropolitan areas the first weekend of a film's release, tracking viewers' demographic background, movie-going habits, and personal film evaluation in form of a letter grade.<sup>255</sup>

---

<sup>254</sup> See, Hayes, Dade & Bing, Jonathan, *Open Wide: How Hollywood Box Office Became a National Obsession* (New York: Miramax Books, 2004).

<sup>255</sup> On Hollywood's approach to market research, see, Wyatt, 2014 and Bakker, Gerben, "Building knowledge about the consumer: The emergence of market research in the motion picture industry," *Business History*, 45, 1, 2003.

The television industry, meanwhile, has largely relied on information input from measurement giant Nielsen, a de facto research monopoly that services all broadcasters.<sup>256</sup> The company started tracking TV programming and audience setup in the 1950s, developing a proprietary ratings system to determine a show’s weekly estimated viewers, their demographic bracket (i.e., adults under 18, aged 18-49, and above 50), and overall content performance compared to competing titles. Nielsen ratings determine the advertising rates of the entire broadcast ecosystem, setting prices for ad time based on audience size, share, and demographic background. As such, Nielsen ratings effectively represent the standard reference for a television show’s popularity and financial performance. Additionally, networks have long utilized pilot testing and research studies to evaluate ideas and project financial returns, pressure-testing programming based on initial audience input in a controlled environment.

In this regard, the legacy media industry has operated an established system of information-gathering on the interaction between content and audience, consistently tracking film and television performance as well as audience feedback. While Hollywood has carefully crafted, curated, and conserved the primacy of its creative DNA, the business has always been a “numbers game.”<sup>257</sup> Indeed, studios and networks have invariably drawn on content and audience information to validate investment theses, greenlight productions, and project returns.<sup>258</sup> As one publication put it succinctly, “Hollywood is driven by numbers—Did a movie surpass \$100 million at the box office? How many 18- to 49-year-olds watched last night?”<sup>259</sup>

---

<sup>256</sup> See, Buzzard, 2012

<sup>257</sup> Grover, Ronald, “The Hollywood Numbers Game,” *Bloomberg*, August 2, 2001, accessed June 6, 2020, <https://www.bloomberg.com/news/articles/2001-08-02/the-hollywood-numbers-game>

<sup>258</sup> See, Havens, 2014

The majors have long gathered information on content output (i.e., films, television shows) and audience input (i.e., the number of movie tickets sold, the number of viewers for a television show, viewer feedback on a new film or television show) to manage the business and guide corporate decisions. Significant box office returns (i.e., a sizable profit after expenses) can fast-track new productions based on the same IP or creative team. Television shows with high ratings and/or specific demographics get a greenlight or renewed. Dedicated fan bases drive sequels, spin-offs, and reboots. In reverse, output that does not meet those criteria foregoes renewals or extensions. Hollywood's math is not only simple, but functional, enabling the majors to calculate risk based on a small set of key benchmarks. As such, box office, ratings, and audience feedback do not only represent the key performance indicators of success in Hollywood, the so called "empirical truth"<sup>260</sup> and "lifeblood of the industry,"<sup>261</sup> they also function as the key decision-making variables of the business that guide the forward-motion of cultural production.

While Hollywood is said to operate "from the gut," the legacy media industry has long managed the business under the guidance of a key set of metrics, following an established measurement framework (table 2.4). Indeed, the majors have consistently measured content performance and audience reception, processing box office and ratings reports on a weekly basis, and commissioning numerous audience research studies at all stages of cultural production, from early test screenings and pilots to tracking surveys and exit polls, thereby documenting viewer

---

<sup>259</sup> See, Laporte, Nicole, Netflix: The Red Menace," *Fast Company*, January 7, 2014, accessed June 5, 2020, <https://www.fastcompany.com/3024158/netflix-the-red-menace>

<sup>260</sup> Hayes & Bing, 2004

<sup>261</sup> See, Ng, 2019

feedback and demographics. Thus, studios and networks have invariably drawn on content and audience information as a reference point and analytical variable in managing the business.

**Table 2.4**

Legacy Media Measurement Framework		
Context	Metric	Data
Film/Television Performance	Box office	Box office gross receipts
	Ratings	Nielsen ratings
Audience Reception	Feedback	Reviews Polls/surveys/focus groups Other (e.g., fan mail)
	Demographics	Age and gender based on the 4-quadrant model (over/under 25; male/female)

The digital era has simultaneously cemented and reconfigured Hollywood’s entrenched measurement framework. On the one hand, the rise of digital technologies has given increased prominence and transparency to legacy metrics. In 1997, Bruce Nash, a computer programmer and film fan, launched *The Numbers*, the first website to track box office numbers. The site created a visual display case for theatrical revenues, delivering regular updates on studios’ financial performance, thereby extending Hollywood’s core success metric into the digital world (figure 2.7). *The Numbers* spawned other online box office trackers, such as *Box Office Mojo*, *Box Office Report*, and *Box Office Guru*, while its open API enabled the use of box office numbers across the digital sphere, making it a central element of popular film discourse.<sup>262</sup> The

<sup>262</sup> See, Raviv, Shaun, “Moneyballing the Movies: How the Box Office Became a Sport,” *The Ringer*, August 2, 2018, accessed June 6, 2020, <https://www.theringer.com/movies/2018/8/2/17641822/box-office-reporting-mojo-the-numbers-marvel-star-wars>

concept was later expanded into the television business with the launch of the ratings tracking site *TV by the Numbers* in 2008, founded by Robert Seidman, a statistical analyst, and Robert Gorman, an Internet executive. Both sites effectively ran and reported the numbers on the business of Hollywood online.<sup>263</sup>

In addition to content performance metrics, the Internet increased the exposure of audience metrics. In 1998, three Berkeley undergraduate students, Senh Duong, Patrick Y. Lee, and Stephen Wang, founded *Rotten Tomatoes*, an online aggregator of critic and audience feedback. *Rotten Tomatoes* aggregates film reviews from critics and regular movie-goers, quantifies the information into a percentage, and outputs the result as a numerical score for critics (i.e., the Critic Score) and fans (e.g., the Audience Score). Building on IMDb's popular audience forum feature, introduced in the early 1990s, *Rotten Tomatoes* effectively displayed a centralized repository of audience reviews, opinions, and comments. Initially focused on the film industry, the site expanded to television in 2013.

Aggregators like *The Numbers* and *Rotten Tomatoes* became a central element of media industry discourse, infusing quantitative performance into critical evaluations and industrial developments.<sup>264</sup> Indeed, *Rotten Tomatoes* data was further deployed for marketing purposes by studios and networks, referencing scores in trailers and digital promotions. Additionally, critic and audience scores were utilized as shorthand for success, joining box office and ratings in distilling the corporate story of a film and television show in an instant.

---

<sup>263</sup> *The Numbers* continues to operate as an aggregator of box office and home entertainment data.

<sup>264</sup> The influence of quantitative aggregators on the film and television industry grew exponentially over time, with recent debates focusing on their impact on film and television financial performance. See, Barnes, Brooks, "'Attacked By Rotten Tomatoes,'" *New York Times*, September 7, 2017, accessed June 6, 2020, <https://www.nytimes.com/2017/09/07/business/media/rotten-tomatoes-box-office.html>

*The Numbers* and *Rotten Tomatoes* constitute industrial feedback mechanisms that collect, consolidate, and conform information on Hollywood's content and audience.<sup>265</sup> Effectively, they represent vast data repositories of film and television industry metrics, covering box office, ratings, and audience feedback by tracking reviews and comments. Rather than digital display cases of information, the sites' focus on specific industry data positions them as *digital trackers of business intelligence*. They automatically screen, save, and surface the core metrics of the legacy media industry, framing information as a valuable form of intelligence that drives corporate decision-making. Put differently, they carry the industry's most valuable information, the data that studios and networks consistently draw on to manage all stages of cultural production and move the business forward, the core metrics that define success. In this regard, the commercialization of industry feedback mechanisms, enabled by digital technologies, has significantly increased the exposure and prominence of legacy metrics.

Yet, on the other hand, digital technologies have fundamentally reconfigured the dynamics of Hollywood's legacy metrics, effectively testing the viability of the existing content performance and audience reception framework. Indeed, over the course of the 21st century, the legacy media industry experienced an unprecedented surge in information that challenged its established measurement framework and, by consequence, the way it has measured and managed industry information.

- *Datafication*: Digital technologies introduced a range of feedback mechanisms that consistently convert online activity into quantifiable information, effectively distilling all online actions into an ever-growing ecosystem of data. As a result, the legacy media

---

<sup>265</sup> See, Napoli, 2010

industry faced an industrial environment where content and audience information is consistently traced, tracked, and translated as digitally-coded information.

- *Web 2.0*: The development of interactive digital environments enabled the ongoing exchange between audiences and content owners, yielding an influx in data points on audience interaction with content, including new levels of detail on audience reception.<sup>266</sup> For example, legacy modes of research gathered audience feedback via a set of directed survey questions, while the new digital environment generated waves of unprompted, unfiltered reactions and comments across multiple digital channels and platforms.
- *New Distribution Mechanisms*: Digital technologies fundamentally expanded the legacy media industry's linear distribution model. As content was released across websites, social media, online video sites, mobile apps, and over-the-top services, Hollywood unlocked a variety of new data streams, delivering an unprecedented amount of information on performance and audience reception.
- *New Consumer Habits*: The rise in digital distribution mechanisms enabled a range of new ways to consume and interact with film and television programming, notably practices such as user-generated content and binge-watching, which, in turn, generated consumer and usage data far outside the controlled environment of linear film and television viewing.

---

<sup>266</sup> See, Luna, Taryn, "Social media play big role in movies," *Boston Globe*, March 12, 2013, accessed June 6, 2020, <https://www.bostonglobe.com/business/2013/03/12/movies-depend-social-media-support-for-staying-power-box-office/mDRqLV2AaS1xqmLdFV1N5O/story.html>



- *Rise of Platforms*: Digital platforms built wide-ranging virtual infrastructures of products and services, effectively creating alternative modes of video viewing and consumption, thereby opening up new information streams on the interaction between content and audience in the digital sphere. Consequently, the legacy media industry was presented with data repositories that operated outside of its core metrics, with box office receipts, ratings points, and legacy media reception losing their prominence as mainstream industry reference points.

Accordingly, the legacy media industry faced a media environment overloaded with data, delivering more volume (i.e., more data points) and depth (i.e., more detail per data point) and creating an array of new information categories, which exceeded the logic of the majors' legacy research framework.<sup>267</sup> In this new digital environment, measuring box office and ratings was no longer sufficient in determining the success of film and television programming; categorizing audience segments by age and gender did not capture the granularity of viewer backgrounds; tracking sold movie tickets and television tune-in rates did not reveal the full extent of media viewership. Indeed, the rise in data challenged the industry's information and measurement framework. As a result, the majors could no longer rely exclusively on their core metrics to rationalize the business (table 2.5).

---

<sup>267</sup> See, Canepa, Steve, "Big Data Is Going To Save The Film Industry, Too," *Business Insider*, May 12, 2013, accessed June 6, 2020, <https://www.businessinsider.com/the-film-industry-needs-big-data-2013-5>

**Table 2.5**

Legacy Media Measurement Framework		Digital Media Measurement Framework	
<i>Content</i>	<ul style="list-style-type: none"> <li>● Revenue</li> <li>● Number of viewers                             <ul style="list-style-type: none"> <li>○ Movie-goers</li> <li>○ TV viewers</li> </ul> </li> <li>● Number of tickets sold</li> </ul>	<i>Content</i>	<ul style="list-style-type: none"> <li>● Engagement                             <ul style="list-style-type: none"> <li>○ Video views</li> <li>○ Watch time</li> <li>○ Likes</li> <li>○ Shares</li> <li>○ Comments</li> </ul> </li> <li>● Reach</li> <li>● Impressions</li> </ul>
<i>Audience</i>	<ul style="list-style-type: none"> <li>● Feedback</li> <li>● Demographics                             <ul style="list-style-type: none"> <li>○ Above/under 25</li> <li>○ Male/female</li> </ul> </li> </ul>	<i>Audience</i>	<ul style="list-style-type: none"> <li>● Online activity                             <ul style="list-style-type: none"> <li>○ Viewing history</li> <li>○ Time spent online</li> </ul> </li> <li>● Sentiment                             <ul style="list-style-type: none"> <li>○ Opinion scores</li> <li>○ Positive/negative</li> </ul> </li> <li>● Affinity                             <ul style="list-style-type: none"> <li>○ Favorite genres</li> <li>○ Favorite actors</li> </ul> </li> <li>● Demographics                             <ul style="list-style-type: none"> <li>○ Age</li> <li>○ Gender</li> <li>○ Education</li> <li>○ Lifestyle</li> </ul> </li> </ul>

The new digital media environment delivered unprecedented layers of data on content and audiences, effectively challenging Hollywood’s de facto approach to making sense of information. As such, the exponential rise in digital information called for an evolution of analytical frameworks to navigate the new data dynamics of the industry. As part of Hollywood’s ongoing convergence with Silicon Valley, the legacy media industry had to adjust to the norms of the technology business. Technology companies, platforms in particular, developed, codified, and maintained an expansive set of digital metrics to manage their business and make sense of the global information inflow.

Over the past decade, platforms had experienced significant user growth, effectively accumulating massive data loads. While this access to data quantity has generally been

characterized as a competitive business advantage,<sup>268</sup> platforms enacted a business strategy that focused on the analytical distillation of quality, investing in systematic data analysis to translate information into valuable insights that deliver actionable intelligence for their business. As Netflix's former VP of Data Science noted, "[t]here's a mountain of data that we have at our disposal. That mountain is composed of two things. Garbage is 99 percent of that mountain. Gold is one percent."<sup>269</sup> Indeed, Netflix developed an analytical framework to cut through the overall data noise and manage the constant information overload, filtering its analytics practices through a wide-ranging, yet highly focused set of digital metrics, tailored to its content and audience. The company's framework particularly focused on measuring audience engagement with content, which Chief Content Officer Ted Sarandos highlighted as the company's key success metric: "The longer people watch Netflix and the longer they stay members - they're the criteria of success for us."<sup>270</sup> As such, Netflix identified a core set of digital metrics to enable focused and actionable data analysis.

Other platforms similarly focused on metrics most relevant to their business. For example, social media platforms like Facebook and Twitter, which derive the largest share of their revenue from advertising, emphasized user actions, such as likes, comments, and shares, as well as content reach and impressions, to measure their business, highlighting key advertising metrics in their analysis of digital information. Digital media distribution platforms like Apple's App Store and Google's Play Store, driven by downloads and direct mobile purchases, measured

---

<sup>268</sup> See, Fritz, Ben, *The Big Picture: The Fight for the Future of Movies* (Boston: Houghton Mifflin Harcourt, 2018).

<sup>269</sup> Kwek, Nick, "Most of big data is 'trash' says Netflix's Todd Yellin," *BBC News*, April 21, 2016, accessed June 6, 2020, <https://www.bbc.com/news/av/technology-36093007/most-of-big-data-is-trash-says-netflix-s-todd-yellin>

<sup>270</sup> Jeffries, Stuart, "Netflix's Ted Sarandos: the 'evil genius' behind a TV revolution," *The Guardian*, December 30, 2013, accessed June 5, 2020, <https://www.theguardian.com/media/2013/dec/30/netflix-evil-genius-tv-revolution-ted-sarandos>

mobile app and game downloads as well as revenue derived from in-app user transactions. All platforms meticulously tracked user growth to document the scale of their audience. While platforms have maintained access to a multitude of data points, the focus on specific digital business metrics enabled them to “separate the signal from the noise,”<sup>271</sup> translate broad information into actionable insights, and generate relevant intelligence for their business.

The list of digital metrics derived from platforms is expansive, reflective of the digital economy’s complex and intricate infrastructure.<sup>272</sup> While all platforms share a set of unifying metrics across the digital landscape (e.g., engagement, video views, audience demographics), there are discrete differences by platform business model, feature design, and company strategy. Furthermore, even though the digital economy is grounded in metrics standards, as devised by industry organizations such as the Interactive Advertising Bureau (IAB), platforms have generally maintained their own definitions for metrics, leading to a layered nomenclature<sup>273</sup> (table 2.6). The following provides a high-level, non-exhaustive overview of digital metrics by platform type, including social media, subscription video-on-demand (SVOD), advertising video-on-demand (AVOD), and mobile apps and games.

---

<sup>271</sup> Andrea Mustain & Mike Osborne, “Gold or Pyrite,” *Raw Data Podcast*, November 29, 2016, accessed June 6, 2020, <https://podcasts.apple.com/us/podcast/gold-or-pyrite/id1042137974?i=1000378399788>

<sup>272</sup> See, Gray, Jonathan, “Reviving audience studies,” *Critical Studies in Media Communication*, Volume 34, Issue 1, 2017.

<sup>273</sup> For an overview of metrics standardization, see, “IAB Measurement Guidelines,” *Internet Advertising Bureau*, accessed October 5, 2019, <https://www.iab.com/guidelines/iab-measurement-guidelines/>

**Table 2.6**

Platforms and Digital Metrics			
Social Media	Subscription Video-on-Demand	Advertising Video-on-Demand	Mobile Apps and Games
<ul style="list-style-type: none"> <li>Engagement (e.g., number of likes, shares)</li> <li>Reach (e.g., # of users who see content)</li> <li>Impressions (e.g., # of users who interact content)</li> </ul>	<ul style="list-style-type: none"> <li>Watch time (e.g., time spent watching video)</li> <li>Membership (e.g., time since first subscribed)</li> </ul>	<ul style="list-style-type: none"> <li>Watch time (e.g., time spent watching video)</li> <li>Reach (e.g., # of users who see content)</li> <li>Impressions (e.g., # of users who interact content)</li> </ul>	<ul style="list-style-type: none"> <li>Downloads (e.g., # of app and game downloads)</li> <li>Revenue per user (e.g., average spend per user)</li> <li>Gaming time (e.g., # of monthly active users)</li> </ul>

The digital economy is characterized by a plurality of metrics that consistently track, collect, and quantify the interaction between content and consumers. Content performance is no longer based merely on financials and number of viewers (i.e., box office and ratings), but video views, the time spent watching video, and interactive actions (i.e., likes, shares, comments). Audience reception is no longer based on static demographics and directed feedback questions, but a fluid spectrum of categories capturing personal interests, unfiltered reactions, and deeper backgrounds. In this sense, platforms effectively expanded beyond legacy metrics to more layered concepts of measurement, designed to capture the complex and layered digital media environment of the 21st century (table 2.7).

**Table 2.7**

Legacy Media Metrics	Digital Media Metrics
Box Office and Ratings	Engagement
Audience Feedback	Affinity

Engagement has become the key digital metric to characterize content performance. It is a broad metric that subsumes a range of digital actions and behaviors, all of which are focused on quantifying consumer interaction with content. While the performance of film releases was measured in box office grosses, the release of YouTube videos is measured in video views, watch time, reach, impressions, and actions taken (e.g., likes, shares, comments, clicks). Affinity has become the key digital metric to measure audience reception. While audience response to a new network TV show was measured in static demos (i.e., audiences aged over or under 25)<sup>274</sup> and weekly ratings, the release of a new Netflix show is measured by affinity clusters and behavioral data.

As platforms cemented their positioning in the new media environment, marked by an ongoing increase in information, driven by alternative modes of content consumption and shifting consumer viewing habits, the majors effectively had to adapt and negotiate a new set of digital metrics to manage the implications for the legacy media business. The emergence of digital metrics did not constitute a departure from Hollywood's legacy framework. Rather, it presented a reconfiguration of the industry's measurement approach, effectively and expanding the entrenched concepts of content performance and audience reception.

- *Content Performance*: The majors moved beyond box office and ratings, developing an analytical focus through *engagement*.
- *Audience Reception*: The majors expanded beyond audience feedback and demographics with an analytical focus on *deeper consumer segments and affinity*.

---

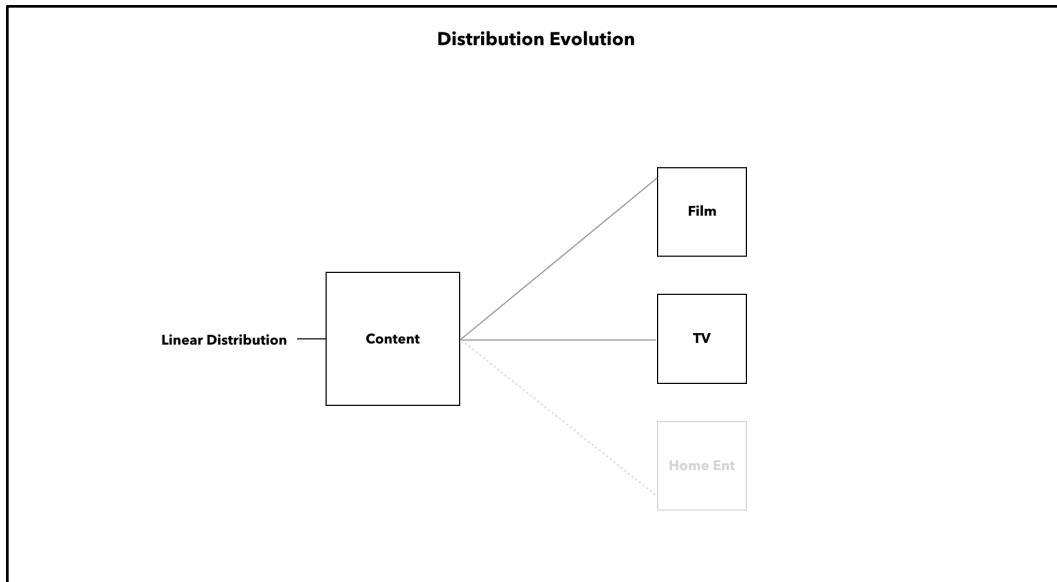
<sup>274</sup> See, Epstein, 2012

By integrating digital metrics into the existing legacy framework, the majors effectively adapted the legacy media industry's measurement approach for the digital economy, incubating analytical frameworks that provided additional insight into the interplay of content and audiences in the shifting digital media environment.

### Content Performance

Hollywood's focus on box office and ratings was grounded in the legacy media industry's linear distribution model. Studios distributed films through theaters and networks broadcast shows through television sets (both further released copies through the ancillary home entertainment market though those revenues were tabulated later). In this model, the metrics of box office and ratings were a direct reflection of audience interaction with content, following a consistent link between the majors and their limited set of distribution channels (figure 2.7).

Box office effectively captured a film's performance, serving as an indicator of financial revenue (i.e., number of tickets sold) and audience scale (i.e., the number of viewers based on ticket sales), and a comparative benchmark for other releases. Ratings similarly illustrated performance, offering a built-in methodology to indicate a show's advertising performance (i.e., the value of ads sold during the show's commercials), audience size (i.e., viewers tuned in), and overall market share (i.e., number of viewers compared to other shows).



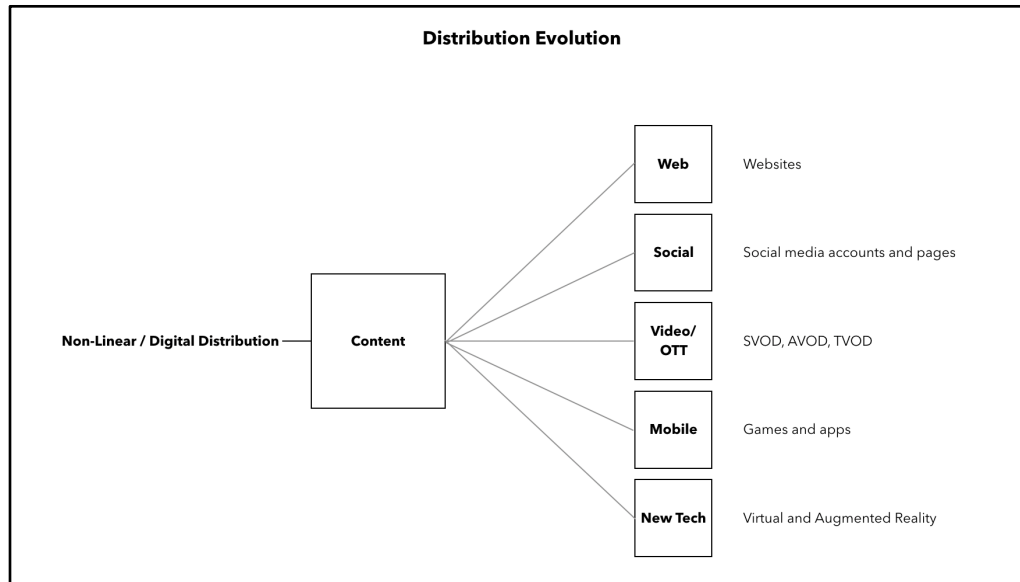
**Figure 2.7:** In linear distribution, the majors distributed programming across three central channels, all disintermediated from the major film studios and television networks.

The digital era introduced a paradigm shift for the linear distribution model with a proliferation of channels and platforms that led to a surplus of content, fragmented content viewing, and various audience clusters (figure 2.8). The resulting distribution environment yielded a flood of content and audience information, differentiated by channel and platform. Effectively, the legacy media industry was no longer able to track content performance based on theatrical and network distribution as its content operated across various digital outlets.

Each digital channel and platform delivered unique forms of data. For example, websites captured the time consumers spend in browsers, the time they spend on specific websites, and their journey across the Internet; social media platforms tracked a film or television show's number of followers, their overall likes and shares, as well as the volume of comments associated with them; and mobile apps and games measured downloads. The focus on engagement enabled a central, unifying metric, consolidating key layers of information across the digital landscape. For example, websites, social media platforms, and mobile apps all track



likes, video views, and time spent. As such, the focus on engagement allowed the majors to gain a holistic perspective on digital content performance.



**Figure 2.8:** Digital distribution expanded the number of distribution channels, creating a network of data.

Hollywood’s growing prioritization of engagement as a central metric to make sense of digital information did not minimize the role of box office and ratings in the industry. Indeed, the legacy media industry has consistently measured content performance through the lens of legacy metrics. At the same time, the majors have increasingly added a focus on measuring engagement to understand the business dynamics of film and television show releases. As such, the concept of digital engagement does not represent a substitute, but a complement to Hollywood’s legacy metrics framework.

### Audience Reception

Hollywood’s legacy measurement framework captured audience reception in two principal ways.

First, the majors defined the composition of the audience through a static segmentation model.

The audience was categorized based on a four-quadrant pattern, which classified viewers by age

(i.e., over/under 25) and gender (i.e., male/female). Second, the majors tracked audience opinion, using directed survey questions to document feedback on various elements of films and television shows.

The information influx brought about by the digital era introduced new layers of data on the media audience that fundamentally expanded the legacy media industry's existing approach. Indeed, by capturing, consolidating, and conforming additional information on viewers, the majors created a more comprehensive and holistic representation of the media audience.

### *Consumer Segments*

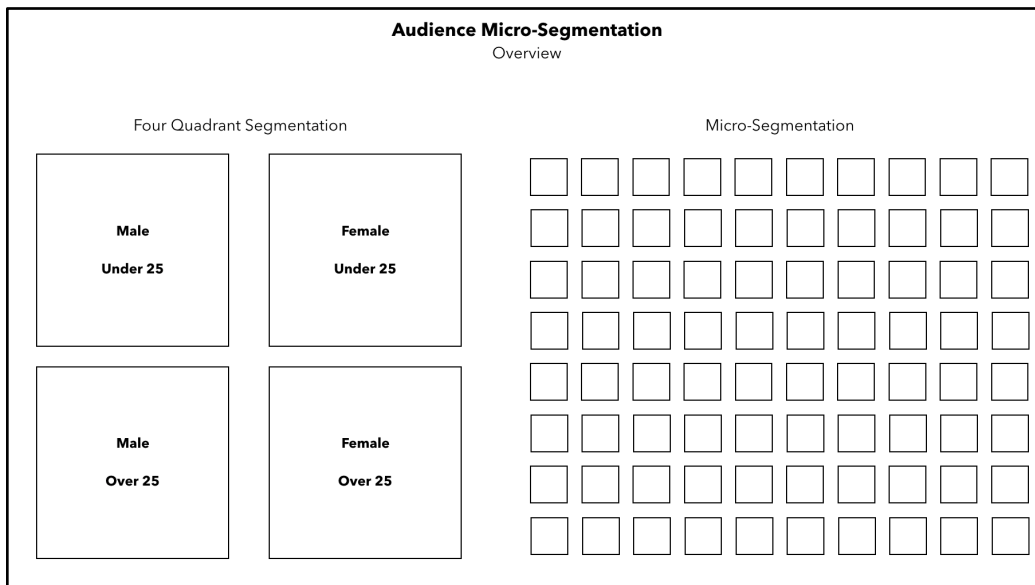
The majors have long captured audience demographics with the help of polls, surveys, and focus groups, using a basic segmentation model to classify viewers into 4 quadrants, grouped by age and gender.<sup>275</sup> The digital landscape introduced a wide range of information categories that added unprecedented layers of data volume and depth on the audience. As a result, the industrial concept of demographics was expanded beyond age and gender, incorporating a variety of incremental data points, such as age, location, and language, as well as personalized classifiers, including interests, hobbies, and likes, and technological descriptors (i.e., technological devices and operating systems used to consume content). To provide an example, Facebook's demographics category lists information on age and gender, lifestyle, education, relationship status, job role and household size.<sup>276</sup>

---

<sup>275</sup> The 4-quadrant audience model emerged as a staple of mainstream Hollywood marketing during the emerging blockbuster period of the 1970s and 1980s. See, Wasko, 2003.

<sup>276</sup> See Facebook's Business Insights page for up-to-date information on metrics: <https://www.facebook.com/business/insights>

The concept of demographics was effectively reconceptualized, from a small staple of static classification categories to a more expansive range of emotional rubrics. As such, the industry moved from a set of segment quadrants to a suite of micro-segments, opening up a broader canvas of metrics to measure, analyze, and conceptualize the audience (figure 2.9).



**Figure 2.9:** Digital dynamics reconceptualized the media audience by adding multiple layers of data.

The shifting focus on micro-segmentation marked a departure from Hollywood’s long-standing 4-quadrant model. Digital micro-segments<sup>277</sup> of the audience increasingly highlighted a consideration of shared behaviors and interests over age and gender, putting emphasis on how audiences feel about content rather than extrapolating preferences based on static categories.<sup>278</sup> As the former head of Amazon Studios put it succinctly: "Demographics are a way to speculate

<sup>277</sup> Digital micro-segments focus on viewing behaviors and interests (e.g., genre, cast, music). See, Napoli, 2010

<sup>278</sup> For example, Todd Yellin, Netflix’s VP of Product, noted: “Geography, age, and gender? We put that in the garbage heap.” See, Morris, David Z., “Netflix says Geography, Age, and Gender are “Garbage” for Predicting Taste,” *Yahoo Finance*, March 27, 2016, accessed June 6, 2020, <https://finance.yahoo.com/news/netflix-says-geography-age-gender-192801834.html>

about what people *might* like to watch, but if you actually know what they like to watch that's more important than the demographics."<sup>279</sup>

### *Consumer Affinity*

The majors have tracked audience feedback on films and television shows by way of focus groups, polls, and surveys, capturing viewers' thoughts, opinions, and reactions in structured and directed patterns. In particular, they gathered feedback on key elements of films and television shows, from the story/plot and characters to the acting and filmmaking display. In the digital era, this manual work has been complemented (at times taken over) by digital technologies, specifically data analysis software applications that collect the opinions and reactions of audiences across a variety of digital feedback mechanisms, consolidate the information into actionable insights, and quantify them in form of a percentage-based reception metric.

Accordingly, the legacy media industry has expanded beyond survey reports and written feedback cards to encompass digital manifestations of the audience's content affinity, tracking the unfiltered and undirected response to a film or television show as part of a larger media discourse driven by the audience. Aggregators like *Rotten Tomatoes*, *IMDb*, and *Fandango* serve as a first layer of this reception, offering stylized forms of digitally-captured and -rendered audience feedback, and detailing the underlying information that informs the overarching metric, such as the number of viewer reactions evaluated.

Additionally, analytics companies track audience opinion on films and television shows across the digital landscape, distilling massive amounts of posts, comments, and reactions into

---

<sup>279</sup> Bishop, Bryan, "From alpha to 'Betas': how Amazon is rethinking the way television is made," *The Verge*, November 26, 2013, accessed June 6, 2020, <https://www.theverge.com/2013/11/26/5147796/betas-how-amazon-is-rethinking-the-way-television-is-made>

digital audience metrics. For example, social media analytics company Crimson Hexagon captures the volume of conversation on a film or television show across social media, offering the number of posts during a specified time period. Moreover, the company enables advanced filtering and classification, effectively slicing and dicing the information to reveal key themes, recurring patterns, and sentiment (i.e., a positive, neutral, or negative) within the conversation.

While focus groups, polls, and surveys emphasized front-end information-gathering, directly soliciting audience feedback in a controlled environment, data analysis software gathers data on the backend, indirectly tracing, tabulating, and translating the opinions of audiences across the entire digital landscape. While digital technologies are neither exhaustive nor infallible, the majors' growing implementation of automated software enabled increased functionality to measure audience reception in form of affinity, revealing hitherto uncaptured information on film and television viewers.

--- --- ---

The rise of digital metrics had significant implications for the legacy media industry. On the one hand, the integration of digital metrics enabled Hollywood to adapt its existing measurement framework for the new media environment, make sense of the growing amounts of data on content and audiences, and deploy an evolving arsenal of insights to inform decision-making through data analysis and business intelligence. Indeed, digital metrics provided a ready-made solution to understand the interaction between content and audiences. Video view totals and percentage-based opinion scores delivered instant insight into content performance and audience reception. Furthermore, the simplicity and accessibility of digital metrics made them highly functional as a determining factor in corporate decision-making. Accordingly, the industry increasingly highlighted digital metrics to frame and rationalize success in the business.

Studios and networks cited trailer views, volume of social media conversations, and likes in evaluating the business,<sup>280</sup> promoting new releases (e.g., the use of *Rotten Tomatoes* scores in marketing materials such as trailers and posters), and negotiating future projects (e.g., choosing IP or cast based on social media followings).<sup>281</sup> Over time, digital metrics emerged as an incremental industry benchmark, an additive industrial practice for Hollywood to not only analyze, but increasingly commercialize, the business.

On the other hand, digital metrics did not constitute a fail-safe solution for the industry, introducing a variety of complications into the evolving measurement framework, for several reasons. One, while box office and ratings were largely simple metrics that lacked context and layers, they were highly transparent, built on an industry-wide set of reporting methods. Digital metrics, by contrast, are derived from platforms which only share select information externally while retaining more substantial data on the backend, effectively operating as “analytical black holes.”<sup>282</sup> Two, platforms have regularly changed their terms of service, seemingly adjusting definitions and values of metrics arbitrarily, which has a notable impact on perceptions of accuracy. Platforms often use varying definitions for metrics, as evidenced in a wide spectrum of rules in what constitutes an online video view (e.g., Facebook’s definition of a view emerged as three seconds long, YouTube opted for ten seconds, and Netflix reportedly saw it as 2

---

<sup>280</sup> See, Kirkpatrick, Scott, *Introduction to Media Distribution: Film, Television, and New Media* (New York: Routledge, 2018). For additional insight into digital metrics and legacy media economics, see, Buzzard 2012 and Napoli, 2010.

<sup>281</sup> See, Pullen, John Patrick, “In Hollywood, social media takes a leading role,” *Fortune*, August 1, 2014, accessed June 6, 2020, <https://fortune.com/2014/08/01/in-hollywood-social-media-takes-a-leading-role/>

<sup>282</sup> Cieply, Michael, “Hollywood Wants Numbers on the Digital Box Office,” *New York Times*, September 15, 2013, accessed June 6, 2020, <https://www.nytimes.com/2013/09/16/business/media/movie-industry-wants-to-get-a-handle-on-the-digital-box-office.html>

minutes),<sup>283</sup> thereby causing confusion around standards. As one agency executive noted, “[t]his space is equivalent to a landfill in an earthquake. All the patterns go haywire.”<sup>284</sup> And, three, the use of digital metrics is not tied to a common structure of standards and practices within the legacy media industry, which complicates long-term institutionalization and undermines efficiencies of the new measurement framework. While platforms have made attempts to provide holistic metrics solutions, they were largely short-lived and too platform-specific to take hold (e.g., Twitter and Nielsen’s collaboration on Social TV Ratings). Notably, a senior digital marketing executive at a major studio noted, “[w]e have to standardize our approach to data as an industry. It’s a huge issue that needs to be addressed.”<sup>285</sup>

Overall, the integration of digital metrics enabled the majors to make sense of digital information and navigate the information influx of the new media environment. At the same time, digital metrics were themselves marked by surplus and complexity, expanding Hollywood’s measurement framework from a handful to a plethora of metrics, creating an industry-wide need for analytical synthesis.

### **Data as Storytelling: The Aesthetics of Data Visualization**

In 2013, big data went public, with the initial public offering (IPO) of Tableau, an analytics software company that specializes in data analysis and visualization. Trading as “DATA,” Tableau priced its IPO on the high end, \$31/share, with initial discussions centering around the

---

<sup>283</sup> John Herrman & Mike Isaac, “The Online Video View: We Can Count It, but Can We Count on It?,” *New York Times*, October 2, 2016, accessed June 6, 2020, <https://www.nytimes.com/2016/10/03/business/media/the-online-video-view-we-can-count-it-but-can-we-count-on-it.html>

<sup>284</sup> Cieply, 2013

<sup>285</sup> Littleton, 2015, “10 Things We Learned at Variety’s Big Data Summit”

lower price mark of \$23-\$26, given Wall Street's hesitation about the technology industry's trajectory in the market, which still reeled from Facebook's lackluster IPO the year before. When Tableau began trading on the New York Stock Exchange (NYSE), the market elevated its price to \$47/share, a 51% increase, with the stock closing at \$50.75/share, registering a 65% growth rate within its first day.<sup>286</sup> The company's tens of thousands of customers across global regions, coupled with its unique data visualization technology, resonated with a large set of the investor base, who ascribed growing value to data analysis in a world increasingly driven by the collection, consolidation, and communication of insights. In 2019, six years after its trading debut, Tableau stock had reached \$125/share. That same year, Tableau was acquired in an all-stock transaction by Salesforce, one of the world's largest enterprise software companies, which valued Tableau at \$15.7 billion. Salesforce positioned the acquisition as a value add for its enterprise customers, noting that "[w]ith Tableau, Salesforce will play an even greater role in driving digital transformation, enabling companies around the world to tap into data across their entire business and surface deeper insights to make smarter decisions [...]." Adam Selipsky, President and CEO of Tableau, equally emphasized the value the company's data analysis would deliver to companies, saying that "[j]oining forces with Salesforce will enhance our ability to help people everywhere see and understand data [...]."<sup>287</sup> Effectively, Salesforce's comprehensive suite of marketing software, global customer base, and limitless IT infrastructure made Tableau the data analysis layer of the digital economy.<sup>288</sup>

---

<sup>286</sup> See, Lunden, Ingrid, "Big Data Analytics Specialist Tableau Software Raises \$254M In IPO," *TechCrunch*, May 17, 2013, accessed June 6, 2020, <https://techcrunch.com/2013/05/17/big-data-visualization-goes-public-tableau-software-raises-254m-as-shares-pop-58-while-marketo-raises-85m/>

<sup>287</sup> Osborne, Charlie, "Salesforce acquires Tableau Software in \$15.7 billion deal," *ZDNet*, June 10, 2019, accessed June 6, 2020, <http://zdnet.com/article/salesforce-acquires-tableau-software-in-15-7-billion-deal/>



While Tableau is but one of multiple data analysis companies that have emerged over the 21st century, its success story reflects key industrial developments that informed Hollywood's integration of data as a decision-making variable through a visual system of reference.

One, Tableau's rise to prominence reflects a growing demand for data analysis software. Indeed, the market for data analysis solutions has grown exponentially with the ongoing influx of digital information, driven by a confluence of factors, including the commercialization of the Internet, the proliferation of platforms, and a steady climb of online activity in the personal and professional spheres. Furthermore, as cloud computing increased data storage capacities while consistently lowering maintenance costs, organizations have invested in collecting, capturing, and conforming data for analysis. Companies specializing in data analysis have thrived under these industrial conditions, leading to an initially highly fragmented marketplace serving multiple industry verticals, that eventually underwent consolidation with larger technology companies adding analytics solutions to their digital ecosystems (e.g., Salesforce's 2019 acquisition of Tableau, Google's 2019 purchase of Looker). Tableau led the wave of data analysis software, building a massive customer base across various industry sectors, including media and entertainment. Indeed, Tableau was widely adopted by studios and networks as an analytics solution, alongside various other products, reflecting Hollywood's growing investment in data analysis software.

Two, as demand for data analysis software grew, companies across industries experienced a need for ready-made, easy-to-use, and highly functional solutions. Tableau's rise illustrated a growing industrial trend toward democratizing data analysis by way of a software-as-a-service (SaaS) offering, effectively enabling organizations and their workforce to manage

---

<sup>288</sup> Salesforce, "Salesforce Completes Acquisition of Tableau," *Salesforce*, August 1, 2019, accessed June 6, 2020, <https://www.tableau.com/about/press-releases/2019/salesforce-completes-acquisition-tableau>

increasing amounts of digital information without any existing capabilities by licensing virtual access to analytics solutions in exchange for a monthly subscription fee. Indeed, Tableau’s corporate mission set a mandate “to help people see and understand data, [...] put the power of data into the hands of everyday people, allowing a broad population of business users to engage with their data, ask questions, solve problems and create value.” This utilitarian approach was seemingly designed to demystify the technical complexities of data analysis and instead emphasize the functional ease and high-value results analysis software can deliver to organizations. As such, data analytics companies effectively sought to institutionalize the automation of data analysis for the digital economy, reaching a wide range of companies that lack analytical capabilities. Tableau’s corporate manifesto expressed this strategic approach succinctly (my emphasis):

“Our software is **designed for anyone** with data and questions. We are **democratizing the use of business analytics software** by allowing people to access information, perform analysis and share results without assistance from technical specialists. By putting powerful, self-service analytical technology directly into the hands of people who make decisions with data, we **seek to accelerate the pace of informed and intelligent decision making** [...].”<sup>289</sup>

Tableau, along with its brethren of data analytics companies, emphasized the role of data analysis software in translating information into insights in an effort to support corporate decision-making. Notably, Tableau did not position its software as the exclusive analytical engine, but rather as a complement to organizations’ operating formulas, which made it less disruptive to established corporate processes.

---

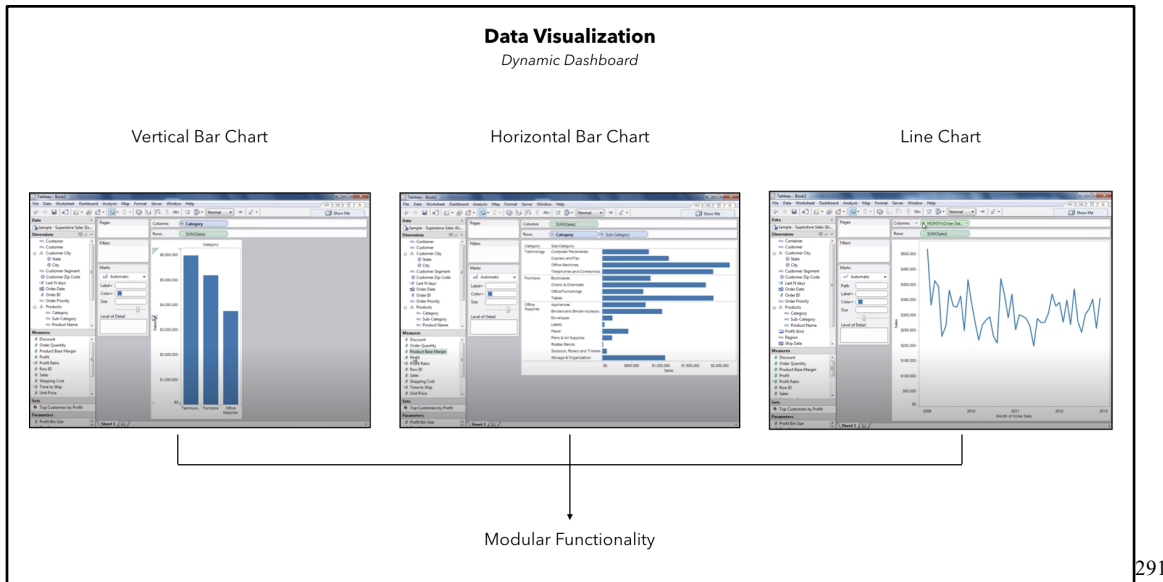
<sup>289</sup> Tableau, “Form S-1 Registration Statement, *SEC*, April 2, 2013, accessed June 6, 2020, <https://www.sec.gov/Archives/edgar/data/1303652/000119312513138700/d469057ds1.htm>

And, three, data analysis software invested in highlighting visual design over technical processes in light of increasing demand from legacy organizations not versed in data. The concept of data analysis by default evokes a suite of technical associations that raise resistance and speculation among non-technical legacy organizations, from highly advanced technical skills (i.e., knowledge of statistics and economics) to computational language requirements (i.e., C, Java, Python), and custom mathematical communications (i.e., Excel spreadsheets, programming-based database systems). Analytics companies accordingly developed a utilitarian SaaS-based approach to minimize the inherent technical complexity and instead focus on highly usable and functional output features. As such, they designed software that shifted focus from the intricacies of the analytical process to the results of the analytical output, enabling organizations to build insights and generate intelligence on their business.

Tableau's software design is particularly instructive in this context. The company's core data analysis product is a custom plug-and-play dashboard that allows users to visualize data using a variety of visual templates. Essentially, the dashboard is a modular digital interface made up of multiple windows where information can be dragged and dropped to form a graphical representation of data<sup>290</sup> (figure 2.10).

---

<sup>290</sup> See, Galloway, Alexander, *The Interface Effect* (Cambridge, UK: Polity, 2012).



**Figure 2.10:** The *Tableau* interface includes various chart features that display data in differentiated views.

In addition to the interactive design feature, Tableau equally offers a suite of templates that automate data visualization, an analytical practice utilized by the majority of analytics companies as a way to minimize the border of entry for non-technical organizations and, in return, maximize software adoption.

Tableau’s focus on data visualization over standard computational database representation is grounded in the idea that data analysis constitutes a form of visual storytelling. The focus on the visual dimension of data likely stemmed from the company’s founding team, which consisted of three Stanford computer science researchers, Christian Chabot, Chris Stolte, and Pat Hanrahan. Hanrahan, notably, was a founding employee of Pixar Animation Studios, who received three Academy Awards for the company’s early animation work, including its debut feature *Toy Story* (1995). On Tableau’s approach to visualization, he noted, “[w]e create

<sup>291</sup> See the Tableau website and YouTube page, accessed June 4, 2020.

pictures that answer questions, but we do it for businesses that want to know things about their own metrics. It has been termed visual analysis—sort of doing a Q&A with data and images.”<sup>292</sup>

Tableau’s dashboard effectively operates as a mechanism for conducting visual analysis. It draws on visualization techniques to discover and communicate data insights via an easy-to-use and intuitive interface. Specifically, Tableau utilizes concepts from neuroscience and psychology to engineer software features that align with patterns of visual understanding, enabling users to make sense of digital information by identifying and highlighting distinct visual patterns, as described in the company’s software product overview (my emphasis):

**“Discovery—We believe that the human mind is better able to process information, discern trends and identify patterns when presented with information in a visual format.** By fundamentally integrating data analysis and visualization, our software allows people to create powerful visualizations and dashboards that can lead to new discoveries. Our software is designed to seamlessly blend, filter and drill down on information [...].”<sup>293</sup>

Effectively, Tableau positioned data analysis not as a technical effort, but a creative exercise in storytelling, generating insights through visual templates and patterns. The company deliberately designed a visual focus to help organizations and their workforce communicate with data, tapping into established patterns of information recognition that transcend technological practices (my emphasis).

**“Communication and sharing—We facilitate more effective communication by empowering people to express themselves creatively and tell better stories with data.** The collaborative features of our software are **designed to foster more sharing of data and to improve the dissemination of information across and among enterprises.** [...]

---

<sup>292</sup> Martinet, Drake, “Almost Famous: Pat Hanrahan of Tableau,” *All Things Digital*, February 26, 2010, accessed June 6, 2020, <http://allthingsd.com/20100226/almost-famous-pat-hanrahan-of-tableau/>

<sup>293</sup> Tableau Mission Statement, Tableau website, accessed June 6, 2020, <https://www.tableau.com/about/mission>

We believe that our software enables our customers to **share more insights** and have richer conversations about their information.”<sup>294</sup>

Tableau’s focus on visualization represents data in a way “which allows [us] to arrive at insights and ‘knowledge’.”<sup>295</sup> While the company’s plug-and-play software design is unique and differentiated, its focus on visual aesthetics has become a holistic feature of the data analysis industry, with analytics companies programming dashboards with charts, graphs, maps, and other visual signifiers to communicate data insights. The data dashboard has become a dominant interface within the information age, acting as an analytical engine and communications tool that drives corporate decision-making through the visual representation of data. The dashboard is at once industrial product, aesthetic expression, and insights medium, de facto operating as a modern data-eye within the digital economy.<sup>296</sup> Borne out of a growing demand for ready-made data analysis, the dashboard emerged as the quintessential analytics practice, collecting, conforming, and communicating data as insight.

Tableau and its ilk of analytics players have effectively incubated, implemented, and institutionalized an automated system to communicate data insights within organizations that lack data analysis capabilities, responding to an increasing demand for data analysis software with a utilitarian software-as-a-service approach that emphasizes information visualization. The dashboard, in particular, operated as the crux of data analysis software, distilling computational,

---

<sup>294</sup> Ibid

<sup>295</sup> Ibid

<sup>296</sup> See, Mamber, Stephen, “Space-Time Mappings as Database Browsing Tools.” In: *Media Computing: Computational Media Aesthetics*, edited by Chitra Dorai and Svetha Venkatesh (Berlin: Springer, 2002) and Manovich, 2001.

statistical, and aesthetic principles into a visual process of information-gathering and insights generation. As such, the dashboard extended its impact across various industries.

The legacy media industry, through its systematic integration of data analysis technology, professionals, and metrics, adopted the data visualization dashboard as a central element of its developing analytics infrastructure. The majors effectively integrated the dashboard as an industrial product, aesthetic interface, and organizational communications tool, a triptych model that relied on visual mechanics to centralize and institutionalize Hollywood's approach to data analysis and analytical decision-making.

- *The Dashboard Industry*: The majors partnered with a wide range of analytics companies, effectively implementing a variety of dashboards across their organizational cultures for various contexts.
- *Dashboard Aesthetics*: The majors implemented dashboards as an extension of digital metrics, leveraging aesthetic principles as a complement to the numerical and textual representations of data in an effort to make sense of digital information.
- *Dashboard Storytelling*: The majors established dashboards as an internal layer to communicate and share data across their organizational structures, establishing virtual access to data insights, and promote growing degrees of data literacy.

Collectively, these industrial developments highlight Hollywood's integration of the dashboard as an interoperable element of its business operations and corporate decision-making.

## The Dashboard Industry

The legacy media industry developed a suite of partnerships with SaaS-based analytics companies that enhanced their data analysis capabilities and established the dashboard as a centralized industrial interface for information-gathering and insights generation. As part of the evolving partnerships network, the majors licensed access to various data analysis dashboards, gradually building an infrastructure of visual data displays across their organizational structures. The dashboards enable studios and networks to visualize data insights across organizational structures, yet differ by feature set (e.g., automated vs. customizable dashboards), data type (e.g., holistic data analysis vs. analysis of specific data like web, social, mobile, etc.), and business model (e.g., freemium vs. paid).

While the majors cultivated a variety of partnerships with analytics companies, following an ongoing test-and-learn approach to experiment with various solutions to expand their technology infrastructure while managing costs, dashboards generally derived from a key set of software categories:

- *Platform Applications*: Major digital platforms offer access to virtual dashboards as part of their larger marketing solutions. Effectively, platforms use dashboards to communicate marketing campaign insights to their customers, primarily showcasing content performance and audience insights. For example, Facebook uses its Insights dashboard to visualize insights for its customer marketing campaigns. Given the global scale of platforms, their dashboard infrastructure is highly transparent and accessible. They effectively represent automated templates that consistently refresh with new campaign data, yet limit customization for users. Customers can choose to focus on specific data views, highlighting specific metrics while minimizing others. However, they are unable



to design their own dashboard logic, with their access limited to a static output template. Furthermore, while platform dashboards are based on a freemium model, enabling customers to access insights for free, access is dependent on marketing spend or other types of investments, effectively requiring customers to run a paid marketing or sales campaign on the platform. Finally, platform dashboards operate as point solutions, only focusing on platform data, rather than holistic offerings that provide access to multiple data types. Notable platform dashboard applications can be grouped by category and include web platforms (e.g., Google Analytics), social media platforms (e.g., Facebook Insights, Instagram Insights, and Snapchat Insights), video platforms (e.g., YouTube Analytics), and mobile platforms (e.g., App Store Analytics, Google Play Console).

- *Startup Applications*: Startups design dashboards that operate as point solutions focused on specific data types, similar to platform applications, yet provide more comprehensive feature sets at a price point. The landscape of startup dashboard applications is expansive and constantly evolving as new entrants consistently seek to capitalize on the data boom. Startup dashboards generally follow a subscription model, which makes pricing a relative commodity, yet they differentiate through feature sets and additive data types. As such, they offer a wider range of automated templates, incremental options for customization, and consulting services that involve designing new dashboards for specific business needs. Furthermore, they deliver dashboards that integrate multiple data types, showcasing content and audience insights across multiple platforms, from web to social media, video, mobile apps, and emerging tech. Notable companies include Parrot Analytics, Listen First, and Relish Mix.

- *Customizable Applications*: Customizable dashboards display the most comprehensive feature sets, data type access, and pricing variability. They enable users to design, build, and update their own dashboards, using a wide range of visualization mechanics, provide infrastructure to integrate with all available data types across the digital landscape, and offer flexible pricing models by company size (e.g., educational institutions vs. small and medium businesses vs. enterprise companies). Given their offering, customizable applications also require higher degrees of technical understanding. While platform and startup applications are built around fully automated templates that self-structure data analysis, customizable applications require analytical input in building out analytical dashboard logic, as in deciding what visualization method to use for a particular source of data. Notable companies include Tableau (owned by Salesforce), Looker (owned by Google), and Alteryx.

The majors developed an infrastructure that leverages a spectrum of dashboard applications across their organizational cultures. The implementation of dashboards is neither focused nor arbitrary, but dependent on a variety of factors, such as internal budget allocation, preferences of key decision-makers (e.g., senior team leaders, head of IT, head of data science), and existing relationships. Given the complexity of legacy media organizations, which consist of multiple business segments with varying data needs, dashboards were generally used across teams and departments, rather than dedicated business units. While the differences in feature sets and data types created ongoing layers of fragmentation in data analysis, the dashboard still functioned as the unifying analytical layer across film studios and television networks, a reference point for information-gathering and insights generation, defined by a suite of familiar aesthetic conventions of visual presentation.

## Dashboard Aesthetics

The majors implemented various versions of the data dashboard across their organizational cultures, establishing it as the key reference for data insights. Indeed, the dashboard effectively acted as the analytical bulletin board, providing automated, automatically refreshing, and visually organized displays of insights to the workforce of film studios and television networks.<sup>297</sup> As such, the dashboard fulfilled the function of what has been characterized as a *digital nervous system*.<sup>298</sup> Bill Gates, the co-founder of Microsoft, offered an instructive definition of this phenomena in 1998 (my emphasis):

“The term 'digital nervous system' is kind of an interesting one. The analogy, of course, is to the biological nervous system where you always have the information you need. You always are alert to the most important things, and you block out the information that's not important. And **companies really need to have that same kind of thing: the information that's valuable getting to the people who need to know about it.**”<sup>299</sup>

In this sense, the dashboard can be conceptualized as a business insights delivery system designed to support corporate decision-making. While the biological metaphor of the nervous system is helpful in understanding the dashboard's high-level value within the corporate media environment, its functionality remains grounded in the visual representation of information. In particular, the dashboard operates under a set of conventions, derived from the disciplines of computer science, visual media, and quantitative science, all of which worked in unison to deliver business insights to the majors.

---

<sup>297</sup> From a technical perspective, the dashboard provides near-universal access. From a business perspective, access to dashboards is sold as part of a subscription model (with pricing based on individual users).

<sup>298</sup> Gates, Bill, “Bill Gates' New Rules,” *Time*, April 19, 1999, accessed June 6, 2020, <http://content.time.com/time/world/article/0,8599,2053895,00.html>

<sup>299</sup> Ibid

The dashboard is first and foremost a computational product. In its elemental form, the dashboard constitutes a graphical user interface (GUI) that incorporates a multitude of visual layers embedded within its horizontal and vertical depths, representing various data types and sources. For example, it displays data on content performance and audience reception, data from websites and social media platforms, as well as historical and real-time data. As such, the interface effectively operates as a composite that synthesizes multiple information elements into a cohesive unit. Furthermore, the compositional nature of the interface is interactive and customizable, enabling users to manipulate the representation of data and thereby re-contextualize data in new ways, using a variety of filters, highlighters, and conditions.

The interface is by nature a virtual space, engineered through digital media software, highly interactive and customizable. Yet, the way the interface organizes data is inherently governed by the visual dynamics of legacy media, particularly the aesthetic strategies of cinema. According to Lev Manovich, cinematic forms of representation “have become basic organizational principles of computer software,” with the dashboard quite literally providing “a window into a datascape.”<sup>300</sup> Indeed, the dashboard is at once interface and visual frame, or more accurately screen, drawing on established camerawork and editing techniques to represent data. It embraces spatial continuity in favor of temporal montage to highlight the consistency of information on-screen; it utilizes a mobile camera to move seamlessly through data space; it zooms in and out to highlight and recast data; it tilts up and down and pans from side to side in order to reveal more information; it incorporates off-screen space to reveal new information in form of menus and additive features; and it plays with multiple windows within the frame, creating a visual dialogue between data sets.

---

<sup>300</sup> Manovich, 2013, 86

The visual elements of the data dashboard further draw on mathematical and statistical forms of representation, expanding the cinematic frame with quantitative levels of measurement. In particular, the dashboard leverages a wide range of graphs, charts, tables, graphics, and maps to put data into context, distilling massive amounts of information, rows of numeric and text data, historical and current data, into visual reference points. Collectively, dashboard aesthetics do not only visualize data, they reference, reframe, and contextualize information to make data visually and intellectually accessible, moving from abstracted concept to relatable reference point; from obscure number to visual pattern; from isolated information to holistic insight. Effectively, the dashboard works to reference data by turning numeric and textual metrics into visual reference; reframe data by aggregating multiple data sets in a cohesive visual frame; and contextualize data by revealing visual connections and disparities.

The dashboard's interdisciplinary approach to data visualization made it more than a repository of insights for the majors. It emerged as a software-based tool to enable, promote, and optimize the communication of data across legacy media organizations, thereby elevating the majors' awareness, understanding, and deployment of data as a decision-making factor.

### Dashboard Storytelling

The growing use of the dashboard across the organizational cultures of legacy media companies enabled the majors to build up data literacy by visualizing data insights as relatable and tangible reference, effectively creating a new “level of data organization, which is made visible and accessible to a user and thus becomes part of their mental model.”<sup>301</sup> The dashboard effectively emerged as the insights interface of Hollywood. Data analysis software facilitated the cross-

---

<sup>301</sup> Ibid, 209

company sharing of dashboards by way of enhanced cloud storage capabilities, enabled the integration with prevalent legacy media software tools (e.g., copying and pasting dashboards into word, PowerPoint, and Keynote software), and contextualized the ongoing flow data points through an existing visual vocabulary. In this regard, the dashboard increasingly normalized the corporate process of leveraging data analysis output in business operations.

Media executives accordingly embraced data analysis software as a key lever of their emerging data analysis infrastructure, leveraging the dashboard to make data available, accessible, and actionable across their organizations. As one executive noted, “We look at it as data democratization: How do I push data into the hands of decision-makers in our business so they can transact and execute on it very quickly.”<sup>302</sup> In particular, the dashboard’s ability to deliver data insights effectively, expeditiously, and engagingly elevated its value for corporate decision-making at the leadership level. One marketing executive at a film and television studio explained thusly: “Putting data front and center actually gets the studio executives excited because it becomes less about their gut and more about how they can react to meaningful information [...]”<sup>303</sup>

Accordingly, the data dashboard enabled a form of corporate storytelling, a way to aggregate, analyze, and action a series of data points as a cohesive insights story that can be communicated, presented, and publicized. The dashboard’s focus on readability, transparency, and analysis effectively positioned it as a discursive vehicle to negotiate corporate decisions. In effect, the dashboard distills the majors’ entire investment in data analysis, from IT and

---

<sup>302</sup> Variety Staff, 2016

<sup>303</sup> Kapko, Matt, “How a small film studio uses Facebook data to compete with Hollywood’s heavyweights,” *CIO*, February 9, 2017, accessed June 6, 2020, <https://www.cio.com/article/3167889/how-a-small-film-studio-uses-facebook-data-to-compete-with-hollywood-s-heavyweights.html>

technology implementation to workforce development and metrics integration, into a simple interface output. It is this ostensible simplicity, the focus on a concentrated visual message, that situates the dashboard in the tradition of other forms of corporate storytelling, including particular Hollywood formats, such as the pitch.

Hollywood's integration of data analysis software involved an increasing focus on the dashboard as an analytical engine and corporate decision-making driver. The majors prioritized the dashboard as a tool to deliver, share, and communicate data insights across their organizational cultures. The dashboard gained prominence within the legacy media industry due to its utilitarian approach to analysis, seamless integration with existing infrastructure, and easy-to-use and -understand functionality, among other factors. More specifically, the dashboard effectively emerged as the central interface of analytical decision-making in Hollywood, acting as the frontend synthesis of various analytical processes conducted by data analysis software applications and professionals on the backend.

--- --- ---

The legacy media industry has consistently taken steps to incubate, implement, and institutionalize a culture of analytics that complements impressionistic forms of corporate decision-making with an analytical, data-driven mindset. The majors effectively developed a comprehensive data analysis infrastructure designed to help legacy media organizations make sense of digital information. To this end, studios and networks built new technological capabilities, expanded existing measurement frameworks, and adopted digital insights tools.

- *Data as a Service*: The majors developed an internal infrastructure of data analysis technology and talent by partnering with a range of data analytics vendors and hiring a

workforce of data professionals, thereby establishing an analytical corporate focus within the legacy media industry.

- *Data as Intelligence*: The majors adopted a suite of new digital metrics to frame data as business insights, expanding their existing measurement framework beyond legacy metrics to negotiate the dynamics of a complex and evolving media environment.
- *Data as Storytelling*: The majors partnered with analytics companies to leverage automated and customizable dashboards to communicate data insights, share them across the organization, and drive corporate decision-making.

Collectively, these industrial developments expanded the legacy media industry's approach to corporate decision-making, imbuing an analytical focus and prioritizing the use of data insights. The process proved long-winding and complicated, involving several challenges, from internal opposition of legacy stakeholders at both the executive and ground level to differences in work cultures, lack of funding, and failed experiments.

While Hollywood had long been classified as a creative industry that operates from the gut, the integration of data analysis software increasingly positioned the business as analytical and data-driven, effectively normalizing the oft-ascribed oppositional dynamic between creativity and data. Philippe Dauman, a long-time industry executive who held chief executive positions at Viacom, summarized the new dynamic succinctly: "Data helps you make more intelligent decisions. Having great creative instincts remains at the core of what we're doing. Data is just another tool that we're going to use." Accordingly, the majors made a concerted effort to position data analysis as a complementary tool in generating insights about the industry, rather than a substitute for the long-standing tenets of cultural production.



At the same time, Hollywood’s culture of analytics gradually expanded its focus toward applying data to concrete business scenarios. While the creative process remained largely sacrosanct,<sup>304</sup> the majors explored industrial strategies and use cases to use data insights to inform how content is produced, marketed, and distributed. As one industry Chief Marketing Officer noted, “I think that insight is the part that people sometimes overlook when it comes to data. There’s also that extra insight piece of what you actually do with the data.”<sup>305</sup> In this regard, film studios and television networks increasingly found themselves negotiating the dynamics of data-driven organizations, not merely rationalizing decision-making through data, but directly executing data-driven decisions.

---

<sup>304</sup> See, Krigsman, 2018.

<sup>305</sup> Verhoeven, Beatrice, “How Data Drives Studio Decisions as Much as Gut Instinct in Netflix Era,” *The Wrap*, October 2, 2018, accessed June 6, 2020, <https://www.thewrap.com/how-data-is-driving-studio-decisions-as-much-as-gut-instinct-in-netflix-era/>

### Part III – Data Application

## Context

### Data and Cultural Production

“Big bets are now being informed by Big Data, and no one knows more about audiences than Netflix.”<sup>306</sup>

David Carr, *New York Times*

“Data application is the lifeblood of how we can push the boundaries of what media companies can be [...]”<sup>307</sup>

Jesse Redniss, Former Executive Vice President of Data Strategy at Turner

The release of *House of Cards*, Netflix’s first original marquee series,<sup>308</sup> proved a major event for the legacy media industry, for several reasons. First, it marked Netflix’s transition from a digital content distributor to programmer, ushering in the company’s era of original programming and creating a new competitive force for the incumbents. Second, it signaled Netflix’s ambition to play an active role in shaping the future of the business, rather than serving as a passive distribution outlet for the majors, creating a new media experience where content is made, marketed, and delivered through a digital interface directly to consumers at home. And,

---

<sup>306</sup> Carr, David, “Giving Viewers What They Want,” *New York Times*, February 24, 2013, accessed June 5, 2020, <https://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html>

<sup>307</sup> Variety, Strictly Business Podcast, 2018

<sup>308</sup> *House of Cards* actually represents the second entry in Netflix’s original series canon after *Lilyhammer*, which premiered in 2012. For Netflix’s originals strategy, see, Rodriguez, Ashley, “Netflix didn’t make many of the “originals” that made it famous. That’s changing,” *Quartz*, February 26, 2019, accessed June 5, 2020, <https://qz.com/1545594/netflix-doesnt-make-most-of-its-originals-now-thats-changing/>

third, it bypassed the established distribution window of the business to make an entire season of television instantly available, offering an alternative to traditional out-of-home releases, originating the binge-watching phenomenon, and challenging the entrenched business model of the industry. What ties this suite of shifts together is that Netflix envisioned, engineered, and executed this series of industrial actions through the application of data.

*House of Cards* has come to represent the blueprint for Netflix's approach to data application. Netflix gave the series an internal greenlight after applying a set of data-driven algorithms, which reportedly indicated that a certain contingent of Netflix subscribers would watch; that the talent involved resonated with certain subscriber segments; and that the investment in the series would generate a financial return. Furthermore, Netflix applied data to market the show, targeting those audience segments most likely to watch to increase reach and viewership. Finally, Netflix applied data to program and surface *House of Cards* to viewers most likely to watch it. Effectively, Netflix relayed the decision to greenlight, market, and distribute *House of Cards* to a complex set of algorithms, developed by internal engineering teams, and informed by Netflix's platform data.<sup>309</sup> As such, the production history of *House of Cards* can be framed as the result of Netflix's data application, the data-driven creation, completion, and commercialization of proprietary algorithms designed to make the series work.

Netflix has systematically applied data to the process of cultural production, beginning with its foundational CineMatch technology, which enabled the company to optimize DVD distribution, guide content licensing deals, and provide early website-based recommendations.<sup>310</sup>

---

<sup>309</sup> For a discussion of algorithms and the differentiation from data, see, Gillespie, Tarleton, "The Relevance of Algorithms." In: *Media Technologies*, edited by Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot (Cambridge: MIT Press, 2014).

<sup>310</sup> Netflix, "Prospectus," *SEC*, May 22, 2002, accessed June 5, 2020, <https://www.sec.gov/Archives/edgar/data/1065280/000101287002002475/d424b4.htm>

The company has consistently made data-driven decisions and implemented industrial actions to ensure its business works, stays competitive, and simultaneously offers a mix of innovative and familiar programming to consumers. *House of Cards* thus does not represent the encapsulation of Netflix's business model. Rather, it constitutes a feature of said model, crystallizing Netflix's commitment to applying data in order to inform the company's approach to the media business.

In essence, Netflix's approach to data application follows a specific cycle. The company leverages its proprietary platform data on consumer behavior and preferences to engineer a complex set of algorithms that inform all stages of cultural production. Netflix accesses, analyzes, and applies data insights to drive business actions, which are executed by algorithms, trained on the company's platform data. As such, Netflix applies data-driven algorithms to initiate, implement, and iterate industrial actions. The New York Times fittingly characterized this model as "running the numbers" and "a product of logic and algorithms as opposed to tradition and instinct."<sup>311</sup> Indeed, Netflix's application of data-driven algorithms introduced an uncommon variable of automation to the business, a system in which the automatic access, analysis, and application of data feeds and fuels a suite of custom-made algorithms that inform content production, marketing, and distribution. Netflix's business model is designed to prioritize data and algorithms over manual institutional processes, endemic to Hollywood's impressionistic mode of decision-making. In this regard, Netflix has envisioned, engineered, and executed cultural production through data application.

---

<sup>311</sup> Carr, 2013

*House of Cards* became the show that was engineered, indeed *made*, by an algorithm, involving “a whole lot of Ph.D.-level math and statistics,”<sup>312</sup> evoking an industrial scenario where creative work is outsourced to a data-driven formula and algorithmic execution.

Conceptually, this idea is neither new nor disruptive given Hollywood’s institutional practice of modifying content based on input from test screenings, focus groups, and survey questionnaires.<sup>313</sup> Netflix’s approach to cultural production effectively added a layer of automation, applying data and algorithms to a largely manual process. As such, it ultimately did not operate as a counter to Hollywood’s traditional process, but a reconfiguration, driven by data and algorithms.

In an interview with the Guardian, Netflix Chief Content Officer Ted Sarandos responded to reigning theories about the company’s application of data-driven algorithms in *House of Cards*, noting that “I didn't use data to make the show, but I used data to determine the potential audience to a level of accuracy very few people can do.”<sup>314</sup> Sarandos here argues that Netflix applied data not to inform the creative vision of *House of Cards*, but the underlying business case. In this sense, data application effectively worked to substantiate the viability of Netflix’s business, to prove out an idea, not to substitute *data* for the *creative*. In characterizing Netflix’s algorithmic approach to programming, Alexis Madrigal, in partnership with Ian Bogost, offered a thorough analysis of this point, noting: “Netflix has created a database of American cinematic predilections. The data can't tell them *how* to make a TV show, but it can tell them *what* they

---

<sup>312</sup> Auletta, 2014

<sup>313</sup> See, Wyatt, Justin, *High Concept: Movies and Marketing in Hollywood* (Austin: University of Texas Press, 1994).

<sup>314</sup> Jeffries, 2013

should be making.”<sup>315</sup> In this sense, *House of Cards* was not engineered by a set of data-driven algorithms. It was enabled by them. Netflix did not change the creative process. Rather, it reconfigured the elements of cultural production that facilitate the creative process.

The production history of *House of Cards* reflects this data logic. Netflix reportedly applied a wide range of proprietary data to model a measurable, built-in audience among its subscribers and thereby validate the financial investment in the series,<sup>316</sup> ultimately revealing a correlation between viewers of the original *House of Cards* series (BBC, 1990), the films of David Fincher, and the filmography of the cast, a trifecta that reportedly sealed the decision to order the series for a straight pick-up and two-season commitment, forgoing the traditional greenlighting process and pilot testing, and releasing all episodes at once for immediate consumption. The dynamics of this data-driven, algorithmic process are notable for several reasons. First, they indicate that Netflix did not apply data to create the idea for *House of Cards*. Rather, it used data to prove that the idea would work on its platform. Second, Netflix applied data to rationalize and commercialize the investment in *House of Cards*, following a standard profit-loss, risk-reward, input-output logic, proving the investment to be commercially viable.<sup>317</sup> And, third, Netflix applied data to recast, revise, and ultimately reduce the impact of a key variable of the business: uncertainty.<sup>318</sup> While Hollywood’s impressionistic mode operated

---

<sup>315</sup> Madrigal, Alexis C., “How Netflix Reverse-Engineered Hollywood,” *The Atlantic*, January 2, 2014, accessed June 5, 2020, <https://www.theatlantic.com/technology/archive/2014/01/how-netflix-reverse-engineered-hollywood/282679/>

<sup>316</sup> See, Harris, Derrick, “Netflix analyzes a lot of data about your viewing habits,” *Gigaom*, June 14, 2012, accessed June 5, 2020, <https://gigaom.com/2012/06/14/netflix-analyzes-a-lot-of-data-about-your-viewing-habits/>

<sup>317</sup> See, Netflix Investor Relations, “Shareholder Letter, Q4, 2010,” January 21, 2011, accessed June 5, 2020, [https://s22.q4cdn.com/959853165/files/doc\\_financials/quarterly\\_reports/2010/q4/Q410-Letter-to-shareholders.pdf](https://s22.q4cdn.com/959853165/files/doc_financials/quarterly_reports/2010/q4/Q410-Letter-to-shareholders.pdf)

<sup>318</sup> See, De Vany, Arthur S., *Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry* (New York: Routledge, 2003).

through patterns of executive confidence, commercial consistency, and institutional intuition, investing in content they *thought* and *hoped* would work, Netflix used data and algorithms in order to *know* what would work.

Netflix's approach to data application has worked to prove, safeguard, and maximize the company's investment in programming, ensuring that any content meets an audience threshold, grounded in data-driven algorithms, rather than intuition, instinct, and impressions. In this context, data operates as a tool to optimize the business of creativity, informing Netflix's actions across production, marketing, and distribution. Indeed, Netflix has consistently applied data across the media value chain to ensure viewers keep watching (i.e., engagement) and stay subscribed (i.e., retention). To this end, the company has used data to predict demand for new content investments (*Production*); segmented the audience on its platform to surface and promote content they are most likely to watch (*Marketing*); and curated content to keep audiences watching (*Distribution*).

Netflix's model of data application is designed to select business actions in the media and cultural production value chain. In this model, data enables the functioning of algorithms, a set of automated computational patterns that select what to *produce* (i.e., identify the films and TV shows to greenlight); what to *promote* (i.e., identify the films and TV shows to market to specific segments of the audience); and what to *program* (i.e., identify the films and TV shows to curate for specific audience segments). In this regard, Netflix's data-driven algorithms select, support, and surface content that maximizes return on investment.<sup>319</sup>

This approach, in concept, is not different from the way Hollywood's legacy media industry has traditionally operated. Indeed, the major film studios and television networks have

---

<sup>319</sup> For a broader application of Netflix's data model across platforms, see, Van Dijck, Jose, Poell, Thomas & de Waal, Martijn, *The Platform Society: Public Values in a Connective World* (Oxford: Oxford University Press, 2018).



long drawn on audience research to greenlight, optimize, and safeguard the creative vision and output, pulling information from focus groups, survey polls, exit interviews, questionnaires, self-reporting diaries, reviews, and even fan mail at various stages of industrial development to ensure film and television output would succeed.<sup>320</sup> Through partnerships with research vendors, the majors have consistently collected, crunched, and consolidated audience feedback to effect changes at the scripting, shooting, editing, and release stage. Hollywood has always carefully packaged its output in order to match audience taste profiles and ensure profitability rather than simply relying on creative instinct entirely, a system which Derek Thompson accurately described as “Hollywood's assembly line of double- and triple-checking [...]”<sup>321</sup> The industry has tailored output to prioritize built-in audiences, as in adapting popular media formats (from mythological stories and historical events to popular books, video games, and podcasts), institutionalizing famous talent (e.g., the star system of the classical period), and relying on repeatable formulas (i.e., sequels, reboots, genre). In this regard, Hollywood has consistently embraced data as a way to tailor the entertainment product and maximize the chance of success.

While Netflix and Hollywood share the same conceptual approach to data application, the execution is markedly different due to a variety of cultural, economic, and technological factors. Netflix’s application of data is systematic across cultural production, applying data holistically across a platform environment to automatically select business actions, fueling a machine that consistently scales through an internal feedback loop.

---

<sup>320</sup> See, Balio, 1985

<sup>321</sup> Thompson, Derek, “The Reason Why Hollywood Makes So Many Boring Superhero Movies,” *The Atlantic*, May 13, 2014, accessed June 5, 2020, <https://www.theatlantic.com/entertainment/archive/2014/05/hollywoods-real-superhero-problem/370785/>

The legacy media industry, by contrast, has applied audience research (i.e., small sets of data, taken at a certain point in time) to selectively alter, iterate, and change the creative output (i.e., reshooting scenes, rewriting the script, adding or deleting scenes, recasting roles).<sup>322</sup> Additionally, the majors have deployed research less consistently, depending on various factors, from the political capital of executives to the relationship with the vendor to the decision-makers involved. Furthermore, audience research application is based on a fragmented, at times siloed, set of data. Test screenings are generally held during the post-production phase; marketing input is routinely gathered prior to the promotional rollout; and distribution dynamics are evaluated a few months prior to the release date. In Netflix's model, data is applied consistently and automatically to fuel algorithms that enable and execute industrial actions. In Hollywood's model, data is applied sporadically, manually, and selectively.

The legacy media industry's approach to data application has not fundamentally changed in the digital era, but it has expanded, driven by an influx of economic, cultural, and technological developments<sup>323</sup> that increasingly challenged the viability of traditional media channels, threatening to erode the core of the legacy media business.<sup>324</sup> While the effects were incremental, rather than immediate, the industry's growing shift toward a digital economy led to downstream impacts on its core business verticals that recast established industry dynamics.

- *Moviegoing*: Theatrical moviegoing has experienced fluctuation over the course of the digital era. While U.S. box office revenue has remained relatively steady, even

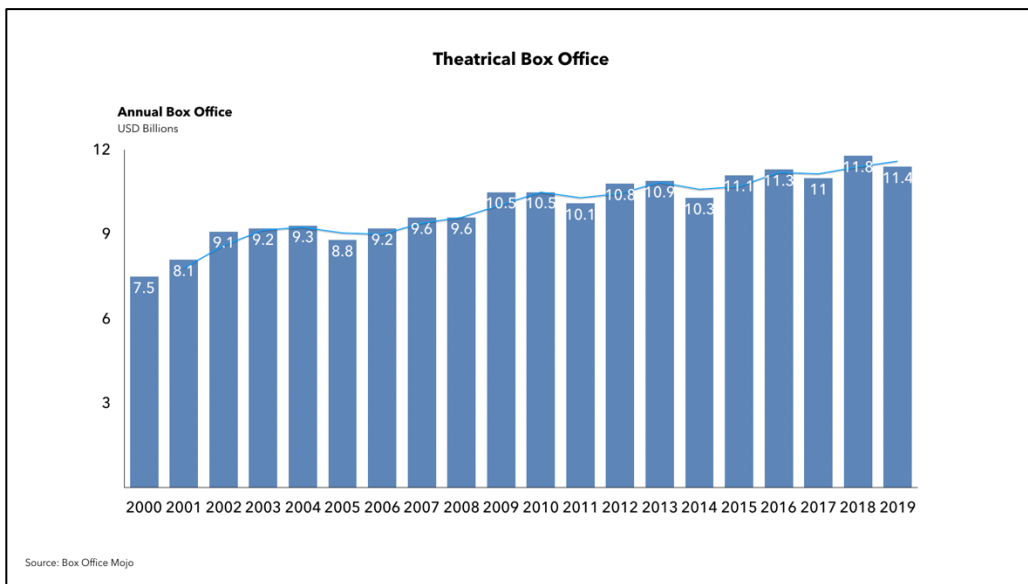
---

<sup>322</sup> See, Wasko, 2003

<sup>323</sup> These developments include, among others, an increase in demand for personalization, the emergence of alternative distribution outlets, the viability of new competitors and partners, and economic pressures.

<sup>324</sup> See, Mann, 2014

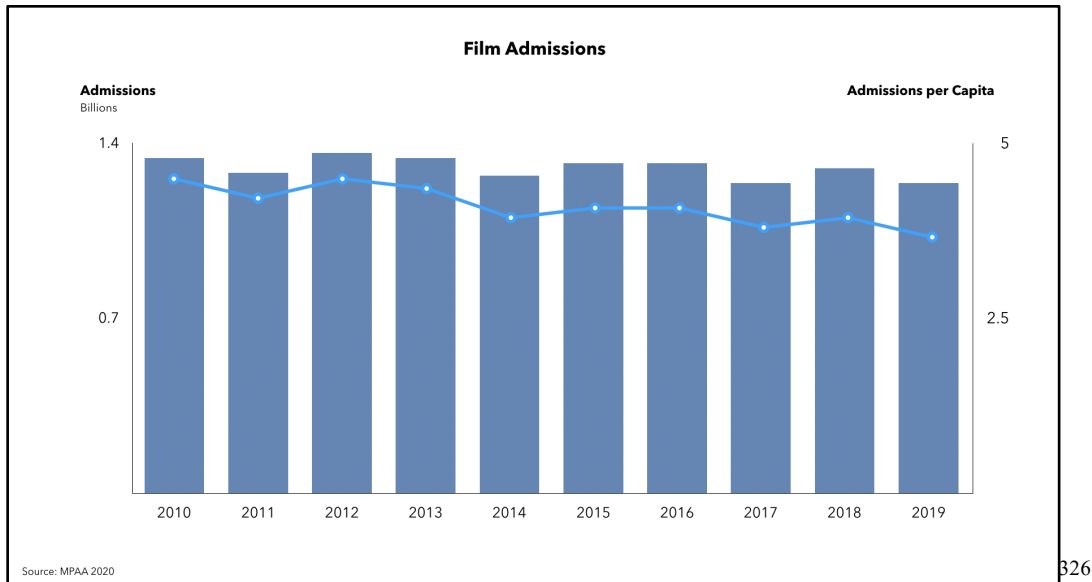
experiencing incremental increases due to growing ticket prices, consolidated release schedules, and a prioritization of global blockbusters, attendance and tickets sold have gradually decreased, effectively shrinking the moviegoing audience. Additionally, the number of frequent moviegoers (i.e., those who visit the theater once a month or more), who traditionally accounted for half of the revenue share, has fluctuated, with younger audiences increasingly turning away from movie theaters and embracing alternative forms of entertainment. Indeed, movie theaters have received increased competition from at-home viewing, driven by Internet-distributed programming, video-on-demand (VOD) services, and mobile viewing devices (figures 3.1 – 3.2).



325

**Figure 3.1:** Theatrical box office has grown incrementally since the early 2000s.

<sup>325</sup> Box Office Mojo, “Domestic Yearly Box Office,” Date as of June 5, 2020, accessed June 5, 2020, <https://www.boxofficemojo.com/year/>

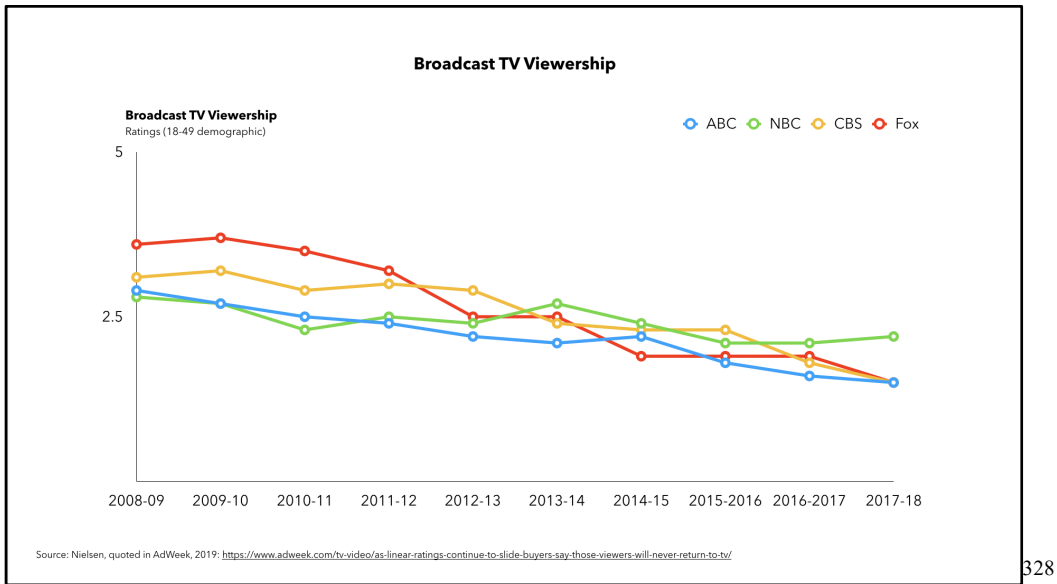


**Figure 3.2:** While film admissions have remained stable, admissions per capita have declined.

- *Television Viewing:* Network and cable television viewing remained steady in the early 2000s, but experienced a viewership decline, fueled by a rise in cord-cutting (i.e., customers cancelling their cable bundle subscription), growth stagnation (i.e., fewer people tuning in), and the proliferation of alternative viewing options (i.e., Internet-distributed television), with the industry struggling to offset viewership and subscriber decline with an increase in affiliate fees and premium advertising.<sup>327</sup> Analysts and pundits forecast the ongoing decline of the medium in light of the platform economy, given the decline in viewers and subscribers and challenged advertising economics due to an increase in viewer age (figures 3.3 – 3.4).

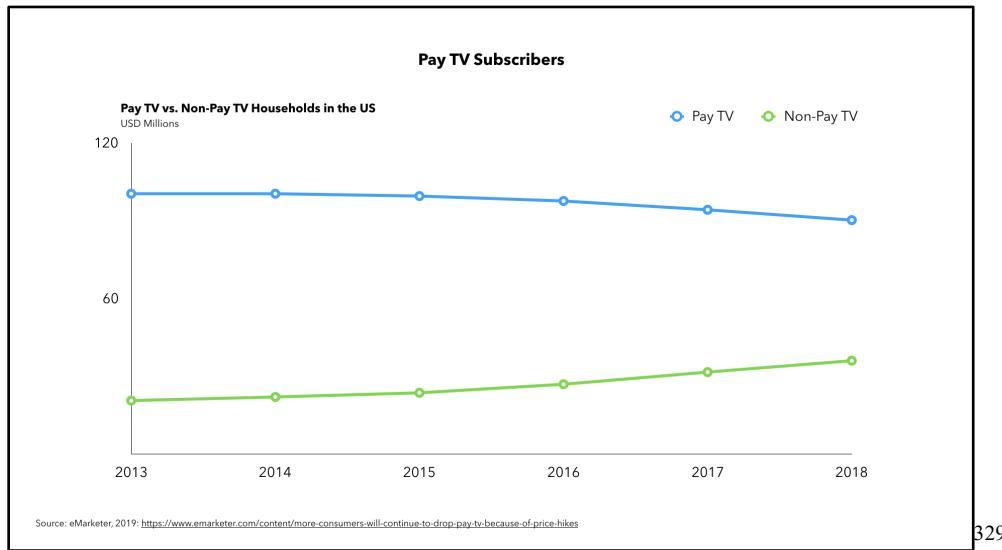
<sup>326</sup> MPAA, “Theme Report 2019,” *Motion Picture Association of America*, March 2020, accessed June 5, 2020, <https://www.motionpictures.org/wp-content/uploads/2020/03/MPA-THEME-2019.pdf>

<sup>327</sup> See, Schomey, Audrey, “Viacom is expanding its streaming distribution to offset ad declines,” *Business Insider*, May 13, 2019, accessed June 5, 2020, <https://www.businessinsider.com/viacom-expands-streaming-distribution-to-offset-ad-declines-2019-5>



328

**Figure 3.3:** Broadcast TV viewership among 18-49 year-olds has declined over the past decade.



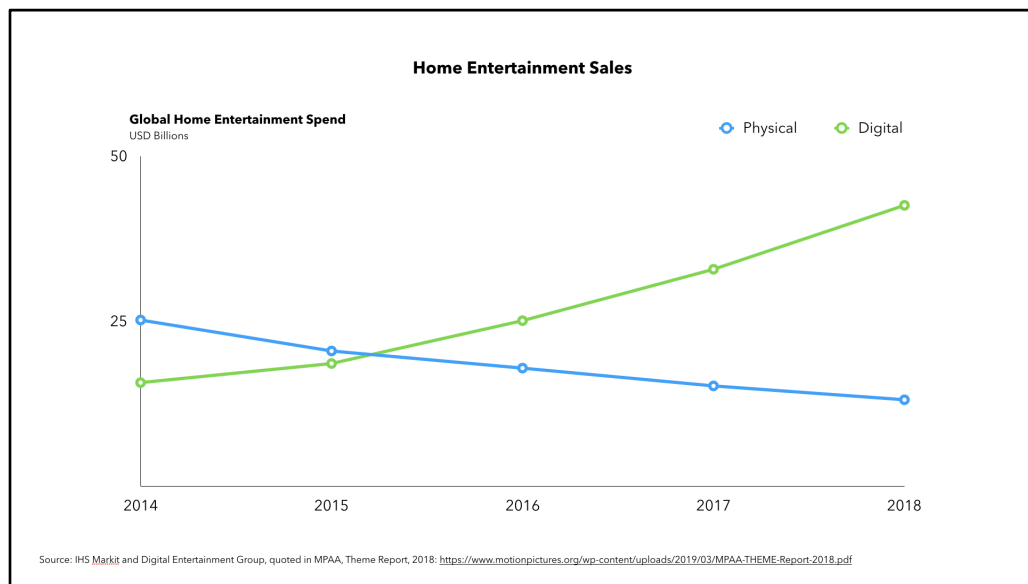
329

**Figure 3.4:** Pay-TV subscriptions have incrementally decreased as direct-to-consumer subscriptions have grown.

<sup>328</sup> Lynch, Jason, “As Linear Ratings Continue to Slide, Buyers Say Those Viewers Will ‘Never’ Return to TV,” *AdWeek*, September 27, 2019, accessed June 5, 2020, <https://www.adweek.com/tv-video/as-linear-ratings-continue-to-slide-buyers-say-those-viewers-will-never-return-to-tv/>

<sup>329</sup> He, Amy, “More Consumers Will Continue to Drop Pay TV Because of Price Hikes,” *eMarketer*, August 8, 2019, accessed June 5, 2020, <https://www.emarketer.com/content/more-consumers-will-continue-to-drop-pay-tv-because-of-price-hikes>

- *Home Entertainment*: Sales of DVDs and Blu-Rays have been in decline since 2005 and 2011, respectively, with viewership shifting to Internet-distributed content, effectively fueling the dissolution of physical media ownership and a once hyper-profitable legacy distribution channel. While home entertainment has shifted to various digital rental and purchasing options to offset the decline in physical sales, platforms have increasingly emphasized direct-to-consumer advertising (AVOD) and subscription (SVOD) options that counter the model of single transactions (figure 3.5).



**Figure 3.5:** DVD and Blu-Ray sales declined as digital home entertainment sales rose, indicating a distinct shift in how consumers consume programming in the post-release windows.

In response to the changing media environment, film studios and television networks have followed multiple strategies to adapt,<sup>331</sup> including increasingly applying data as a risk

<sup>330</sup> MPAA, “Theme Report 2018,” *Motion Picture Association of America*, March 2019, accessed June 5, 2020, <https://www.motionpictures.org/research-docs/2018-theatrical-home-entertainment-market-environment-theme-report/>

<sup>331</sup> See, Perren & Holt, 2011

mitigator and optimization tool to manage the legacy media business. To this end, they have come to apply data across key parts of the cultural production value chain: to decide on future programming slates, identify priority audience segments for marketing initiatives, and increase attendance and tune-in rates across legacy channels. As such, the industry effectively reverse-engineered Netflix's data strategy to reduce uncertainty. While Netflix incubated a new system of production, applying data and algorithms to automatically increase the success rate of its programming, Hollywood added data as an incremental layer to increase business success without disrupting traditional practices.

Part I traced the industrial strategies and developments that established an evolving infrastructure of data access within the legacy media industry. Part II outlined the formalization of data analysis as a new variable in Hollywood's approach to corporate decision-making. Part III builds on this narrative by examining how the application of data has impacted the way content is produced, marketed, and distributed. It explores how film studios and television networks have applied data to the traditional process of cultural production, using data as an overlay to existing legacy practices, complementing industrial actions informed by instinct, institutional knowledge, and audience research.

Using this approach, Part III aims to illustrate several points. One, data application is a complex industrial practice with different manifestations, ranging from the holistic and automated to the manual and selective use of data. Two, data application is a business optimization tool that carries significant implications for what content is produced, how it is marketed, and where it is distributed. And, three, data application operates dynamically across the legacy media business and the digital economy, responding to cultural, economic, and technological forces.

As such, data application continues the data integration narrative set out in the previous two parts of this project, further illustrating how data has come to function as a growing element of Hollywood's operating logic. The legacy media industry has come to leverage data access to generate, license, and own information (*Data Access*); employ data analysis to translate the information into business insights that shaped corporate decisions (*Data Analysis*); and apply these decisions to effect industrial actions that directly impact the business (*Data Application*). The result is not a linear set of procedural steps, but an ongoing feedback loop, an evolving system of data-driven practices and developments. The actions resulting from data application generate more data, feeding the access infrastructure and analytical machine, which again drive corporate decisions and industrial executions. Hollywood has built its own version of the platform flywheel, a system that evolves through ongoing information input.<sup>332</sup>

In effect, Part III posits data application as an industrial strategy to reduce uncertainty, eliminate risk, and maximize return on investment in cultural production. This rationale runs the risk of positioning data as infallible and foolproof, evoking a false sense that the application of data invariably ensures success. This project aims to take a more balanced approach, building on the concept that data represents capital, while exercising caution in ascribing it too much power. In particular, the project does not aim to make any claims about the economic viability of data application and thus does not intend to prove whether data application actually works in improving the balance sheet of industrial output. While data adds a quantifiable and provable dimension to cultural production, data invariably constitutes an interpretation, an inference, and an intellectual construct. Data manifests as a contextual reading that informs corporate thinking, decision-making, and action, not a fact-based predictor that guarantees a specific outcome.

---

<sup>332</sup> For examples of the flywheel business model, see, Petruska, Karen, "Amazon Prime Video: Where Information is Entertainment." In: Johnson, 2018.



Indeed, the application of data by tech companies has not proven to be a steady guarantor of success. For example, Amazon Studios, the entertainment division of retail giant Amazon, has shifted its programming strategy on multiple occasions, incurring losses along the way, even though it has maintained access to massive amounts of data.<sup>333</sup> Thus, the application of data is a dynamic and iterative process that depends on a variety of factors, a strategic measure to manage the business, not a magic formula to increase success.

To this end, Part III seeks to engage with a core set of questions around Hollywood's integration of data application:

- How did the legacy media industry apply data across the media value chain and what were the implications for the processes of production, marketing, and distribution?
- What software applications did the major studios and television networks make use of in order to manage cultural production through data?
- How did the majors negotiate and reconcile the dynamics of data application in a shifting media environment?

To answer these questions, it is vital to think about data application as a working set of experiments rather than a proven scientific formula. To drive return on investment and increase the rate of success in the shifting digital landscape of media and entertainment, the majors have

---

<sup>333</sup> Admittedly, Amazon Studios programming has helped Amazon increase Amazon Prime subscriptions, achieving the company's overall goal, yet Amazon's movies and TV shows have often underperformed in the market based on legacy metrics. See, McClintock, Pamela, "As Amazon Suffers String of Box Office Flops, Executives Struggle to Find Winning Strategy," *The Hollywood Reporter*, June 26, 2019, accessed June 5, 2020, <https://www.hollywoodreporter.com/news/amazon-studios-film-division-tumult-string-box-office-flops-1220968>

applied data in a variety of ways, seeking to elevate traditional homegrown practices while negotiating newly developed external practices.

### **A Framework for Data Application**

Part III examines an integrated approach to data application in cultural production. To this end, it posits data application as the use of data and algorithms in content production, marketing, and distribution to minimize uncertainty, increase the probability of success, and generate a return on investment in a shifting media environment. In particular, it considers how the legacy media industry applied data to preserve, support, and optimize the processes and profits associated with the legacy media business. As such, part III traces three ways data application has manifested in the legacy media industry:

- *Data-Driven Production:* The majors have increasingly drawn on a variety of data sources to manage content production, applying data to greenlight ideas, select talent, and inform creative direction. In effect, the legacy media industry has looked at data to validate content investments, aiming to minimize uncertainty and reduce the margins of error inherent to impressionistic modes of decision-making.
- *Data-Driven Marketing:* The majors have worked to apply data as a marketing tool to measure, segment, and engage the media audience in ways previously considered both unconventional, uncommon, and unproven, thereby reconfiguring a system built on institutional wisdom, entrenched behavior, and inflexible methodologies. In particular, they have looked at data as a way to customize general marketing practices, introducing flexibility, dynamics, and nuance to the long-held tenets of mass media outreach.

- *Data-Driven Distribution*: The majors have applied data to boost the performance of legacy media channels, driving attendance in theaters and tune-in on television. Specifically, they have used data to more efficiently allocate resources, update distribution strategies, and implement new business models.

In this sense, data application has been deployed not to redefine but to reconfigure, enhance, and optimize cultural production, with the goal of driving return on investment, minimizing uncertainty, and boosting performance. Rather than adopting the data-driven model of platforms such as Netflix, the majors have applied data as an additional layer to the way they have always done business, as a way to balance, counter, and at times check the legacy motivators driving industrial actions. As such, the majors did not restructure their business to become platforms. Rather, they applied data to operate more like them. In this reverse constellation of convergence, the majors effectively sought to adopt platform features in order to preserve the legacy media business. As such, Part III further looks at the legacy media industry's application of data to negotiate the emerging demands of the digital economy built on the business model of platforms, balancing the rise and influence of new technologies, shifting consumer sensibilities, and evolving business economics, effectively charting a path to operate in a new media environment that is markedly different from the legacy media business.

Indeed, the business focus of the legacy media industry, long centered on *pushing out* content periodically to draw in the audience for a specific time period (i.e., a theatrical event, primetime viewing schedules, or a home entertainment release), has expanded to managing content constantly in order to *pull in* data consistently and keep the audience locked in at all

times.<sup>334</sup> Keeping viewers watching provides a constant stream of data, which, in turn, enables the application of data to keep the audience engaged. This direct connection with the audience, the legacy media industry's inherent digital feedback loop, has introduced new dynamics into the process of cultural production, opening up new opportunities and challenges for the legacy media business in data application. The majors have effectively applied data to manage a media business in transition, at once adopting the business model of platforms while simultaneously revising data-driven practices to fit alternative legacy modes of cultural production. As such, Part III seeks to illustrate that the majors have built a dynamic mode of cultural production to preserve the legacy media business while equally navigating the new media experience, hybridizing operational and strategic elements from the digital economy and the legacy media industry into a new data-driven business logic.

This idea of complex convergence is further illustrated by shifting perceptions and attitudes on data application in the relationship between Hollywood and Silicon Valley. While platforms initially emerged as data purists, deploying algorithms to automate cultural production and promoting data as an industrial differentiator, the new entrants have increasingly adopted a more impressionistic tone, following the conventions of the legacy media business. Ted Sarandos, Netflix Chief Content Officer, initially emphasized the impact of data in the process of media production, noting that data informed 70% of Netflix's approach, while later reversing the formula, stating that "[i]t's 70 percent gut and 30 percent data [...]."<sup>335</sup> Similarly, the former head of Amazon Studios clarified that the retail giant's approach to entertainment balanced data and

---

<sup>334</sup> See Annette Hill & Janette Steemers, "Media Industries and Engagement," *Media Industries Journal*, Volume 4, Issue 1, 2017.

<sup>335</sup> Adalian, Josef, "Inside the Binge Factory," *Vulture*, June 10, 2018, accessed May 4, 2020, <https://www.vulture.com/2018/06/how-netflix-swallowed-tv-industry.html>

creative decision-making: “It’s not like you can come in on Tuesday and the computer says: ‘Doot, doot, doot. Here are the shows you are going to do’ It’s not ‘The Barefoot Executive.’ You have to use some judgment as well.”<sup>336</sup> By the same token, the majors have balanced skepticism about data with a seemingly growing curiosity and openness. Former Paramount Chief Executive Officer Philippe Dauman noted: “Data helps you make more intelligent decisions. Having great creative instincts remains at the core of what we’re doing. Data is just another tool that we’re going to use.”<sup>337</sup>

The converging views on data application highlight an ongoing negotiation of different approaches to the media business, a complex dynamic where legacy media companies and platforms have consistently reconfigured established practices to explore new ways to produce, market, and distribute content. The result is a media environment in which Hollywood and Silicon Valley maintain a complicated cycle of convergence, with data emerging as a central, underlying crux of cultural production.

---

<sup>336</sup> Steel, Emily, “For Its New Shows, Amazon Adds Art to Its Data,” *New York Times*, August 15, 2014, accessed June 5, 2020, <https://www.nytimes.com/2014/08/16/business/media/for-its-new-shows-amazon-adds-art-to-its-data.html>

<sup>337</sup> Littleton, 2015, “Viacom’s Philippe Dauman Talks New Metrics, Creative Tools and Wall Street’s ‘Short-Termism’”

## Chapter 3

### The Algorithmic Age: Optimizing the Legacy Media Business

“Nobody knows anything. Nobody, *nobody* — not now, not ever — knows the least goddam thing about what is or isn’t going to work at the box office.”<sup>338</sup>

William Goldman, Screenwriter and Author

“[N]obody had access to the Holy Grail of Hollywood, which is its film data. [...] We have more film data than anyone in this business. And information is power.”<sup>339</sup>

Ryan Kavanaugh, Founder of Relativity Media

In 2004, Ryan Kavanaugh founded Relativity Media, a brokerage firm that secured investments from banks and hedge funds to develop long-term slate financing deals for film studios. After completing a suite of co-financing packages, including deals with Sony Pictures Entertainment and Universal Pictures involving funds for a total of 19 feature films,<sup>340</sup> Relativity gradually transformed into a dedicated media company, leveraging private equity investments to expand into production and distribution, home entertainment, music, sports, and digital entertainment through a suite of acquisitions and production deals. As such, the company produced, distributed, and merchandised a range of film and television programming, building a sizable content

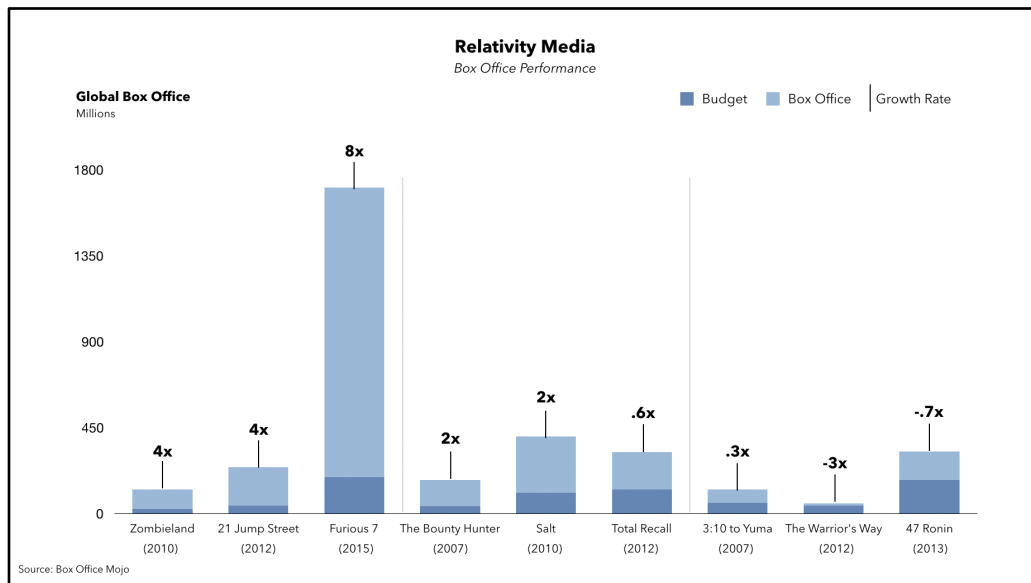
---

<sup>338</sup> Goldman, 1983

<sup>339</sup> Block, Alex Ben, “Ryan Kavanaugh’s secret to success,” *Hollywood Reporter*, September 29, 2010, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/ryan-kavanaughs-secret-success-28540>

<sup>340</sup> See, Kay, Jeremy, “Relativity Media fires up Gun Hill Road package,” *Screen Daily*, May 11, 2006, accessed June 6, 2020, <https://www.screendaily.com/relativity-media-fires-up-gun-hill-road-package/4027164.article>

repository that generated robust returns (e.g., Sony’s *Zombieland* and *21 Jump Street*, and Universal’s *Furious 7*), modest hits (e.g., Sony’s *The Bounty Hunter*, *Salt*, and *Total Recall*), as well as a string of underperformers (e.g., Lionsgate’s *3:10 to Yuma*, Rogue’s *The Warrior’s Way*, and Universal’s *47 Ronin*) (figure 3.6).



**Figure 3.6:** A snapshot of Relativity’s box office reveals a wide spectrum of financial performance.

With its shift into cultural production, Relativity effectively rose from external financier to recognized mini-major within a decade, before ultimately declaring bankruptcy in 2015. At the time, the company counted over 100 films under its banner, marking an annual output that nearly doubled the average of the major film studios, with billions of dollars in capital invested in Hollywood. Relativity’s legacy, however, is not *what* the company invested its capital in, but the underlying logic of *how* it decided to invest.

Relativity promoted a business strategy that emphasized an automated approach to cultural production. The company’s core industrial product was *the model*, a data-driven set of computational calculations that reportedly enabled more secure investments into the media and entertainment business. Effectively, it positioned its value proposition as “a risk-assessment

algorithm that [...] takes a lot of the mystery out of the movie business.”<sup>341</sup> In effect, the company programmed, patented, and promoted a proprietary algorithmic model as a risk mitigator for the legacy media industry, reportedly leveraging a broad range of industry data to identify financially profitable investments. Based on the Monte Carlo method,<sup>342</sup> Relativity’s predictive model crunched a variety of data points to simulate and test random financial scenarios for films, thereby evaluating the probability of possible box office outcomes. According to Relativity, the model worked to algorithmically determine and then exclude investments that did not meet its internal profit threshold. As such, it functioned as a “movie-rejection system,”<sup>343</sup> filtering out high-risk ventures to focus on profitability and return on investment.

While Relativity applied outside Wall Street hedge fund methodology to the film business, its model effectively ran on inside Hollywood numbers. Kavanaugh reportedly built a comprehensive database of studio financial data, aggregating ten years of ultimates (i.e., studio revenues across all available platforms, from theatrical to physical and digital home entertainment) alongside massive amounts of metadata on film and television programming, including information on cast and crew, genre, rating, and release year, among others.<sup>344</sup> The prospect of receiving outside capital infusion, mitigating risk, and achieving financial return made for an appealing pitch to studios and incentivized them to engage in a data-sharing

---

<sup>341</sup> Digiaco, Frank, “The Theory of Relativity,” *Vanity Fair*, February 9, 2010, accessed June 6, 2020, <https://www.vanityfair.com/news/2010/03/kavanaugh-201003>

<sup>342</sup> The method describes the use of algorithms to conduct random sampling in order to determine patterns.

<sup>343</sup> Digiaco, 2010

<sup>344</sup> Ibid



partnership with Relativity, which provided the company with unprecedented industry information. On his studio pitch, Kavanaugh noted (my emphasis):

“I went to all these guys and said, ‘You're going to need a lot of money and there is only one place to get it: Wall Street. [...] **But the only way I can do that is to have pure access to your data.**’ We got data from virtually every studio.”<sup>345</sup>

Accordingly, Relativity assembled a data repository that documented “[film] revenues, line by line; costs, line by line; and profitability.”<sup>346</sup> The company leveraged its model to crunch the available data and generate risk profiles on new film ideas, effectively detailing “the percentage of time the movie will be profitable, and the average profit for each profitable run.”<sup>347</sup> In effect, Relativity seemingly developed a formula for film studios, and subsequently television networks, to determine success of their investments up front.

Relativity characterized the process as running a “geek squad” of data analysts focused on assessing the risk and profitability of new studio ventures. They fed studio data into the company’s proprietary algorithmic model and evaluated the commercial prospects of ideas by considering the historical track records of everyone involved, from the studio’s past investments into similar genres to the actors’ and directors’ box office history as well as the financial outcome of related investments from competitors. Effectively, they used studio financial data as input and applied metadata filters (i.e., industry-related information such as studio revenues, cast and crew, and genre) to generate modulated scenario output (i.e., different financial outcomes by input and filter). In this regard, the company applied insights, generated from a repository of

---

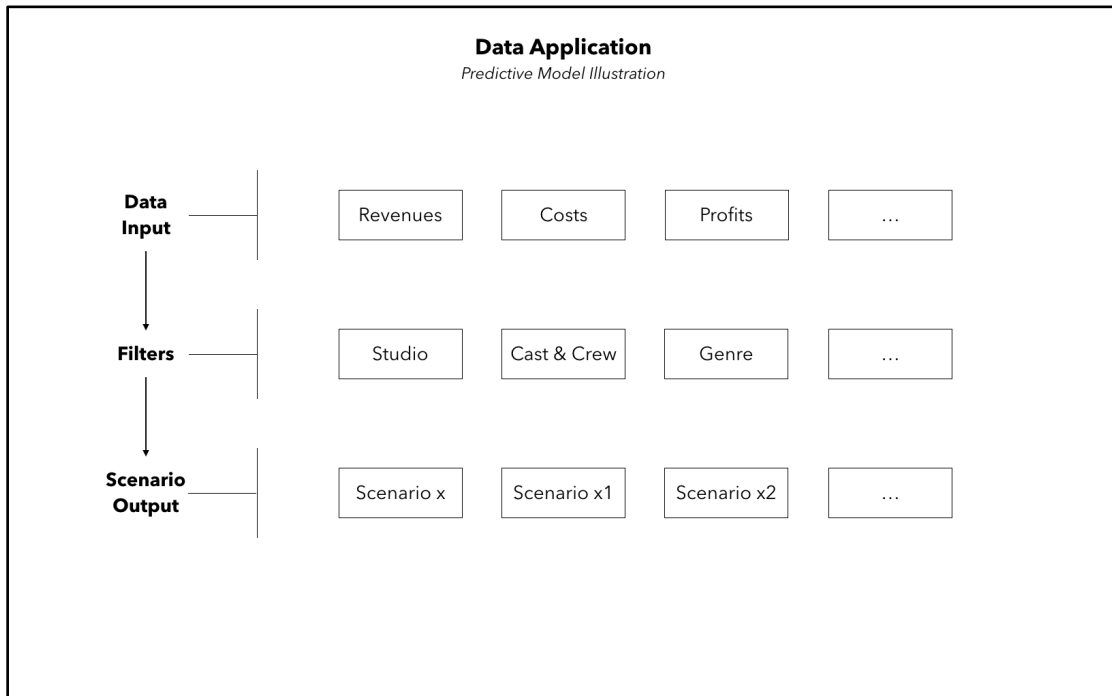
<sup>345</sup> Block, 2010

<sup>346</sup> Digiacommo, 2010

<sup>347</sup> Ibid

data, to an algorithmic model that automatically forecast the commercial prospects of early-stage film ideas. As such, Relativity performed an exercise in data application to manage the economic and cultural dynamics of the legacy media business.

In this regard, Relativity constitutes a *data application* company, a venture that deployed data-driven insights to fuel an algorithmic model designed to automate critical practices of cultural production. Relativity positioned its work in data application as a more efficient way to allocate capital and secure returns on investment,<sup>348</sup> appealing to the business needs of Hollywood’s economic brass. At the same time, Relativity’s pragmatic and rational approach evoked an abstract air of alchemy, arguably mixed with an undercurrent of technical obfuscation and faux omniscience, in an industry driven by impressionistic decision-making (figure 3.7).



**Figure 3.7:** The data application model leverages data input to filter and predict the outcome of a given scenario. Output is dependent upon data input and the applied filters in the model.

<sup>348</sup> Waxman, Sharon, “Ryan Kavanaugh’s Relativity Turns 10: The CEO on Succeeding Despite Predictions and What’s Next,” *The Wrap*, May 19, 2014, accessed June 6, 2020, <https://www.thewrap.com/relativity-ceo-ryan-kavanaugh-on-10th-anniversary/>

Relativity's model of numbers-based certainty consequently offered an effective strategy to balance Hollywood's long history of picking the right content based on institutional knowledge, creative intuition, and gut feeling, representing an incremental approach to managing the erratic, uncertain, and fluctuating nature of the business. Indeed, the legacy media industry's impressionistic formula has always been marked by uncertainty. On the one hand, the industry displays multiple cases of unexpected success, from global low-budget phenomena (e.g., Artisan's *The Blair Witch Project* and Paramount's *Paranormal Activity*) to massively successful B-movie prestige ventures (e.g., Warner Bros.' *The Exorcist*) and multi-season TV smash hits (e.g., NBC's *Seinfeld* and *Friends*). On the other hand, there are an array of surprising failures, including supposedly sure-thing IP- and talent-driven blockbuster productions that did not work (e.g., Disney's *John Carter*), failed vehicles for bankable stars (e.g., Dwayne *The Rock* Johnson's *Baywatch*), and IP translations with in-built audiences gone awry (e.g., Sony's *Men in Black: International*). As Hollywood shifted to a system with reduced output at increasingly high budgets on a global scale,<sup>349</sup> putting more pressure on individual investments, the majors were looking for additional ways to manage risk and control the financial narrative. At the same time, the growing availability of digital content and the maturation of platforms as alternative distribution channels challenged Hollywood's legacy media business model, putting increased strain on the studios to generate returns across traditional channels.

In this environment, Relativity positioned its offering as a tool to manage the shifting dynamics of the legacy media business. The company effectively sold the legend of a magical media industry algorithm, an automated, data-driven system to reduce risk, control costs, and,

---

<sup>349</sup> See, Elberse, 2013

most notably, remove uncertainty, a true “Hollywood Houdini trick.”<sup>350</sup> The prospect of establishing control over the fickle vicissitudes of the legacy media business shaped Relativity’s position in Hollywood as an innovative and forward-looking venture. While the company garnered supporters, including studio heads and actors, it was not universally accepted, notably drawing resistance from old-guard media financiers dubious about the company’s purported certainty model. As one executive noted, “[t]here’s no black box, no magic bean that you can plant and make the movie business more profitable.”<sup>351</sup> Indeed, Relativity’s algorithm ultimately did not deliver on its conceptual logic. While it reduced the overall risk of film investments by slate, minimizing the probability of large-scale financial fallouts, it effectively spread smaller losses over time, thereby reducing long-term gains for investors.<sup>352</sup> As such, the model, for all its input, did not work.

Yet, despite the failed technical dynamics and business value of its algorithmic model, Relativity still crafted a differentiated industrial narrative of risk mitigation and data-driven certainty that enabled the company to shape the legacy media business for over a decade. Responding to a permeating thirst for knowledge and growing desire to “minimize guesswork”<sup>353</sup> in an industrial environment marked by new competitors, shifting consumer habits, and alternative monetization models, Relativity packaged an all-in-one solution that promised to

---

<sup>350</sup> Block, 2010

<sup>351</sup> Digiacomo, 2010

<sup>352</sup> Zhang, Jack, “Why the Algorithm That Promised to Save Hollywood Destroyed Relativity Media,” *No Film School*, April 13, 2016, <https://nofilmschool.com/2016/04/why-algorithm-promised-to-save-hollywood-destroyed-relativity-media>

<sup>353</sup> Barnes, Brooks, “Solving Equation for a Hit Film Script, with Data,” *New York Times*, May 5, 2013, accessed June 6, 2020, <https://www.nytimes.com/2013/05/06/business/media/solving-equation-of-a-hit-film-script-with-data.html>

forecast commercial success based on historical numbers and a distinct set of industry variables. Relativity's approach resonated with an industry where "nobody knows what commercial is [...]"<sup>354</sup> and success is invariably determined by "a roll of the dice."<sup>355</sup> As former Hollywood executive and studio head David Picker, who helped greenlight successful ventures such as the *James Bond* franchise (1962 - Present), *Saturday Night Fever* (Paramount Pictures, 1977), and *Grease* (Paramount Pictures, 1978), pointedly explained: "If I had turned down every picture I greenlit, and greenlit every picture I turned down, I'd have the same number of hits and flops."<sup>356</sup>

Relativity's vision of an automated control mechanism for legacy media profits, driven by data and algorithms, was generally discussed as a unique industrial phenomenon, but it did not represent an isolated incident. In fact, it reflected a broader trend of data application within the evolving assembly lines of the legacy media industry. Over the 2000s, Hollywood attracted an increasing number of data application companies, a range of startups, mid-market companies, as well as established platforms specializing in media and entertainment business optimization and risk mitigation via the application of data-driven, algorithmic software solutions. The time period from 2010 to 2015, in particular, saw an influx in data application players, positioned as algorithmic service partners for the major studios and networks, fueled by emerging technology

---

<sup>354</sup> IBM

<sup>355</sup> Carr, 2013

<sup>356</sup> Picker, David H., *Musts, Maybes, and Nevers: A Book About The Movies* (Scotts Valley: CreateSpace Independent Publishing, 2013).

development, rising venture funding rounds, and growing industry demand.<sup>357</sup> As the New York Times observed pointedly at the time, “[d]ata crunchers are invading Hollywood.”<sup>358</sup>

Indeed, the wave of data application companies effectively built and expanded on Relativity’s value proposition of the *model*, promoting a proprietary set of algorithms that automate idea testing to identify profitable film investments.<sup>359</sup> For example, Piedmont Media Research, a Los Angeles-based startup, utilized an algorithm to project the correlation between an audience’s interest in plot and actors with box office outcome, providing a quantitative score to inform greenlighting decisions.<sup>360</sup> Similarly, Pilot, a software company from Boston, developed an algorithm that mined industry-specific data inputs - from the cast and creative team to genre, budget, release date, and others - with the goal of predicting a film’s box office revenue and thereby supporting early-stage production rationales.<sup>361</sup> Companies like Relativity, Piedmont, and Pilot effectively functioned as film performance forecasters, running algorithms against a growing corpus of internal and external data sets to model demand and inform the dynamics of cultural production.

---

<sup>357</sup> The majority of companies profiled in this project emerged during this time period, as based on public data compiled from Crunchbase.

<sup>358</sup> Barnes, 2013, “Solving Equation for a Hit Film Script, with Data”

<sup>359</sup> See, Vincent, James, “Hollywood Is Quietly Using AI To Help Decide Which Movies To Make,” *The Verge*, May 28, 2019, accessed June 6, 2020, <https://www.theverge.com/2019/5/28/18637135/hollywood-ai-film-decision-script-analysis-data-machine-learning>

<sup>360</sup> See, Avirgan, Jody, “Podcast: The Guy Who Predicts Whether A Movie Will Bomb, Months Before It’s Made,” *Five Thirty Eight*, September 17, 2015, accessed June 6, 2020, <https://fivethirtyeight.com/features/podcast-the-guy-who-predicts-whether-a-movie-will-bomb-months-before-its-made/>

<sup>361</sup> See, Adams, Dan, “Can this Boston startup predict films’ success?,” *Boston Globe*, October 22, 2016, accessed June 6, 2020, <https://www.bostonglobe.com/business/2016/10/21/can-this-boston-startup-predict-film-grosses/ib6zL7GR0Qj6FoKc7OpCRN/story.html>

At the same time, data application companies expanded beyond production financing, vertically integrating with the entire media and entertainment value chain by delivering a suite of solutions across film and television production, marketing, and distribution (table 3.1). Production-focused solutions applied data-driven algorithms to ideas, pitches, scripts, and casting lists to model box office outcomes and increase production efficiencies (e.g., Epagogix, IMDb Pro, Vault). Marketing-based software solutions analyzed online audience behavior and feedback on creative concepts and promotional materials to identify viable audience segments to efficiently target campaign outreach for new releases (e.g., Fizziology, Canvs, Cinelytic). Distribution-centric solutions evaluated release and ticketing strategies to project turnout and tune-in, unlocking opportunities to drive incremental demand during the traditional release window (e.g., Atom Tickets, Get Glue, MoviePass).<sup>362</sup> Effectively, the companies applied data to automatically determine demand at various stages of cultural production, working to mitigate risk, remove uncertainty, and increase probability of success for film and television output.

While individual companies pursued distinct business strategies in an effort to create competitive differentiation (e.g., MoviePass's Netflix-inspired model of cinema subscriptions to increase movie-going attendance during the theatrical release window) and capitalize on emerging trends (e.g., Get Glue's social check-in feature to drive television tune-in and appointment viewing), their overall focus positioned them as media and entertainment value chain add-ons, automated solutions that deliver incremental value to film studios and television networks. Indeed, the solutions offered a new way to improve the economics of the business by applying data-driven software solutions to the established process of film and television cultural production. As such, they collectively formed a developing ecosystem of partner services for the

---

<sup>362</sup> Several data application companies operate across all stages of the media and entertainment value chain, offering integrated software products and services across production, marketing, and distribution.

majors to optimize the way content is produced, marketed, and distributed, effectively representing a new lever to manage the legacy media business.

**Table 3.1**

Cultural Production Value Chain			
<i>Value Chain</i>	<b>Production</b>	<b>Marketing</b>	<b>Distribution</b>
<i>Use Cases</i>	<ul style="list-style-type: none"> <li>● Financing / greenlighting               <ul style="list-style-type: none"> <li>○ Idea testing</li> <li>○ Script analysis</li> <li>○ Casting evaluation</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>● Marketing outreach               <ul style="list-style-type: none"> <li>○ Creative testing</li> <li>○ Promotional analysis</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>● Release strategy               <ul style="list-style-type: none"> <li>○ Rollout assessment</li> <li>○ Release timing</li> </ul> </li> </ul>
<i>Sample Companies</i>	<ul style="list-style-type: none"> <li>● Piedmont</li> <li>● Pilot</li> <li>● Epagogix</li> <li>● Worldwide Motion Picture Group</li> <li>● Vault</li> <li>● ScriptBook</li> <li>● theAudience</li> <li>● Canvs</li> <li>● Parrot Analytics</li> <li>● Cinelytic</li> <li>● IMDb Pro</li> </ul>	<ul style="list-style-type: none"> <li>● Crimson Hexagon,</li> <li>● Fizziology</li> <li>● ListenFirst</li> <li>● Relish Mix</li> <li>● theAudience</li> <li>● Canvs</li> <li>● Parrot Analytics</li> <li>● Cinelytic</li> </ul>	<ul style="list-style-type: none"> <li>● Get Glue</li> <li>● Atom Tickets</li> <li>● MoviePass</li> <li>● Moviefone</li> <li>● Parrot Analytics</li> <li>● Cinelytic</li> </ul>

While the focus on data-driven and technology-enabled optimization has been characterized as an innovative practice in Hollywood,<sup>363</sup> the rise of data application companies effectively represents an extension of the legacy media industry’s traditional research approach. Indeed, the majors have long practiced research initiatives to shepherd film and television output, drawing on various audience feedback mechanisms like focus groups, survey polls, exit interviews, questionnaires, self-reporting diaries, reviews, and even fan mail to effect changes at

<sup>363</sup> As an example, see Variety’s investment in an annual Big Data innovation summit featuring industry executives.



the production, marketing, and distribution stage.<sup>364</sup> As such, studios and networks have always carefully packaged programming in order to meet audience expectations and maximize return on investment. In this regard, data application companies effectively mark a continuation of the industry's legacy practice of applying data to the system of cultural production.

At the same time, the new data entrants introduced a set of digital features that differentiated their offering from the legacy media industry. One, while legacy research vendors focused on small representative samples and directed feedback (i.e., structured data from survey and focus group questionnaires), data application companies emphasized data across the digital landscape, managing increasing information volume and depth as well as free-flowing, unfiltered feedback (i.e., unstructured data from various formats, such as posts, images, and video). Additionally, data application companies increasingly drew on data inputs from across the legacy media industry, either directly sourcing metadata inputs from the majors (as in Relativity's case) or indirectly accessing information through application programming interfaces (APIs) from industry aggregators such as IMDb and Studio System, thereby adding incremental industry-specific details to their models.

Two, legacy vendors largely maintained a manual process in accessing, analyzing, and applying audience feedback, with incremental investments in digital technologies and software,<sup>365</sup> whereas data application companies invested in software automation, which enabled them to regularly update their work output with new information, respond to changes in content

---

<sup>364</sup> See, Ohmer, 2006 and Ward, Karen, "Hollywood's conception of its audiences in the 1920s." In: *The Classic Hollywood Reader*, edited by Steve Neale (New York: Routledge, 2012).

<sup>365</sup> See, Wyatt, 2014

and audience data, and scale their services across the value chain. The use of automation ensured advanced information scale and a capability to dynamically react to incoming flows of data.

And, three, while legacy vendors positioned their output as mainly directional, data application companies increasingly worked to substantiate their business optimization guidelines as scientifically robust and definitive, promising a higher probability of success. Indeed, the companies constructed a promotional data discourse, running ongoing model displays to *prove out* the validity of their predictions. To this end, data application companies partnered with the trades on weekly box office prediction columns,<sup>366</sup> major platforms published white papers outlining the industrial benefits of predictive models,<sup>367</sup> and startups partnered with studios and networks on one-time experiments, presenting the results in ready-made and easily readable dashboards to self-promote and -publicize their value proposition across the broader industry.

The result was an industrial environment increasingly geared toward “a process of rationalization,”<sup>368</sup> in which the majors more and more relied on data application as an incremental strategy to manage cultural production. The New York Times captured the shifting environment pointedly (my emphasis):

“Production executives are relying more on data crunchers to guide their decisions. **Armed with algorithms, Facebook data and focus group research, a new breed of consultant is swarming Hollywood:** cast this actress, don’t cast that actor, invest in this script, burn that one.”<sup>369</sup>

---

<sup>366</sup> For example, Variety and ListenFirst launched Digital Audience Ratings (DAR), a metric for audience engagement across social media. See, Klein, Jason, “Digital Audience Ratings” *Variety*, October 1, 2015, accessed June 6, 2020, <https://variety.com/2015/digital/news/digital-audience-ratings-strong-showing-for-sony-as-family-favorites-top-the-week-1201606163/>

<sup>367</sup> See, Watercutter, Angela, “To Predict the Next Box Office Hit, Look at Google Movie Trailer Searches,” *Wired*, June 6, 2013, accessed June 6, 2020, <https://www.wired.com/2013/06/google-box-office-prediction/>

<sup>368</sup> Napoli, 2014

Accordingly, the majors developed a suite of partnerships with the new wave of data application companies, cultivating a dynamic of short-term experimental ventures and long-term systemic deals. While studios and networks generally classified these deals as research and development expenses on technological innovation, a holistic investment into the future, they in fact signaled an industry-wide prioritization of the past. Indeed, the majors rationalized the application of data-driven, algorithmic models not as a strategy to chart the new media future, but to preserve the legacy media business of the past, which had experienced fundamental challenges due to a “perfect storm” of disruption that led to steady declines in theatrical, advertising, and home entertainment revenues.<sup>370</sup>

- *New Distribution Platforms*: The rise of new distribution platforms across the digital landscape enabled a growing content surplus, coupled with non-linear viewing schedules (i.e., time-shifting) and format variety (i.e., user-generated content, short-form content) that increasingly pressured the economics of the legacy media business. In particular, cloud-based distribution facilitated a shifting focus on *anything, anytime, anywhere* programming as a viable alternative to time-based theatrical releases and appointment viewing across network and cable television.
- *New Competition*: Distribution platforms such as Netflix, Amazon Prime Video, and YouTube initially maintained an exclusive focus on revenue-sharing, enabling the majors

---

<sup>369</sup> Barnes, Brooks, “Save My Blockbuster,” *New York Times*, June 28, 2013, accessed June 6, 2020, <https://archive.nytimes.com/www.nytimes.com/interactive/2013/06/28/movies/BLOCKBUSTER.html>

<sup>370</sup> Smith & Telang, 2016

to maximize revenue across various digital destinations at set intervals (i.e., windowing), yet eventually moved into programming, effectively emerging as direct competitors to film studios and television networks. Fueled by rich balance sheets, grounded in subscription models (e.g., Netflix's recurring subscription revenue model) and alternative funding sources (e.g., Amazon's cloud business funding its content, Google's search business funding YouTube), competitors increasingly took market share from the majors.

- *Shifting Consumer Behavior*: The rise of new distribution and content platforms created alternative viewing options for audiences (e.g., web video, social media short-form video, mobile video, set-top boxes), which, in turn, contributed to new consumption trends (e.g., binge-watching) and ensured instant access to content beyond the traditional mechanisms of delivery (e.g., theaters, cable subscriptions, physical home entertainment releases).

This confluence of disruptive factors effectively impacted all stages of Hollywood's value chain. In production, film and television slates were cut and financing experienced erratic downturns. At the same time, studios increasingly turned to blockbuster productions in order to balance cost and benefit, effectively minimizing investments in mid-tier productions.<sup>371</sup> In marketing, costs grew as digital advertising mechanisms proliferated and audience attention spans were consumed by a plurality of digital content options, with studios investing heavily through ever-increasing budgets to stand out in a crowded marketplace. In distribution, traditional release windows shrank and fragmented with the infusion of new delivery mechanisms while in-person attendance and appointment viewing declined. In effect, the legacy media business faced considerable obstacles to profit in an industrial environment increasingly marked by pressured foundations,

---

<sup>371</sup> See, Elberse, 2013 and Balio, 2013

shifting dynamics, and overall uncertainty. As a result, film studios and television networks displayed a growing willingness to ground industrial decision-making in data-driven, algorithmically-engineered models to contain uncertainty and increase their success rates.

Data application companies effectively gained traction with the majors by offering an additive approach to risk mitigation, designed to reduce uncertainty, elevate probability of success, and increase profitability through a data-driven and algorithmic business model, deploying numbers-driven automation to preserve the legacy media business. This turn toward rationalization was not an instant pivot, but a long and winding process driven by ongoing experimentation under ever-changing circumstances. It effectively raised Hollywood's technological profile, evolving the legacy media industry's long-held status as an impressionistic institution driven by intuition and instinct, and underscoring the legacy media industry's ability to adapt and negotiate industrial shifts. As Derek Thompson noted succinctly, "Hollywood, like other entertainment industries in the era of big data, is better than ever at figuring out how to give audiences exactly what we say we want."<sup>372</sup>

At the same time, the majors' approach to data application remained largely experimental and incremental. Indeed, studios and networks did not holistically replace manual processes with automated solutions. Rather, they utilized data-driven, algorithmic models as an overlay, an additive feature to pressure-test, enhance, and validate legacy processes, thereby effectively preserving the underlying traditional dynamics of the legacy media business. In this sense, the majors negotiated the integration of existing and emerging solutions, tools, and practices as a way to sustain, secure, and safeguard, rather than disrupt, the way content is produced, marketed, and distributed. The implementation of data application solutions manifested across an industrial

---

<sup>372</sup> Thompson, 2014

spectrum, from one-time experiments to sustaining infrastructure changes, all dependent upon various factors, from the buy-in of senior leadership to budget and financial dynamics, technological capabilities of legacy media organizations, and perspectives of employees.<sup>373</sup>

Overall, Hollywood's growing engagement with data application companies was grounded in an industrial strategy to overcome uncertainty in an increasingly shifting media environment. Film studios and television networks took steps to apply data-driven, algorithmically-engineered decisions to the media and entertainment value chain, implementing a range of data application solutions that reconfigured the process of cultural production while working to preserve its foundational tenets and associated monetization dynamics.

Accordingly, this chapter examines the majors' key data application strategies to preserve the legacy media business, applying data-driven algorithms to key processes of cultural production, incrementally informing how content is produced, marketed, and distributed in the legacy media industry.<sup>374</sup>

- *Data-Driven Production:* The majors partnered with data application companies to automate the selection of ideas, stories, and talent that would increase profitability of film and television programming. In effect, the majors incorporated a layer of data and algorithms into production dynamics.
- *Data-Driven Marketing:* The majors engaged data application models to optimize the impact of marketing campaign outreach, leveraging data-driven algorithms to reconfigure

---

<sup>373</sup> For an example of data-driven organizational changes in the legacy media industry, see, Knowledge@Wharton, "How Viacom Sparked Its Digital and Cultural Transformation," *Wharton School of Business*, February 28, 2020, accessed June 6, 2020, <https://knowledge.wharton.upenn.edu/article/viacom-sparked-digital-cultural-transformation/>

<sup>374</sup> The goal of this chapter is to document and analyze the operational dynamics of data application solutions, not to prove their viability in actually optimizing the business outcome of the legacy media industry.

audience segmentation and creative production practices in an effort to optimize return on investment.

- *Data-Driven Distribution*: The majors drew on data application as a way to increase movie-going turnout and television tune-in during the traditional release window, effectively working to expand established distribution dynamics, including theatrical windows and linear broadcasting times.

Together, these industrial developments signaled the growing application of data to preserve the legacy media business, its approach to cultural production, associated business model, and underlying dynamics, while, at the same time, illustrating the challenges inherent in applying data as a risk mitigator.

### **Data-Driven Production: Validating Investments**

In 2013, Netflix and Amazon Studios released their first original series, *House of Cards* (2013-2018) and *Alpha House* (2013-2014), effectively evolving from content distributors into dedicated programmers. While the shows were made independently from one another, with different production and creative teams involved, they shared a set of distinct similarities. For one, both shows were political dramas set in Washington D.C. and starring prominent actors with notable credits. Furthermore, they were produced by technology platforms seeking to vertically integrate into the media and entertainment business by owning production and distribution capabilities. They also were ready-made for streaming, releasing all of their episodes directly to consumers, accessible anytime and anywhere, enabling seamless viewing, unencumbered by

windows and time slots. And, most notably, they were made in an unconventional way, reflecting a differentiated take on Hollywood's traditional production process.

Yet, despite their commonalities, they were received and discussed differently. *House of Cards* drew significantly more cultural and industrial attention. The show's debut season was the first web television series to receive awards recognition, garnering four Emmy nominations, including one win for best direction. It received critical acclaim from critics and industry insiders.<sup>375</sup> Its production process, meanwhile, the way it found its way onto the screen, has effectively become Hollywood lore. Netflix reportedly utilized its massive repository of proprietary viewer data, along with custom-developed algorithms, to rationalize the investment in the show. Accordingly, the work of data and algorithms enabled the company to offer an unprecedented two-season, \$100 million-dollar commitment to the *House of Cards* producers and creative team and forego any conventional pilot testing while granting unlimited, note-free creative freedom, thereby outbidding established cable and network impresarios and effectively becoming an upper-echelon player in the television production business.<sup>376</sup> Netflix's approach to validating an investment idea based on data-driven algorithms became an emblem of technological innovation and disruption, the de facto use case of revolutionizing the way television is made.

Amazon's take on *Alpha House* did not engage in revolutionary tactics to break into the legacy media business. While the company equally invested in data and algorithms to validate its marquee idea, it developed an approach that emphasized continuity and variation over disruption,

---

<sup>375</sup> See, Dietz, Jason, "House of Cards: Reviews for the Complete 1st Season," *Metacritic*, February 20, 2013, accessed June 6, 2020, <https://www.metacritic.com/feature/house-of-cards-netflix-full-season-1-reviews>

<sup>376</sup> See, Burroughs, Benjamin, "House of Netflix: Streaming media and digital lore," *Popular Communication, The International Journal of Media and Culture*, Volume 17, Issue 1, 2019.



not revolutionizing, but reconfiguring the way television is made. Indeed, Amazon applied data-driven algorithms to automate key elements of the traditional pilot process. Whereas Netflix predicted demand up front, training an algorithm on its data to determine if *House of Cards* could *work* on its platform, Amazon followed the traditional production process, adding software automation, data scale, and algorithmic logic in order to determine a success scenario and then make it *work*.

While both Netflix and Amazon implemented data application strategies in order to predict demand, ensure a receptive audience, and thereby validate production investments, their approaches effectively differed in practice. Netflix worked to change the television production system to gain a competitive advantage over the incumbents, removing the pilot process from the industrial equation to offer creative teams unprecedented control. Amazon, meanwhile, worked to enhance the system to gain entry into the legacy media industry, aiming to advance, improve, and optimize the pilot process to give producers and creative teams more certainty. In this regard, Amazon's application of data effectively worked to preserve core tenets of the legacy media business.

Amazon developed, produced, and released 14 new series pilots on its video platform, culling ideas from a wide range of submissions. The company made the content available to over 250 million users across its website, mobile app, and IMDb platform in order to “get customer feedback.”<sup>377</sup> While Amazon effectively followed the traditional pilot development system, which typically solicited feedback by way of focus groups with up to 50 people, the company reconfigured the process by incorporating a set of differentiators. First, Amazon expanded the scope of the focus group audience, moving from a small representative sample to a large-scale

---

<sup>377</sup> Farber, Dan, “Amazon Studios debuts 14 pilots for free viewing,” *Cnet*, April 19, 2013, accessed June 6, 2020, <https://www.cnet.com/news/amazon-studios-debuts-14-pilots-for-free-viewing/>

digital audience, thus significantly increasing the available data volume on audience feedback. Indeed, over 1 million Amazon users reportedly watched one or more of the available pilots across Amazon's video ecosystem.<sup>378</sup> Second, Amazon increased the depth of audience feedback, considering a broader set of quantitative data points, beyond the conventional focus group questions, including video views (i.e., the number of times a pilot episode was viewed), viewing behavior (i.e., the number of viewers who watched an episode more than once, the number of viewers who watched more than one pilot episode, the moment at which viewers stopped watching), the average user rating (i.e., feedback ranging from 0-5 stars), reviews (i.e., the number of viewers who wrote reviews in addition to ratings), and the share rate (i.e., the number of times an episode was shared across Amazon's platforms). And, third, it automated the collection, consolidation, and communication of audience feedback, distilling large amounts of information into readily accessible insights for the Amazon Studios leadership team.<sup>379</sup>

Amazon's use of data application effectively reconfigured Hollywood's traditional approach to cultural production, leveraging data to update the conventional pilot development process in distinct ways:

- *From relying on sample-size focus groups to incorporating digital information scale:*

Amazon signaled the use of digital information as a practical indicator of future demand, prioritizing data volume and depth across its multi-platform ecosystem.

---

<sup>378</sup> See, Sharma, Amol, "Amazon Mines Its Data Trove to Bet on TV's Next Hit," *Wall Street Journal*, November 1, 2013, accessed June 6, 2020, <https://www.wsj.com/articles/amazon-mines-its-data-trove-to-bet-on-tv8217s-next-hit-1383361270>

<sup>379</sup> See, Sharma, 2013

- *From focusing on small feedback samples to considering a variety of digital metrics:* Amazon illustrated a broader framework to classify audience feedback, including viewing behavior (e.g., video views, watch time) and content engagement (e.g., share rate, reviews).
- *From reviewing feedback manually to generating insights automatically:* Amazon exemplified the process of distilling massive amounts of information into actionable, business-focused insights that complement established creative parameters (e.g., the greenlighting process) in guiding corporate decision-making.

Similar to Netflix, Amazon's approach to data application derived from an internal ambition to evolve Hollywood's traditional practice of impressionistic and insular decision-making. In particular, Amazon sought to prioritize the voice of the customer and expand beyond the perspective of studio tastemakers to gain a more holistic understanding of the audience. As Bill Carr, Amazon's former Vice President of Digital Video and Music, explained succinctly (my emphasis): "**We've always operated in a way where we let the data drive what to put in front of customers. We don't have tastemakers deciding what our customers should read, listen to and watch.**"<sup>380</sup>

Amazon effectively applied its data repository and algorithmic principles as an overlay to the traditional television greenlighting process, seeing an opportunity to optimize the system by ensuring a built-in audience on its platform for shows that receive a season commitment. The result is a more holistic picture of the audience, driven by tangible insight into customer demand. As the former head of Amazon Studios noted, "You can just see what's working on the site, and

---

<sup>380</sup> Sharma, 2013

then that gives you a little context, you know? That *these things* are resonating."<sup>381</sup> At the same time, Amazon's leadership team emphasized that data-driven and algorithmic automation is not a substitute for the creative process, but rather an underlying check system to test, evaluate, and substantiate institutional rationales. "It's not like you can come in on Tuesday and the computer says: 'Doot, doot, doot. Here are the shows you are going to do' It's not 'The Barefoot Executive.' You have to use some judgment as well."<sup>382</sup>

Overall, the production history of *Alpha House* illustrates the impact of data application on traditional production dynamics. Amazon maintained the tenets of the legacy media industry's development process, producing a pilot episode for review, yet added an additional layer of automation, applying incremental data points and algorithmic logic to "take some of the guesswork out of the creative process"<sup>383</sup> and increase the overall probability of success. Indeed, by preserving the pilot format, Amazon was effectively able to validate the entire show, not just the vision, from concept idea to casting and creative execution. While Amazon's approach drew resistance from industry executives, creatives, and pundits raising concerns about premature computationally-manufactured judgment, effectively removing a show's ability to gain an audience over time, and *Alpha House* getting cancelled prematurely after two seasons, the company's investment nevertheless reflected an increasingly viable formula for data-driven production in Hollywood.

Accordingly, the majors displayed a growing tendency to emulate, or at least experiment with, data-driven production models by partnering with a range of data application companies

---

<sup>381</sup> Bishop, 2013

<sup>382</sup> Sharma, 2013

<sup>383</sup> Ibid

across the media and entertainment value chain. While platforms like Amazon leveraged their digital ecosystem of capabilities to manage and control all aspects of the production process (i.e., access to online audiences, interactive user features, distribution platforms, expansive data repositories), film studios and television networks primarily relied on external service providers to validate production investments incrementally. In particular, the majors prioritized three data application strategies to determine demand and profitability in production:

- *Idea/Concept Validation*: The majors utilized data application software solutions to forecast demand and estimate profitability as a way to validate ideas and concepts.
- *Script Validation*: The majors used data application software solutions to evaluate the commercial viability of film and television scripts during the greenlighting process.
- *Talent Validation*: The majors drew on data application software solutions to identify talent that would increase demand and profitability of a production.

By investing in data-driven production dynamics, the majors increasingly made algorithmic use of digital information in an attempt to mitigate risk, control costs, and negotiate uncertainty in the new media environment.

### Idea/Concept Validation

The majors developed partnerships with a suite of data application companies to validate the viability of new programming ideas and concepts at two distinct stages of the production process. One, they drew on data application software solutions during the *pre-production stage*, evaluating concepts and pitches to determine a production greenlight. Two, they leveraged data

application software to estimate demand and profitability for a completed film or television show at the *final production stage*, prior to release, determining any outstanding production decisions (e.g., edits, re-writes and re-shoots, release schedule updates) (table 3.2).

**Table 3.2**

<b>Data Application Production Investments</b>		
<i>Value Chain</i>	<b>Pre-Production Stage</b>	<b>Final Production Stage</b>
<i>Goal</i>	<ul style="list-style-type: none"> <li>● Forecast demand / profitability up front</li> <li>● Validate production investment</li> <li>● Evaluate greenlighting decision</li> </ul>	<ul style="list-style-type: none"> <li>● Forecast demand / profitability months and weeks before release</li> <li>● Evaluate production and release decisions</li> </ul>
<i>Use Case</i>	<ul style="list-style-type: none"> <li>● Generate box office / ratings estimates</li> <li>● Determine demand across regions and audience segments</li> </ul>	<ul style="list-style-type: none"> <li>● Generate box office / ratings estimates</li> <li>● Determine demand across regions and audience segments</li> <li>● Rationalize editing / production / release updates to increase demand and profitability</li> </ul>

Data application companies developed algorithmic models that drew on a wide range of digital information in order to validate production investments, effectively collecting and consolidating data across the digital landscape, including websites, social media, video platforms, and mobile apps. In particular, they utilized data-driven algorithms to identify and analyze leading indicators of audience demand and, by extension, content profitability:

- *Audience Scale*: Audience scale can be characterized as the total number of potential viewers for a film or television show across the global digital landscape.
- *Audience Engagement*: Audience engagement comprises the number of actions taken by audiences to express varying levels of interest in a film or television show, from liking

social media posts to viewing trailers, sharing promotional content, and paying for consumption of related content.

- *Audience Reception*: Audience reception can be characterized as general feedback on related content (i.e., reviews and ratings) as well as initial interest in a specific film or television show.

These indicative data sources can be filtered to reveal more detail and granularity in determining demand across a sliding scale. For example, audience scale can be segmented by country and demographic; engagement can be defined by purchase (i.e., digital purchase or rental of a film or television show), video views, likes/shares/comments, and downloads; reception can be classified by reviews (e.g., Rotten Tomatoes scores and IMDb ratings) and sentiment (e.g., positive or negative sentiment across social media conversations) (table 3.3).

**Table 3.3**

<b>Demand and Profitability Indicators</b>			
<i>Indicators</i>	<b>Audience Scale</b>	<b>Audience Engagement</b>	<b>Audience Reception</b>
<i>Metrics</i>	Total # of potential viewers	Total # of actions taken	Feedback and sentiment
<i>Data Types</i>	<ul style="list-style-type: none"> <li>● # of subscribers</li> <li>● # of viewers</li> <li>● # of followers</li> </ul>	<ul style="list-style-type: none"> <li>● # of views</li> <li>● # of likes</li> <li>● # of shares</li> <li>● # of comments</li> </ul>	<ul style="list-style-type: none"> <li>● Review ratings</li> <li>● Social media sentiment</li> </ul>
<i>Examples</i>	<ul style="list-style-type: none"> <li>● OTT subscribers (i.e., Netflix, Amazon, Hulu)</li> <li>● Video subscribers (i.e., YouTube)</li> <li>● Social media followers (i.e., Facebook, Twitter)</li> </ul>	<ul style="list-style-type: none"> <li>● Video views (i.e., Netflix, Amazon, Hulu, YouTube)</li> <li>● Facebook likes, YouTube thumbs-up</li> <li>● Re-tweets</li> </ul>	<ul style="list-style-type: none"> <li>● Rotten Tomatoes Score</li> <li>● IMDb rating</li> </ul>

Data application companies further applied algorithms to weigh data indicators by impact, ensuring that information would be representative and contextual in the model output. For example, demand is weighted differently in smaller countries with less developed digital infrastructures than in fully developed countries with a high penetration of digital services to account for the difference in scale. Similarly, digital content purchases are considered to be more indicative of audience demand for content than likes across social media accounts given that a purchase shows a clear willingness to own, watch, and engage with content.

Additionally, data application companies incorporated media and entertainment industry metadata into their algorithmic model, prioritizing information sets specifically associated with film and television programming in order to add industry-specific context. In particular, they focused on holistic data points, commonly shared across film and television programming, such as genre, cast, creative team (e.g., directors, producers), type (e.g., theatrical or network television content vs. direct-to-streaming), release year, and rating, among others.

The focus on these multidimensional indicators enabled data application companies to algorithmically filter demand data at various levels, thereby offering a range of scenarios to contextualize model output within the production process. Parrot Analytics, an analytics company based in Los Angeles, provides an instructive use case in this context. The company's patented data application solution enables various *demand views* across a suite of dashboard outputs. It tracks global demand and can segment by country; it can filter metadata to reveal elements that drive demand (e.g., genre); it pinpoints demand at various stages of the production process (e.g., pre-release); it reveals how demand evolves over time; it highlights key indicators



that affect demand; and it compares demand against competitors or similar titles.<sup>384</sup> Effectively, the company algorithmically tracks, analyzes, and expresses demand for films and television shows along the production lifecycle.<sup>385</sup>

While data application solutions enabled granular representations of demand, relying on data visualization and dashboard aesthetics to express layers of insights that informed business decisions, their main model output generally manifested as a quantitative score. For example, Piedmont Media Research promoted a “Cumulative Engagement Score,” that measured the correlation between a film release and its anticipated box office performance. Ranging from 1 to 1,000, the score effectively delivered a statistical reference to estimate financial performance, allowing studios to draw comparisons to previous titles or competitors.<sup>386</sup> By relying on quantitative scores, data application companies thus provided an easily digestible and comparable benchmark for the legacy media industry to validate investment ideas.

### Script Validation

The majors drew on data application solutions to examine, evaluate, and elevate the commercial viability of scripts during the greenlighting process. While the reliance on algorithmic automation in script evaluation introduced technological efficiencies, it represented a continuation of the legacy process. Indeed, Hollywood has long drawn on readers to classify scripts and provide coverage based on a set of inputs in order to determine the potential success

---

<sup>384</sup> See, Wallenstein, Andrew, “Netflix in Japan: Tepid Demand for Originals, Data Indicates,” *Variety*, September 1, 2015, accessed June 6, 2020, <https://variety.com/2015/data/news/tepid-demand-for-originals-on-netflix-in-japan-1201581943/>

<sup>385</sup> Seger, Wared, “Parrot Analytics: Understanding Audience Demand for TV Shows - Variety Big Data 2016,” *Parrot Analytics*, December 13, 2016.

<sup>386</sup> See, Avirgan, 2015

of a film.<sup>387</sup> Data application solutions followed the coverage process in considering a wide range of industry metadata in their analysis of films' financial performance, including character, plot, themes, genre, and length.<sup>388</sup> At the same time, they triangulated additional industry data, adding information on the production and financials into the evaluative mix. As such, data application companies assigned measurable commercial value to script coverage.

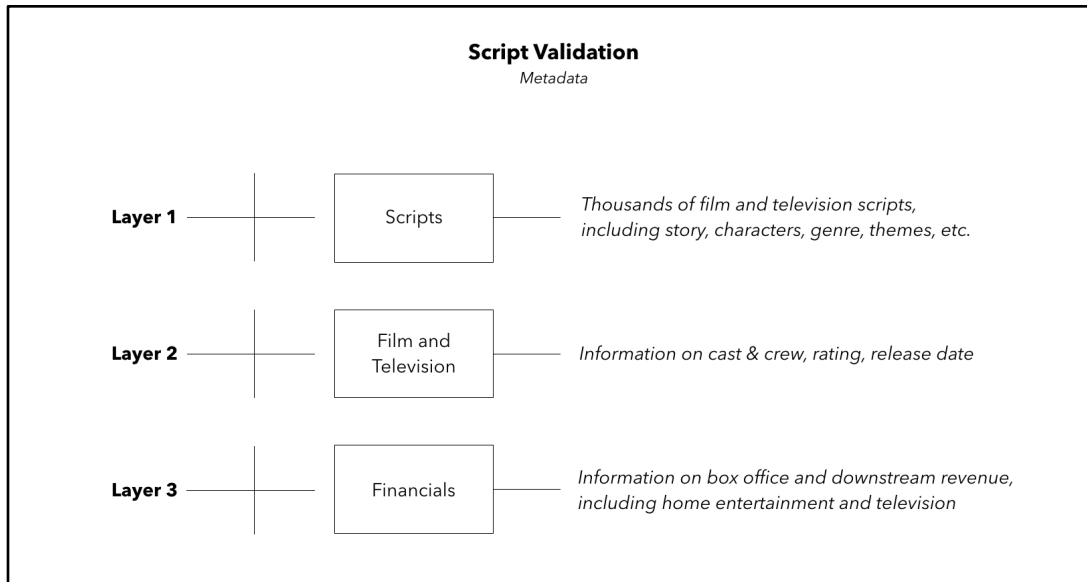
To validate scripts, data application companies accordingly developed algorithms trained specifically on three sets of industry metadata. First, they built repositories of scripts, consolidating, classifying, and categorizing a wide range of information on storytelling. For example, Scriptbook, a Belgian company, built a library of over 30,000 scripts.<sup>389</sup> Second, they collected data on the finished films and television shows based on the scripts, including cast and crew, rating, and release date. Third, they incorporated data on the financial performance of scripts. As such, they created a comprehensive information set, detailing and quantifying key layers of a script (figure 3.8).

---

<sup>387</sup> See, Ross, Alexander G., "Creative decision making within the contemporary Hollywood studios," *Journal of Screenwriting*, Volume 2, Number 1, 2011.

<sup>388</sup> See Barnes, 2013, "Solving Equation for a Hit Film Script, with Data"

<sup>389</sup> See Scriptbook website, <https://www.scriptbook.io>



**Figure 3.8:** The process of script validation applies various layers of metadata, which includes information gathered from scripts, related information on the completed film (e.g., cast, genre), and the associated financials.

Data application solutions analyzed the layers to determine the commercial viability of a script, generating various box office scenarios based on variable inputs. For example, the financial performance of a film could be adjusted by changing the ending or altering the location of a pivotal scene (i.e., the script layer), re-casting the lead actor (i.e., the production layer), or reducing the budget (i.e., the financial layer). This type of algorithmic exercise resembles a plug-and-play model, which automatically generates new script options based on prevailing industry conventions. While the approach created efficiencies via automation, with data application companies frequently citing the quick turnaround of script coverage, it equally streamlined variety, effectively prompting the model of a successful script to “regress to the mean.”<sup>390</sup>

Overall, data application solutions thus worked to reinforce reigning script formulas, correlating conventional practice with box office outcome. While the application of data-driven

<sup>390</sup> Boone, Christopher, “Paint by Numbers? Hollywood Rewriting Scripts Based on Statistical Analysis to Boost Box Office,” *No Film School*, May 7, 2013, accessed June 6, 2020, <https://nofilmschool.com/2013/05/hollywood-rewrite-script-statistical-analysis-box-office>

algorithms consistently predicted demand and profitability of scripts at a higher accuracy rate than industry insiders,<sup>391</sup> giving data application companies an air of oracle-like wisdom, their overall reliability remained limited. Data application companies train algorithms based on historical data, which lacks contextual detail. As such, data application is largely driven by past audience behaviors and industrial conventions. Accordingly, the approach does not take account of future developments, such as cultural shifts, new modes of production, or industrial outliers, effectively limiting its operational scope. As one technology executive noted, “[T]he data just tells you what happened in the past. It doesn’t tell you anything that will happen in the future.”<sup>392</sup>

In this sense, script validation operated as an extension of traditional legacy media practices, leveraging data-driven algorithms to introduce automation and predictive models to substantiate creative investments, while highlighting distinct limitations of data application within the creative environment of media and entertainment.

### Talent Validation

As part of their overarching service offerings within the media and entertainment value chain, data application companies promoted solutions designed to optimize the packaging of programming. In particular, data application solutions increasingly informed casting decision dynamics, enabling the majors to validate the impact of specific actors on overall content demand and profitability. As such, they focused on validating audience demand for the cast of films and television shows, in three principal ways. One, they estimated quantitative demand for a particular actor or actress based on overall audience size and engagement. Two, they

---

<sup>391</sup> See, Vincent, 2019

<sup>392</sup> Adalian, 2018

determined the overall audience perception based on qualitative input, filtering actors' public image and likability into demand predictions. And, three, they measured actors' financial track records to determine their role in making content succeed.

- *Audience Size and Engagement*: Data application companies scanned the digital landscape to compile the number of an actor's online followers (i.e., the number of digital users who receive automatic updates from a digital channel, a social media account, or other platform mechanisms) as well as the engagement rate (i.e., the number of followers who actively interact with the actor's digital content on a regular basis).
- *Audience Perception*: Data application companies measured how an actor is perceived by audiences across the digital landscape, consolidating critical feedback and user comments to determine a likability score.
- *Commercial Value*: Data application companies compiled information on actors' box office history to develop a holistic overview of financial performance.

To validate audience size and engagement, data application companies trained algorithms on various data sources, specifically social media followers across major platforms and subscription-based services (e.g., blogs, podcasts), given their global scale and interactive nature. For example, theAudience, a social media publishing and analytics company, consolidated actors' social media presence by platform to illustrate their online following and associated strength in commanding digital attention as a leading indicator of demand.<sup>393</sup>

---

<sup>393</sup> See, Hod, Itay, "How Hollywood Actors' Twitter Followings Have Become as Important as Talent," *The Wrap*, March 10, 2015, accessed June 6, 2020, <https://www.thewrap.com/how-hollywood-actors-twitter-followings-have-become-as-important-as-talent/>

Data application companies utilized solutions that tracked, measured, and analyzed online conversations to determine positive or negative sentiment, cross-referencing the results with an actor's critical review track record (e.g., Rotten Tomatoes, IMDb) to project actors' ability to drive digital conversation and support demand and profitability measures for new content.<sup>394</sup>

To validate commercial value, data application companies deployed a system that aggregates actors' box office histories and, using a predictive algorithm, enables the forecast of future box office performance based on varying input of actors' financial information. As such, data application allowed the majors to evaluate an actor's impact on box office and downstream channels (i.e., home entertainment). For example, Cinelytic, a Los Angeles-based startup, developed a solution to model a film's entire commercial lifecycle (i.e., theatrical, in-home, and television) by evaluating actors' financial performance history.<sup>395</sup> Meanwhile, Vault, a startup based in Los Angeles and Tel Aviv, enabled the tracking, classification, and categorization of actors' financial performance over time, outlining the economics of programming.<sup>396</sup>

Effectively, data application companies validated actors' online presence based on three key criteria. First, they measured actors' in-built audience and engagement to determine an addressable audience for the film or television show, thereby modelling potential ticket sales or tune-in rates. Second, they examined actors' public image among audiences as a contributing

---

<sup>394</sup> See, Barnes, Brooks, "A-Listers, Meet Your Online Megaphone," *New York Times*, November 10, 2012, accessed June 6, 2020, <https://www.nytimes.com/2012/11/11/business/oliver-luckett-of-theaudience-building-online-fan-bases.html>

<sup>395</sup> See, Vincent, 2019

<sup>396</sup> McAlone, Nathan, "This startup uses artificial intelligence to predict whether a Hollywood film will be a hit or a flop — just by scanning the script," *Business Insider*, July 29, 2015, accessed June 6, 2020, <https://www.businessinsider.com/this-startup-uses-artificial-intelligence-to-tell-whether-a-hollywood-film-will-be-a-hit-or-a-flop-just-from-the-script-2015-7>

factor for demand and profitability. And, third, they considered the financial performance histories of actors to enable comparative benchmarking.

Talent validation manifested as a recurring industrial practice, with actors' online presence increasingly functioning as a yardstick to measure mainstream appeal, indicating a tangible ability to promote, open, and sustain new films and television shows. As a veteran casting director noted, "If it came down to two professional actors, one of whom had great visibility in social media and one who was barely recognizable, we'd go with the one who could get the numbers."<sup>397</sup>

The practical approach to data application in the casting process can take different forms. For example, productions looking to unlock a broad digital audience in order to promote digital word-of-mouth might exclusively focus on the scale of actors' online audience, as in the case of SyFy's *Sharknado* series (2013-2018), which deliberately cast talent with large social media followings in order to boost ratings.<sup>398</sup> In other contexts, the criteria can be more narrow and discrete, focused on specific audience segments or commercial markets. For example, Matt Damon was reportedly cast as the lead in *The Great Wall* (Universal Pictures, 2016), a U.S.-Chinese co-production, because he "over-indexed in web searches in China,"<sup>399</sup> indicating popularity across regional markets based entirely on audience engagement and perception. Overall, casting effectively took the form of a quantitative numbers game, with references to online followings increasingly influencing corporate decision-making.<sup>400</sup>

---

<sup>397</sup> Hod, 2015

<sup>398</sup> Pullen, 2014

<sup>399</sup> Moore, 2016

While the scale of information and automated nature manifested as a new element in the casting process, the conceptual principle of measuring popularity based on audience feedback has long played a role in the legacy media industry, dating back to the use of fan mail as an incremental input in evaluating actors' studio contracts.<sup>401</sup> As such, the use of data application constitutes a continuation of existing casting practices.

As part of its ongoing impact on production dynamics, the practice of data-driven, algorithmic talent validation instituted a vibrant industry discourse, with data application companies collaborating with the trades to provide regular trackers on actors' online presence and digital capabilities. For example, The Hollywood Reporter launched the *Top Actors Social Media Ranking*, "which ranks the most popular actors based on data from Facebook, Instagram, Twitter, YouTube and Google Plus."<sup>402</sup> Additionally, legacy media players partnered with data application companies to incubate new commercial software products. For example, Variety, in partnership with analytics company Viralheat, created *VScore*, a measurement solution covering 17,000 film and television actors, which "tracks social listening across all social channels and news sites and correlates it with an actor's box office and other data across TV and film, [...] the actor's nominations and wins in all major award categories, as well as his or her upcoming projects, [...] adjust[ed] in real time according to the actor's recent roles and media coverage."<sup>403</sup>

---

<sup>400</sup> See, Luna, 2013 and Tryon, Chuck, *On-Demand Culture: Digital Delivery and the Future of Movies* (New Brunswick: Rutgers University Press, 2013).

<sup>401</sup> See, Ohmer, 2006

<sup>402</sup> Rutherford, Kevin, "Dwayne Johnson Leads Top Actors Social Media Ranking for 10th Week," *Hollywood Reporter*, September 9, 2017, accessed June 6, 2020, <https://www.hollywoodreporter.com/lists/dwayne-johnson-leads-top-actors-social-media-ranking-10th-week-1036720/item/gal-gadot-mvp-9-13-1036339>

<sup>403</sup> Variety Staff, "Variety Launches Vscore to Measure Actors' Value," *Variety*, August 6, 2014, accessed June 6, 2020, <https://variety.com/2014/biz/news/johnny-depp-value-jennifer-lawrence-variety-vsore-1201263164/>



The rise in industrial data application products effectively assigned a new value measure to actors, highlighting and dissecting their digital presence into demand and profitability indicators. As such, the commercial value of actors became increasingly datafied, defined by a range of algorithmically captured data points.

Overall, data application companies worked to reconfigure and enhance the traditional casting process by validating the online presence of talent as an incremental indicator of demand. Indeed, their solutions were primarily applied as a preliminary evaluative measure, delivering a set of insights to inform corporate negotiations as part of the traditional process, rather than supplanting established modes of decision-making. As one executive summarized, “We’re not really at the [casting] table; we’re before the table.”<sup>404</sup> As such, while data application companies generally did not influence the final deal-making output, their input increasingly enabled the majors to make a demand-driven case.

--- --- ---

By partnering with data application companies, the majors effectively automated key stages of the production lifecycle, from development to greenlighting, casting and pre-release, to validate their investments in film and television programming. In particular, the use of data-driven algorithms enabled the majors to predict demand and forecast profitability for films and television shows based on variable inputs, including digital information (e.g., web, social, video, mobile data), industry metadata (e.g., cast and crew, genre, and release date information), and financials (e.g., box office, budgets, downstream revenues). While outcomes and success rates differed, the majors increasingly relied on data application as an incremental way to mitigate risk, control costs, and reduce uncertainty.

---

<sup>404</sup> Pullen, 2014

The integration of data application strategies is not a new industrial practice, but a reconfiguration of legacy production dynamics. The majors have consistently drawn on data sources to validate investments, from internal financial benchmarks (e.g., box office, advertising and home entertainment revenues) to external focus groups and surveys. The use of data-driven algorithms effectively automated this approach, introducing incremental information scale, creating efficiencies, and enabling new comparisons across larger data sets.

While data application companies delivered a suite of new automated software solutions that enabled the majors and other industry stakeholders, from independent production companies to financiers and filmmakers, to realize incremental value, their algorithms neither proved infallible nor guaranteed success. Effectively, they instituted another layer of managing the uncertainty of the legacy media business.

### **Data-Driven Marketing: Programming for the Audience**

In 2014, Warner Bros. rebooted one of the longest-running media franchises in the history of the entertainment business with its take on the pop-culture icon *Godzilla*. Made in partnership with Legendary Pictures, the film's May release window was positioned as a summer movie blockbuster, celebrating the 60th anniversary of *Godzilla*'s first on-screen appearance, which launched a massive media empire counting over thirty films and numerous spin-offs in television, video games, literature, and music, all grounded in a vivid stream of popular fandom. Yet, despite its cultural cachet, historical significance, and in-built fan audience, the film was widely considered a commercial risk in Hollywood, for several reasons. For one, *Godzilla* flaunted a \$160 million-dollar price tag, which put immense financial pressure on the film, with analysts forecasting that it would need to generate close to half a billion dollars in box office

revenue in order to fully recoup its costs.<sup>405</sup> Additionally, as an entry in the monster movie genre catalog, *Godzilla* faced an uphill battle in appealing to mainstream audiences. While Peter Jackson's remake of *King Kong* (Universal Pictures, 2005) had proved moderately successful, other high-budget monster movies did not fare as well. 20th Century Fox's 1998 *Godzilla* adaptation had barely broken even at the box office and more recent iterations, like *Pacific Rim* (Warner Bros., 2013), did not meet the expectations set by their exorbitant budgets, ultimately underperforming at the box office. Despite the global success of comic book and sci-fi fare, monster movies were still considered niche fandom, unlikely to break through the mainstream mold and achieve widespread appeal.<sup>406</sup> Finally, *Godzilla* was seemingly promoted as an ensemble drama, with the lack of bankable stars raising concerns about the film's profitability in a highly competitive marketplace.<sup>407</sup>

Accordingly, industry analysts marked *Godzilla* as the first box office bomb of the year, citing it as another example of the legacy media industry's oft-ascribed paradigmatic overspending on IP, with Forbes going as far as asking the prophetic question, "Is *Godzilla* the John Carter of 2014?,"<sup>408</sup> evoking one of the most dire financial miscalculations in Hollywood history.<sup>409</sup> *Godzilla* was expected to open between \$60-70 million at the domestic box office,

---

<sup>405</sup> See, Lang, Brent, "'Godzilla' Stomps to \$36 Mil Debut in China," *Variety*, June 15, 2014, accessed June 6, 2020, <https://variety.com/2014/film/news/godzilla-china-box-office-1201220922/>

<sup>406</sup> See, Grant, Barry Keith, ed., *Film Genre Reader IV* (Austin: University of Texas Press, 2012).

<sup>407</sup> Khatchatourian, Maane, "'Godzilla' Crushes Box Office With Largest Opening Day of Year, Set for \$98 Million Weekend," *Variety*, May 17, 2014, accessed June 6, 2020, <https://variety.com/2014/film/box-office/godzilla-98-million-weekend-opening-1201184456/>

<sup>408</sup> Furrier, John, "Hollywood: Is 'Godzilla' The 'John Carter' Of 2014?," *Forbes*, January 23, 2014, accessed June 6, 2020, <https://www.forbes.com/sites/siliconangle/2014/01/23/hollywood-is-godzilla-the-john-carter-of-2014/#115c1968301c>

<sup>409</sup> See, Sellers, Michael D., *John Carter and the Gods of Hollywood* (Lichfield, UK: Universal Media, 2012).

indicating that the film would face tremendous obstacles in remaking its budget and achieving profitability, with several trade outlets projecting that it would not cross the 100 million-dollar mark in the U.S..

*Godzilla*'s return to the big screen recast reigning industry dynamics. During its debut weekend, the film opened to \$93.2 million in the U.S., a 40% average increase to the initial measurement benchmark, scoring an additional \$103 million in receipts overseas, which made it the second-highest opening of the year.<sup>410</sup> The film went on to gross \$529.1 million globally, effectively launching a modern expansion of the *Godzilla* franchise that resulted in a suite of sequels, including *Godzilla: King of Monster* (2019) and *Godzilla vs. Kong* (2020), as well as an array of commercial tie-ins. As such, the film effectively defied industry expectations, demonstrating that big-budget monster genre fare could perform on the mainstream stage and be financially profitable for Hollywood. While *Godzilla* accordingly emerged as an oft-cited industrial phenomenon, its success is not solely grounded in quality, but its unique promotional rollout that elevated the film above industry expectations. Indeed, *Godzilla* represents a layered use case of data-driven marketing in Hollywood, a procedural approach that enabled the film to bypass and reconfigure traditional marketing practices along the legacy media value chain.

Hollywood's conventional marketing approach for blockbuster and tentpole productions has consistently followed a core set of tenets.<sup>411</sup> First, major studios effectively seek to maximize turnout and attendance in the first month of release at the box office in order to gain exposure and drive profitability in an increasingly crowded marketplace. The overarching goal is to create

---

<sup>410</sup> See, McClintock, Pamela, "Box Office: 'Godzilla' Opens to Monstrous \$93.2 Million in North America," *Hollywood Reporter*, May 17, 2014, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/box-office-godzilla-opens-monstrous-705168>

<sup>411</sup> See, Kerrigan, Finola, *Film Marketing* (New York: Routledge, 2009).

awareness in order to achieve as much reach as possible, maximizing exposure as a way to capture the attention of a broad set of movie-goers. To this end, marketing departments have primarily relied on television advertising to capitalize on its wide audience reach. While broadcast and cable viewership has incrementally declined over the 2000s,<sup>412</sup> with digital advertising models providing efficient and dynamic alternatives (e.g., social media advertising, online video advertising),<sup>413</sup> the number of television households in the U.S. has largely continued to grow,<sup>414</sup> albeit across older demographics, thereby cementing the widespread belief that the medium remains the most productive advertising mechanism for studios to reach a mainstream audience. As one Hollywood executive noted, television “is the most efficient and still casts the widest net [...] to get enough people.”<sup>415</sup> Accordingly, studios have consistently allocated the majority of their increasingly growing marketing budgets to television, while spending incrementally, and indiscriminately, across all other available channels (i.e., print, out-of-home) to ensure wide exposure. This mass-media approach has increasingly ballooned the cost structure of marketing campaigns as Hollywood expanded its market structure across the world, with the average blockbuster demanding up to \$100 million in promotional costs.

---

<sup>412</sup> Buchanan, Kyle, “How Will Movies Survive The Next 10 Years?,” *New York Times*, June 20, 2019, accessed June 6, 2020, <https://www.nytimes.com/interactive/2019/06/20/movies/movie-industry-future.html>

<sup>413</sup> See, Shields, Ronan, “U.S. Digital Ad Spend Will Surpass Offline in 2019,” *AdWeek*, February 20, 2019, accessed June 6, 2020, <https://www.adweek.com/programmatic/u-s-digital-ad-spend-will-surpass-offline-in-2019/>

<sup>414</sup> See, Nielsen, “Nielsen estimates 120.6 million TV homes in the U.S. for the 2019-2020 TV season,” August 27, 2019, accessed June 5, 2020, <https://www.nielsen.com/us/en/insights/article/2019/nielsen-estimates-120-6-million-tv-homes-in-the-u-s-for-the-2019-202-tv-season/>

<sup>415</sup> McClintock, Pamela, “\$200 Million and Rising: Hollywood Struggles With Soaring Marketing Costs,” *Hollywood Reporter*, July 31, 2014, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/200-million-rising-hollywood-struggles-721818>

Second, studios have long tied the mass-media approach to a four-quadrant audience model, emphasizing the reach of specific audience segments as the key variable in marketing outreach. The model conceptualizes the mainstream Hollywood audience as a suite of four quadrants, organized by gender (i.e., men and women) and age (i.e., over and under 25 years old). Studios have consistently positioned blockbusters as four-quadrant in that they appeal to all core audience segments. The focus on four-quadrant audiences has enabled the studios to effectively allocate marketing budgets across legacy advertising channels, primarily television, in order to sustain broad-based targeting. In effect, the legacy media industry has rationalized broad, cross-platform marketing campaigns as the most efficient way in order to drive audience engagement at scale.

Third, while studios generally create a suite of materials to populate marketing channels with promotional content, they have primarily relied on a small set of trailers (i.e., the teaser, the launch trailer, and the final trailer) in order to engage audiences.<sup>416</sup> The trailer triptych has effectively served as the central foundation of marketing content and is consistently edited, modified, and customized, by channel (e.g., theatrical trailer, television spot, radio ad), theme (e.g., action, romance, suspense, comedy), and audience segment (e.g., men and women over/under 25). Studios have fashioned variations on the same trailer theme, crafting a unified repository of promotional materials.

Finally, studios have long measured audience feedback in order to evaluate the performance of marketing campaigns in the lead-up to a film's release. In particular, they have partnered with legacy research vendors in order to determine audience interest in a film, extrapolate potential attendance, and model ticket sales. Generally referred to as *tracking*, this

---

<sup>416</sup> See, Wyatt, 1994 and Balio, 2013

process has incorporated audience feedback by way of telephone polls as well as online surveys, mixed with historical financial data from benchmark releases. Tracking begins up to three months ahead of the film's release, providing studios with intermittent pulse checks on a film's marketing rollout performance, theoretically as a way to anticipate, implement, and deploy changes to a campaign. In practice, tracking has largely functioned as a financial forecast tool, leaving studios little time to affect marketing changes. As one industry executive characterized the practice's impact on marketing, "[w]ith tracking, there's not much you can do besides a Hail Mary three weeks out."<sup>417</sup>

Hollywood's conventional marketing approach has consistently relied on a key set of steps: allocating increasingly growing marketing budgets to a concentrated set of mainstream marketing channels, with a long-established priority focus on television; promoting a core set of trailers in intervals during the lead-up to release; and targeting a four-quadrant mainstream audience to engage as broad a viewership as possible. Additionally, the legacy media industry has tracked audience feedback in the lead-up to releases, though the manual and static nature of the process has generally not allowed for impactful changes to marketing campaigns, but instead served as a tracking tool to allocate incremental advertising funds. In this regard, studios have long emphasized broad audience reach and manual information-gathering in an effort to maximize exposure to a small set of promotional materials in an effort to drive opening weekend attendance. Effectively, the industry has practiced a "spray and pray" model, broadly distributing content across a variety of channels, and relying on exponential exposure for results.<sup>418</sup>

---

<sup>417</sup> Barnes, 2014, "Hollywood Tracks Social Media Chatter to Target Hit Films"

<sup>418</sup> See, Kringsman, 2018

*Godzilla* reconfigured key elements of Hollywood’s conventional marketing approach by applying data holistically across the marketing value chain in order to create a more dynamic, responsive, and focused promotional process. Led by Legendary Pictures’ in-house analytics division, the film’s marketing outreach was driven by data-driven insights and algorithmically-based practices. First, Legendary reconceptualized the notion of the mainstream blockbuster audience, moving from a general rank of four-quadrant movie-goers (i.e., the top-down model) to a more focused set of viewers specifically interested in seeing *Godzilla*, or displaying interest in key elements associated with *Godzilla*, such as genre, tone, or the cast (i.e., the bottom-up model). By drawing on a vast data repository, culled from the digital landscape and industry metadata, Legendary created a proprietary algorithmic system designed to identify a focused set of audience segments for *Godzilla*, composed of two major groups:

- *Fans*: This group comprised movie-goers keen on seeing the film in theaters. Fans generally display high degrees of familiarity, awareness, and interest. For *Godzilla*, the fan segment can be broken down into a multitude of layers, including fans of the character, fans of previous films, fans of related media spin-offs, fans of the overall IP and franchise, as well as fans of the monster movie genre, among others.
- *Persuadables*: This group referred to a wide range of audience members who are “open-minded about being convinced to see the film.”<sup>419</sup> As such, persuadable audience segments may not have the same level of interest and familiarity as fans, yet they display

---

<sup>419</sup> Gaudiosi, 2014



willingness to see a film if presented with the right incentive (i.e., marketing materials that appeal to their interests).<sup>420</sup>

Legendary's analytics team thus effectively shifted focus from a broad set of general movie-goers to a focused segmentation of dedicated and potential *Godzilla* movie-goers. Indeed, the company programmed this segmentation horizontally, creating various micro-segments within both groups.<sup>421</sup>

Legendary conceptualized the *Godzilla* fan as a dynamic group of individuals with varying interest in the franchise, which enabled the company to court long-time fans of the IP while engaging emerging fan segments with interest in monster movies, world-building, and sci-fi action. Additionally, Legendary created various layers of persuadable audience segments, identifying key interest drivers that would incentivize them to see the film. Legendary's Chief Analytics Officer described the process as assigning a numerical value to determine the degree of persuadability (my emphasis):

“The first step we'll do is take that persuadable audience and define them exclusive of the givens and the nevers. Then, within the persuadable audience, **we will effectively score every single person**. In the U.S., for a movie of a scale we typically would work on, it could be 40 million or 50 million people. **They'll get a score from zero to 100. 100 being very likely, zero being very unlikely.**”<sup>422</sup>

---

<sup>420</sup> See, Moore, 2016

<sup>421</sup> Marolda, 2014, “Changing Hollywood Paradigms with Analytics”

<sup>422</sup> Krigsman, 2018

Legendary effectively applied data to expand the concept of audience segments, enabling the company to focus its marketing efforts on a core set of fans, coupled with an incremental set of persuadable movie-goers, which allowed for a more focused and efficient approach.

Based on this new segmentation model, Legendary created a dynamic marketing campaign rollout. For fans, the company positioned its marketing within a relevant discursive framework, targeting cultural events such as *Comic-Con* as the catalyst to create resonance with the audience. For persuadables, Legendary found that women aged 25-34 would be open to seeing the film, among others, effectively providing another addressable audience segment to target with custom marketing materials across a dedicated set of channels. In this regard, Legendary's application of data enabled the company to adjust the broad mass media marketing approach with an automated layer of audience segmentation, shifting from a broad to a focused audience target.

Second, the Legendary team applied data in order to consistently maintain an understanding of audience reception in the lead-up to the film's release and adjust marketing outreach accordingly. To this end, the team deployed analytics software solutions to regularly measure audience feedback on the film, effectively tracking reactions and sentiment on the film's release and associated marketing materials across a wide range of digital channels to measure the audience's reactions and further identify popular characters, themes, and scenes to spotlight in the ongoing rollout. While studios typically hold a suite of test screenings across core markets prior to a film's release, a manual process that involves written feedback cards designed to pinpoint what audiences liked and disliked, Legendary automated the process by collecting and analyzing information across the digital landscape, enabling the company to dynamically track and respond to audience feedback on a daily basis as opposed to a limited set of intervals. By

establishing an ongoing feedback loop with the audience, Legendary thus gathered a stream of insights, establishing a dynamic capability to refine its marketing output, effectively creating promotional materials tailored to the preferences of fans and persuadable audience segments. While studios generally crafted trailers for a four-quadrant audience ahead of time, considering only minimal adjustments based on research input, Legendary created, curated, and customized a suite of materials for its key audience segments, indeed generating “hundreds or thousands of combinations of subsegments and creative”<sup>423</sup> to be deployed across the digital landscape.

Finally, Legendary drew on audience insights to optimize its marketing spend, effectively determining the marketing and advertising channels most suited for reaching its key audience segments. Founder and CEO Thomas Tull cited an opportunity to spend marketing capital more efficiently, noting that the legacy system’s mass media approach lacked context and detail, which could lead to a mismatch between a film’s marketing materials and the marketing channel’s associated audience. In an anecdote on television-based film marketing, he noted, “[y]ou see a *Dark Knight* ad on the *Golden Girls* ... that may not be the best use of capital.”<sup>424</sup> Furthermore, he observed that mass media marketing would allocate spend across channels that were not relevant to a film’s core audience focus: “We got a seven million dollar bill on one of the *Hangover* movies for newspaper display ads and I thought that’s probably not a great or efficient use of capital.”<sup>425</sup>

Accordingly, the company allocated the budget of its marketing mix to channels that over-indexed on its target audience segments’ behaviors and interests, while simultaneously

---

<sup>423</sup> Ibid

<sup>424</sup> Knowledge Project Podcast, 2019

<sup>425</sup> Ibid

removing under-indexing channels. As a result, Legendary increased the efficiency of its marketing outreach, ensuring that marketing materials would reach an audience that was already primed, or open-minded, to see *Godzilla*, which helped the company reduce its marketing spend by up to 15%, while simultaneously increasing return on investment at the box office.<sup>426</sup> While studios prioritized television advertising, and spreading budgets across all available channels through a “spray and pray” approach, Legendary emphasized efficiencies by focusing on channels that matched the behavior and interest of its key audience segments.

*Godzilla* illustrates how Legendary reconfigured Hollywood’s conventional marketing approach, not replacing, but automating key elements of audience engagement:

- *From gathering audience feedback manually and sporadically to automating insights consistently:* Legendary leveraged data application software to collect, consolidate, and conform audience feedback on a regular basis, enabling the company to dynamically respond to new insights. Studios previously relied on static, limited, and time-consuming information samples.
- *From focusing on broad audience quadrants to prioritizing select micro-segments:* Legendary applied data to identify and classify audience segments by their likelihood to see a film, grouping fans, persuadables, and nevers (i.e., general movie-goers not interested in seeing the film) into categories to develop a focused and addressable audience overview. Studios previously operated in broad strokes, focusing on an audience’s overall presence across channels rather than level of interest.

---

<sup>426</sup> Ibid

- *From following a “spray and pray” model to using a more efficient approach of engagement:* Legendary distributed promotional materials across select channels in order to incentivize audiences to see a film. Studios previously allocated marketing budgets primarily to television, segmenting by TV channels’s audience profiles rather than audience behavior and interest.

Legendary’s data-driven marketing was grounded in a growing availability of information on content and audience, enabling film and television producers to increasingly program marketing for the audience, customizing outreach based on audience behavior and varying interest levels, while at the same time optimizing efficiencies to maximize the impact of marketing channels and thereby reduce overall marketing budget spend. As Thomas Tull explained, “[t]here’s more information available today than has ever been available in terms of people putting their preferences and all kinds of information freely up online. We want to take advantage of that and be much more efficient about the way we run our business.”<sup>427</sup>

While Legendary’s marketing model appeared de novo, its reliance on data-driven and algorithmic considerations to optimize audience engagement reflected a broader trend towards managing costs, creating efficiencies, and maximizing returns in Hollywood’s high-stakes culture of high-budget investments. While marketing has remained “the single most discussed and debated issue in Hollywood,”<sup>428</sup> the increasing presence of data application solutions signaled an attempt to make campaigns more cost-effective and efficient. To this end, the legacy media industry increasingly incorporated data-driven marketing strategies to reconfigure the

---

<sup>427</sup> Gaudiosi, 2014

<sup>428</sup> McClintock, 2014, “\$200 Million and Rising: Hollywood Struggles With Soaring Marketing Costs”

process of audience engagement, designed to automate feedback, optimize targeting, and customize creative. In particular, the majors implemented a key set of data-driven practices:

- *Health Measurement*: The majors drew on data application solutions to measure the health of marketing campaign outreach, effectively enabling a more responsive and dynamic outreach model characterized by ongoing information input and tailored marketing output.
- *Targeting Optimization*: The majors utilized data application solutions to identify audience segments with a high movie-going quota, creating an opportunity to optimize marketing outreach, target campaigns more efficiently, and control budget spend.
- *Creative Customization*: The majors engaged data application solutions to tailor and customize promotional materials to match the interest and preferences of target audience segments, thereby crafting custom marketing campaigns.

By applying data across the marketing value chain, the majors effectively reconfigured the existing promotional system to engage the audience.

### Health Measurement

The majors partnered with data application companies to automate the measurement of marketing campaigns on an ongoing and dynamic basis. In particular, film studios and television networks applied data to evaluate the *health* of a campaign, consistently tracking audience feedback in order to determine the impact of marketing materials on campaign efficiency and

return on investment.<sup>429</sup> Accordingly, data application solutions automatically collected and analyzed information across the digital landscape to measure audience engagement and reception, beginning with the release of the first piece of digital marketing content, such as an announcement post on social media, a teaser trailer on online video platforms, or a concept pitch made available on websites. Data application solutions measured how audiences would interact with marketing materials, detailing video views, likes, shares, and comments to project a quantitative view on audience engagement. At the same time, they examined sentiment surrounding audience activity and interaction, offering a qualitative assessment of audience opinion on a new film or television show. As such, they enabled the majors to track campaigns early on, reviewing how audiences interact with, and feel about, upcoming films and television shows. As a result, the majors established an on-demand digital survey panel, delivering insights on the state of marketing campaigns at an accelerated and more expansive rate.

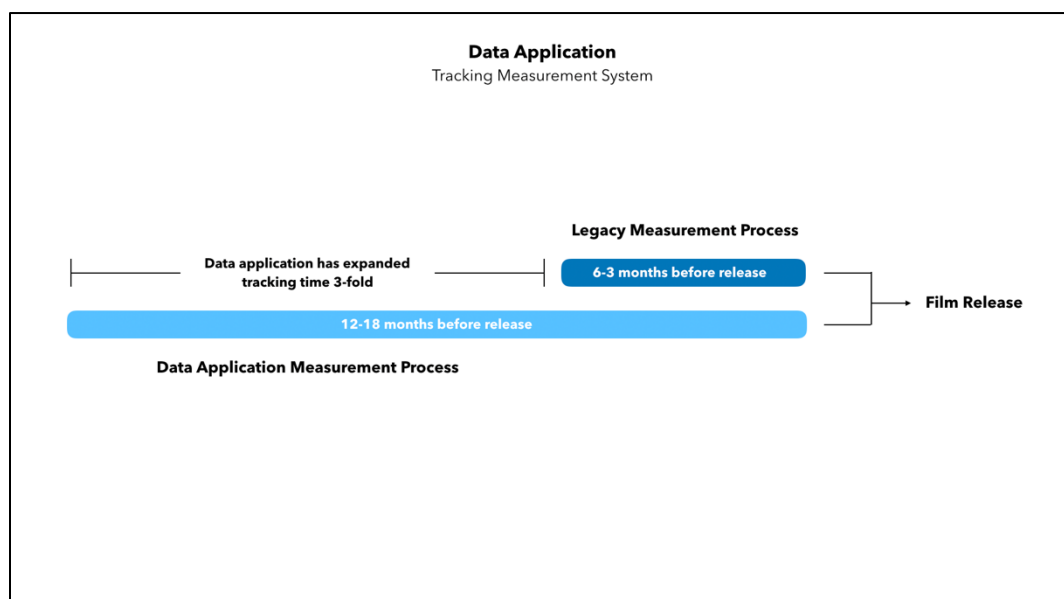
The integration of data application solutions expanded Hollywood's largely static legacy measurement system, in several ways. One, they extended the measurement timeline by enabling the majors to track audience engagement and reception on an ongoing basis. Effectively, studios and networks were able to measure online response after releasing the first piece of marketing content into the digital sphere. This dynamic, on-demand approach to measurement was not inherent to the legacy media business. In film, the tracking process customarily set in 3-6 months ahead of the release.<sup>430</sup> The use of data application allowed studios to gather information and develop insights much earlier. For example, PreAct, a joint venture between United Talent

---

<sup>429</sup> The term marketing campaign health is adopted from Marketing industry jargon and generally used with analytics measurement software. See, Boyd, Joshua, "Brand Health and How to Measure It," *Brandwatch*, November 6, 2017, accessed June 6, 2020, <https://www.brandwatch.com/blog/brand-health-how-to-measure/>

<sup>430</sup> Barnes, 2014, "Hollywood Tracks Social Media Chatter to Target Hit Films"

Agency, one of Hollywood’s preeminent deal-making institutions, Rentrak, an entertainment data company, as well as Crimson Hexagon, an analytics startup, applied a data-driven software program to collect and consolidate data a year ahead of the release date, thereby helping studios read, review, and respond to audience feedback over a longer period of time. The extended period of insight into audience feedback provided studios with more flexibility in adjusting marketing materials and improving outreach. As one industry executive noted, “[m]easuring things a year out is a difference between course correction and bracing for impact. It’s truly changing the way things are done”<sup>431</sup> (figure 3.9).



**Figure 3.9:** Traditional tracking starts 6-3 months ahead of a film release while data application technology enables a more expansive and dynamic timeline to help studios measure a film’s marketing campaign early on.

In television, audience measurement generally involved data sets culled from pilot test screenings, monthly surveys, and Nielsen ratings, delivered via monthly reports, which enabled networks to track feedback pre-release and during release of new shows. These research

<sup>431</sup> Ibid



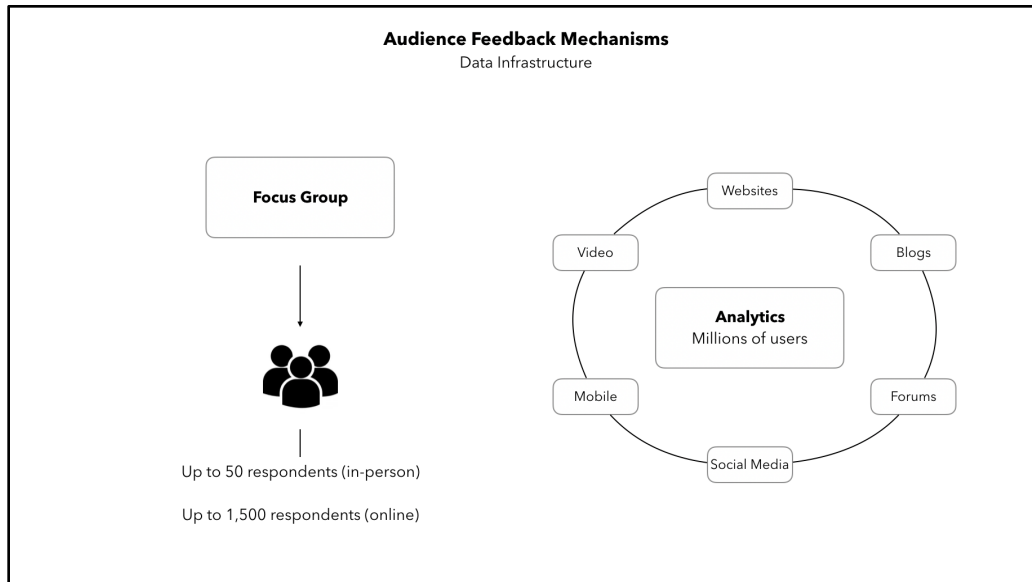
mechanisms functioned as intervals-based information displays rather than real-time insights generators. Data application companies expanded on this process by executing ongoing data pulls in order to model real-time feedback. For example, EDO, a TV analytics startup, invested in real-time data collection and analysis to help networks adjust their marketing based on ongoing data input.<sup>432</sup>

Two, data application solutions expanded the volume and depth of information available to the majors, moving beyond focus groups and surveys to establish digital data repositories. While the majors previously relied on legacy research vendors to conduct in-person focus groups or conduct telephone and online survey polls with expanded user groups, data application companies designed algorithms that leveraged access mechanisms such as application programming interfaces (APIs) to collect data across the entire digital landscape. For example, Parrot Analytics promoted its platform as using the “world’s largest audience behavior datasets,” which included “peer-to-peer downloading/streaming networks, social media sites, photo sharing sites, blogging and micro-blogging sites, video streaming sites, fan and critic sites, and informational sites.”<sup>433</sup> As such, data application solutions enabled the majors to expand their tracking scope, considering larger quantities and detail of information in identifying key themes within audience feedback (figure 3.10).

---

<sup>432</sup> See, Spangler, Todd, “Edward Norton-Backed TV-Analytics Startup EDO Raises \$12 Million,” *Variety*, November 1, 2018, accessed June 6, 2020, <https://variety.com/2018/digital/news/edward-norton-edo-analytics-funding-12-million-1203014461/>

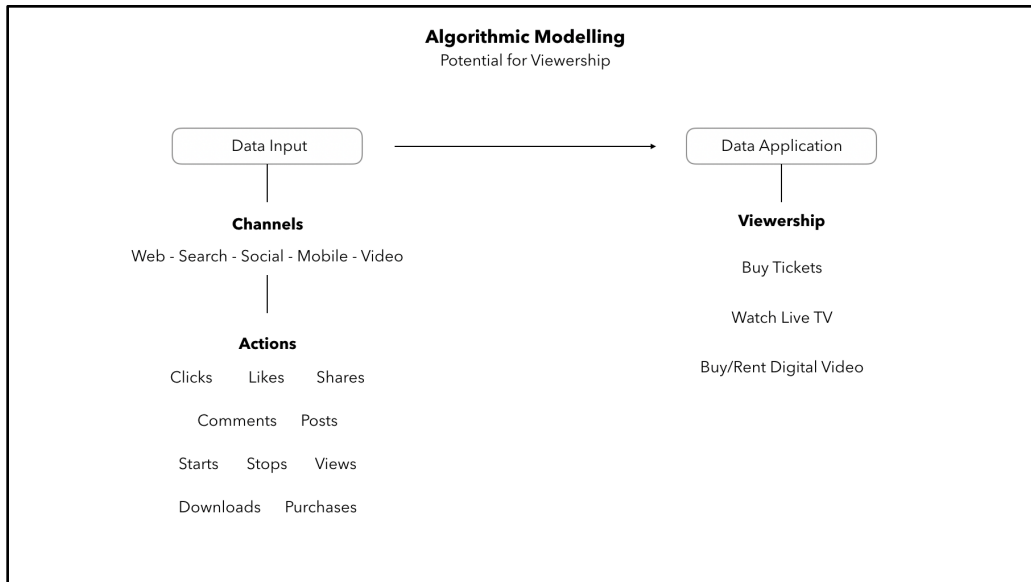
<sup>433</sup> Website, “What are your data sources in more detail?,” *Parrot Analytics*, accessed June 6, 2020, <https://support.parrotanalytics.com/hc/en-us/articles/222713988-What-are-your-data-sources-in-more-detail>



**Figure 3.10:** While focus groups have expanded the number of respondents with investments in online formats, the rise of digital analytics software has increased the intake of information exponentially.

Three, data application solutions enabled the majors to track discrete actions within audience behavior as a foundation to model future intent. While focus groups and pre-release surveys generally captured hypothetical information (i.e., would you see this film, would you watch this ...?),<sup>434</sup> data application solutions provided empirical insight into recorded audience interaction with content. By drawing on actionable indicators (e.g., video views, video shares, likes, and comments), also known as signals, they effectively quantified audience engagement, using algorithmic projections to showcase the probability of certain audience segments going to see a film in theaters, tune in to primetime viewing, or watch content on-demand (figure 3.11).

<sup>434</sup> See, Willens, Michele, “Putting Films to the Test, Every Time,” *New York Times*, June 25, 2000, accessed June 6, 2020 <https://www.nytimes.com/2000/06/25/movies/film-putting-films-to-the-test-every-time.html>



**Figure 3.11:** Data application companies consider a variety of signals (i.e., consumer and usage data) across the digital landscape to model and predict viewership actions (e.g., buying tickets or tuning into live TV broadcasts).

By partnering with data application companies, the majors effectively expanded the legacy media industry’s established measurement practices to more closely monitor and examine the health of marketing campaigns:

- *From tracking information in intervals to measuring on-demand / anytime:* While the majors previously gathered information at regular intervals, providing snapshots of audience feedback, data application solutions enabled the consistent and ongoing tracking of audience information.
- *From tracking select groups to measuring an expansive audience:* The use of data application solutions enabled the majors to expand beyond focus groups and survey-based methodologies to measure audience activity across the entire digital landscape, covering a more expansive information set.

- *From tracking audience intent to measuring action:* Data application solutions expanded the measurement beyond audience intent, captured through self-reporting estimates and hypothetical scenarios, to audience actions, recorded across the digital landscape.

Data application solutions enabled the majors to analyze more information and develop granular insight into the health of a marketing campaign over an extended time period in order to identify, process, and react to changes in audience behavior. Effectively, data application solutions functioned as a course-correction tool, helping the majors adjust marketing campaigns to account for ongoing audience feedback.

### Targeting Optimization

The ongoing integration of data application solutions enabled the majors to reconfigure and optimize the marketing outreach to audiences by increasingly investing in audience targeting. While studios and networks previously largely operated under the dynamics of the four-quadrant model, targeting audiences by static categories across a concentrated set of marketing channels, the use of data application solutions facilitated a more focused, dynamic, and agile approach. Indeed, by automating audience feedback at scale and, at times, in real-time, data application solutions provided the majors with a consistent stream of insights to update and adjust marketing campaigns dynamically over time.

The majors utilized data application solutions to establish an ongoing feedback loop with the audience, tracking information across the entire digital landscape to identify audience segments that responded to and engaged with promotional content for a film or television show. Data application solutions measured a wide range of quantitative and qualitative signals to identify interest among potential audience members, from discrete actions (e.g., video views,

digital transactions) to behavioral patterns (e.g., likes, shares), sentiment (e.g., online posts or comments that relate to the film or television show), and background information (e.g., highlighted interest in a genre, cast member, or related media by way of likes and/or follows on social media as well as downloads across mobile marketplaces and e-commerce platforms). Furthermore, they applied algorithms to arrange audience behavior and feedback by impact level, modelling micro-segments of viewers that showed a high likelihood of seeing a new film or television show. These micro-segments, often characterized as “taste clusters,”<sup>435</sup> united various audience members across the digital landscape into addressable marketing targets for the majors, enabling the legacy media industry to deploy budget across specific marketing channels to optimize their outreach. Rather than allocating marketing budget across a small set of traditional channels, with a primary focus on television, the majors deployed marketing funds strategically across a wide range of digital channels, from web to search, social media, mobile, and online video, with the ultimate channel mix tailored to audience behavioral patterns and preferences.

Accordingly, the majors increasingly relied on data application solutions to collect, curate, and commercialize audience feedback, deploying data-driven algorithms to identify, assemble, and target addressable audience segments across digital marketing channels. The startup TV Time offers an interesting use case in this context. TV Time, launched in 2012 as TV Show Time and later acquired by Santa Monica-based Whip Networks and rebranded as TV Time, represents a film and television tracking platform that operated an integrated business-to-consumer (B2C) and business-to-business (B2B) model. On the B2C side, TV Time allowed users to track new television shows, document their viewing habits, and follow what other users

---

<sup>435</sup> Nguyen, Nicole, “Netflix Wants To Change The Way You Chill,” *BuzzFeed News*, December 13, 2018, accessed June 6, 2020, <https://www.buzzfeednews.com/article/nicolenguyen/netflix-recommendation-algorithm-explained-binge-watching>

are watching, effectively functioning as a social television recommendation tool. On the B2B side, TV Time utilized proprietary algorithms to collect, consolidate, and communicate the user data to television networks through its TVlytics platform, showcasing viewing trends and sentiment patterns among its audience. With over 13 million global users, up to one million daily active users, and a catalog of over 60,000 shows, TV Time effectively operated as an on-demand focus group for the television industry, providing “a read on what fans of a particular show are especially passionate about (or, by the same token, what they hate) and benchmark that against other content,”<sup>436</sup> thereby helping networks program their marketing approach and target specific audience segments to achieve more efficient return on investment.

Data application solutions effectively consolidated, conformed, and curated the broader media audience for the majors, delivering a set of audience data cuts that highlighted overlapping interest levels based on industry metadata, such as genre, theme, and cast. As a result, the majors were able to determine an addressable audience for their new film and television releases by analyzing data on existing content benchmarks. For example, rather than promoting a new comedy show to men and women under 25 through ads across comedy channels on TV, data application solutions enabled the targeting of specific comedy fans, with interest in the cast, the theme of the show, as well as its comparatives (i.e., other shows that have a similar theme or tone) across the digital platforms they spend most of their time on and display the highest likelihood of engaging with marketing content. In this regard, the majors expanded four-quadrant targeting to be more interest-based and channel-/platform-specific.

---

<sup>436</sup> Spangler, Todd, “TV Time Launches Social-Analytics Tool to Break Down Fan Reaction to Shows,” *Variety*, July 31, 2018, accessed June 6, 2020, <https://variety.com/2018/digital/news/tv-time-fan-reaction-social-analytics-1202888016/>

While data application solutions enabled the majors to define and target addressable audience segments based on general viewing behavior and programming interests up front, they further facilitated the dynamic development of audience targets based on specific feedback over time. Filmmaker Eli Roth, known for low-budget independent horror fare such as *Cabin Fever* (Lionsgate, 2002) and *Hostel* (Lionsgate, 2005), promulgated a data-driven marketing approach to optimize and control costs for films with limited budgets by incorporating audience feedback into the segmentation of the addressable audience. In particular, his marketing team consistently tested audience feedback in linear fashion, distributing marketing materials across specific platforms to gather insight into audiences' response to the creative materials and adjusting the audience targets from there, re-targeting audiences that displayed interest while removing those segments that did not respond. This dynamic testing approach further revealed interest among unconventional audience members outside the typical horror demographic, enabling Roth and team to expand their audience target. As recounted by Roth (my emphasis):

**"Everything that we are doing informs who we market to next. We are letting the data tell us who's responding. Like, OK, we know we've got the males. The horror fans are going crazy, but as a test we tested against Selena Gomez's audience and Taylor Swift's, just to see, we found teenage girls went crazy for the hardcore gory clips. It wouldn't have been my intuitive pairing, but the data is telling us that the audience is going insane for it. So we can now retarget spots and redirect."**<sup>437</sup>

The use of data application enabled Roth and team to uncover addressable audience segments and expand the targeting of their marketing to optimize audience engagement. Effectively, data application enabled the identification and targeting of audiences with a high level of interest and

---

<sup>437</sup> Crucchiola, Jordan, "Eli Roth Has a Data-Happy Vision for Cinema's Future," *Wired*, September 25, 2015, accessed June 6, 2020, <https://www.wired.com/2015/09/eli-roth-data-crunching/>

a corresponding likelihood of seeing a film in theaters, increasing the probability of return on investment. As Roth noted, data application surfaced information on who was interested in the marketing materials and thus enabled effective targeting (my emphasis):

"[W]e can really see in like a 12 or 24 hour period whether it was males or females, or by interest groups like people that like Danny Trejo or Quentin Tarantino. We have their data so let's retarget the people that have liked the Facebook groups. The people that responded, we can tell they like Gossip Girl. They like all this different stuff I could have never anticipated. **So you're just listening to the data. Our goal is to only reach people that are interested.**"<sup>438</sup>

Roth subsequently operationalized this approach with the founding of Crypt TV, a digital entertainment company focused on horror genre content which distributes content across digital platforms. Crypt TV drew on data application solutions to identify fans and potential viewers for its horror programming across the digital landscape, relying on data-driven audience feedback to optimize targeting and control costs. As Crypt TV co-founder and CEO Jack Davis explained, "[w]e have our tracking pixel on all those sites. When we put a piece of content out we can see exactly who is watching so we can hyper-target those groups [...]."<sup>439</sup> Crypt TV further implemented this approach to test the viability of programming concepts for linear television and film releases, measuring audience feedback to determine increased programming investment. While this approach effectively extends the legacy media industry's greenlighting and investment process, it equally constitutes an exercise in data-driven marketing, utilizing promotional materials to identify an addressable audience prior to production.

---

<sup>438</sup> Ibid

<sup>439</sup> Ibid



The Crypt TV model reflects a broader trend of digital media companies testing the viability of short-form digital content for the bigger screens of film and television. For the legacy media industry, this model constituted a cost-effective strategy in identifying existing audience segments for new programming, reducing downstream marketing and development costs. During the early 2010s, in particular, with audiences increasingly fragmenting across digital platforms in a shift away from legacy media, the majors acquired stakes in digital media companies as a way to access proprietary data-driven and algorithmic solutions, with the long-term goal of consolidating addressable audience segments across the digital landscape to efficiently market legacy media releases (table 3.4).

**Table 3.4**

<b>Data Application Marketing Investments</b>			
<b>Digital Media Company</b>	<b>Legacy Media Investor</b>	<b>Year</b>	<b>Investment</b>
BuzzFeed	NBC Universal	2015 and 2016	\$400 million
Vox Media	NBC Universal	2015	\$200 million
Snap, Inc.	NBC Universal	2017	\$500 million
Crypt TV	NBC Universal, Blumhouse	2017 and 2018	Undisclosed (part of two venture rounds totaling \$3.5 and \$6.2 million)

The integration of data application solutions effectively reconfigured Hollywood’s “spray and pray” marketing approach into a “test and learn” framework, enabling the majors to iterate on marketing campaigns incrementally by incorporating audience feedback on a dynamic and agile basis, adjusting the marketing outreach across channels to target the most interested viewer segments. Indeed, the use of data application facilitated an increasing focus on digital marketing channels and prioritization of data-driven targeting, based on audience behavior and feedback.

## Creative Customization

The partnerships with data application companies recast the legacy media industry's approach to marketing content creation. While the majors previously relied on a small set of trailers to promote a new film or television release over the course of the pre-release window, the growing availability of digital information on media audiences increasingly enabled a more customized approach. In particular, film studios and television networks created a suite of marketing materials designed to appeal to a focused set of audience segments, customized to their media behavior and interests. Thus, rather than crafting a few trailers geared toward a broad, four-quadrant audience, the majors invested in producing a wide range of promotional video content to customize the marketing experience for data-driven target audience segments. The integration of data application into the process of marketing content creation had several implications for the legacy media industry.<sup>440</sup>

First, the legacy media industry applied data to scale and differentiate marketing output. While the majors previously relied on a key set of trailers, differentiated by channel, to reach a broad four-quadrant audience, the integration of data application solutions enabled an exponential increase in the output of creative marketing materials. As studios and networks defined addressable audience segments by their viewing behavior and interest, creating more granular and multi-dimensional targets for marketing outreach, the demand for differentiated marketing materials grew accordingly. To this end, the majors increasingly expanded from a small set of trailers to a portfolio of materials, customized to audiences' viewing patterns and interests, and distributed across the digital landscape. To manage the increase in marketing

---

<sup>440</sup> Streaming companies also applied data to reconfigure marketing practices. Netflix cut different trailers for *House of Cards* to appeal to viewing patterns and interests of key audience segments. Netflix also implemented a new data-driven algorithmic model to automatically customize thumbnails and cut trailers based on viewer profiles.

output, the majors continued an established legacy practice, outsourcing production to boutique vendors specializing in marketing production. Given the rise in demand across the legacy media business, the number of vendors grew from a dozen to more than a hundred, working consistently to customize marketing materials for differentiated audience segments.<sup>441</sup>

Second, the legacy media industry applied data to implement a more dynamic and agile production model, incorporating data-driven insights on audience viewing behavior and interests to customize output on a dynamic and ongoing basis. As such, studios and networks expanded from developing trailers up front to producing promotional content during the marketing rollout. In effect, the majors measured audience feedback on marketing output with the goal of identifying popular elements, such as characters, themes, and creative aesthetics (e.g., the choice of music), to inform content creation, effectively tailoring marketing output to audience reaction.

And, third, the legacy media industry applied data to reconfigure and expand the underlying format of marketing materials, thereby creating new creative templates to drive audience engagement. In particular, the majors drew on insights into how audiences interacted with content across the digital landscape to modify content formats, customizing promotional materials to align with dominant behavioral patterns across digital channels and platforms. For example, while the proliferation of digital platforms enabled studios and networks to expand beyond linear advertising modules (e.g., the 30-second ad, weekly print ads), it equally engendered more opportunities for consumers to actively skip ads (e.g., YouTube's automatic 5-second skip button). To this end, the majors experimented with marketing formats that put engaging content in the first few seconds. The *bumper trailer* concept is a case in point. It

---

<sup>441</sup> See, Faughnder, Ryan, "Movie trailer makers multiply as online viewing of previews soars," *Los Angeles Times*, July 21, 2015, accessed June 6, 2020, <https://www.latimes.com/entertainment/envelope/cotown/la-et-ct-hollywood-trailers-20150721-story.html>

represents a typical teaser or trailer, preceded by a 5-second bumper that distills key moments from the trailer in a vivid barrage of images, effectively telling the story of the trailer within five seconds. According to an industry executive, the use of bumpers increased overall trailer retention, ensuring more viewers watched the marketing content in full. The bumper format was further extended into previews for trailers, with studios and networks deploying short-form previews to announce the release of new trailers ahead of time, an overarching attempt to establish the release of marketing content as an event within the digital marketplace.<sup>442</sup>

Accordingly, the integration of data application solutions reconfigured the legacy media industry's established approach to marketing content creation and enabled studios and networks to drive audience engagement in several new ways:

- *From relying on a core set of trailers to prioritizing an expansive portfolio of promotional materials:* The majors increased the output of marketing materials to respond to growing demand among audiences, partnering with boutique vendors to create a wide range of data-driven promotional content.
- *From setting a trailer vision early to implementing a dynamic approach of ongoing customization:* The majors customized marketing materials based on ongoing audience feedback, incorporating data-driven insights to adjust promotional content over the course of the marketing rollout.

---

<sup>442</sup> See, Dickey, Josh, "Movie trailers have an effective new strategy you might not have noticed," *Mashable*, June 19, 2017, accessed June 6, 2020, <https://mashable.com/2017/06/19/movies-trailers-bumpers-online-pre-roll/>

- *From utilizing a traditional trailer approach to embracing data-driven formats:* The majors reconfigured the traditional trailer format, creating new content containers that align with audience behavioral patterns across the digital landscape.

As such, the majors' use of data application solutions enabled a set of enhanced promotional practices, automating the incorporation of audience viewing behavior and interests into the marketing outreach of studios and networks.

--- --- ---

The integration of data application solutions enabled the major film studios and television networks to reconfigure key practices of the marketing value chain. In particular, the majors were able to automate the measurement of marketing campaigns, identify data-driven insights to algorithmically segment and select addressable audience segments, and optimize audience targeting by delivering customized marketing content based on audience behavioral patterns and interests. As a result, the majors were able to drive audience engagement in a shifting media environment, marked by ongoing audience fragmentation, shifting consumer behaviors, and the proliferation of new marketing channels.

At the same time, while the implementation of data-driven marketing practices enabled the majors to increase the efficiency, elevate the reach, and control the costs of marketing campaigns, data-driven marketing was not universally implemented across the legacy media industry. Indeed, studios and networks used data application solutions as an incremental extension of traditional marketing outreach, generally deploying their digital budgets to experiment with data-driven, algorithmic practices in an effort to navigate and manage the shifting digital media environment. Data application, while innovative, equally proved fickle,

costly, and challenging to implement at scale. In general, the majors continued to prioritize legacy media channels, which effectively positioned data application as an experimental add-on, rather than an established mainstream phenomenon.<sup>443</sup>

Overall, the growing number of use cases in data-driven marketing, from independent, low-budget films and television series launches to mainstream blockbusters, signaled an increasing viability of data application in Hollywood's marketing value chain. By partnering with data application companies, the majors effectively cultivated a robust model of marketing efficiency, implementing a system that did not prove infallible, but worked to identify, target, and engage audience segments with a higher likelihood of turning out at the movies and tuning into television broadcasts.

### **Data-Driven Distribution: Maximizing the Window**

In 2010, Time Warner led a \$6 million investment in Get Glue, a digital media company specializing in second-screen software applications.<sup>444</sup> In particular, Get Glue programmed a website that enabled users to *check-in* to films and television shows, adopting a digital feature popularized by location-based service Foursquare, and further interact with content by writing reviews, comments, and ratings as well as sharing updates and news with friends. To incentivize users to apply the check-in feature via its website, the company implemented a rewards system that offered users virtual stickers to showcase their entertainment viewing across the digital

---

<sup>443</sup> See, Sweeney, Erica, "Digital represents 14% of movie ad budgets but drives 46% of revenue, study finds," *Marketing Dive*, December 13, 2018, accessed June 6, 2020, <https://www.marketingdive.com/news/digital-represents-14-of-movie-ad-budgets-but-drives-46-of-revenue-study/544253/>

<sup>444</sup> See, Time Warner Press Release, "Time Warner leads \$6M round of funding in Get Glue, the leader in social entertainment." *Time Warner Group (WarnerMedia)*, December 7, 2010, accessed June 6, 2020, <https://www.warnermediagroup.com/newsroom/press-releases/2010/12/07/time-warner-leads-6m-round-of-funding-in-getglue-the-leader-in>

landscape.<sup>445</sup> As users checked in to a film or television show, Get Glue automatically surfaced stickers across its website and social platforms like Facebook, allowing users to highlight their viewing experience and alert their friends, and effectively functioning as an amplification device. At the same time, the company functioned as an online version of TV Guide, offering recommendations based on users' viewing behavior and content preferences. Effectively, Get Glue operated as an entertainment-focused social network that made it possible for users to signal and share their film and television viewing experience with their friends and followers on the digital screen, thereby facilitating conversations, bringing the “water cooler effect online,”<sup>446</sup> and encouraging co-viewing sessions. As such, Get Glue fueled the early stages of the social TV phenomenon,<sup>447</sup> incorporating social media technology into the traditional media experience and blending the digital environment with the television screen.<sup>448</sup>

Get Glue subsequently developed a suite of partnerships with legacy media players, signing deals with over 75 broadcasters and cable companies across 680 shows, including Time Warner (e.g., HBO, CNN), Comcast's NBC Universal (e.g., USA Network, SyFy, MSNBC), The Walt Disney Company (e.g., ABC, ESPN), Fox, Showtime, and MTV, all of which sought to tap into Get Glue's engaged television audience.<sup>449</sup> Additionally, the company raised new

---

<sup>445</sup> See, Rose, Frank, *The Art of Immersion: How the Digital Generation Is Remaking Hollywood, Madison Avenue, and the Way We Tell Stories* (New York: W. W. Norton & Company, 2012).

<sup>446</sup> Vidyarthi, Neil, “Social TV: Water Cooler Effect Gone Online,” *AdWeek*, June 19, 2012, accessed June 6, 2020, <https://www.adweek.com/digital/social-tv-water-cooler-effect-gone-online-at-140conf-live-video/>

<sup>447</sup> See, Gillan, Jennifer, *Must-Click TV* (New York: Routledge, 2010).

<sup>448</sup> See, Mike Proulx & Stacey Shepatin, *Social TV: How Marketers Can Reach and Engage Audiences by Connecting Television to the Web, Social Media, and Mobile* (Hoboken: Wiley, 2012).

<sup>449</sup> See, Mann, Denise, “Introduction: When Television and New Media Work Worlds Collide.” In: *Wired TV: Laboring Over an Interactive Future*, edited by Denise Mann (New Brunswick: Rutgers University Press, 2014).

funding rounds to incubate a mobile app and implement an API (application programming interface) platform to integrate its software into digital distribution mechanisms such as set-top boxes (e.g., Get Glue launched its first integration with Pay-TV provider DirectTV in 2011, offering a Social TV feature as part of the company's wireless television service. Furthermore, Get Glue developed features to reward users for actively viewing live television programming, partnering with networks on incentives and prizes. For the legacy media industry, the holistic investment in Get Glue offered a number of strategic benefits. One, the company counted close to one million active users, generating over 10 million monthly interactions on its website, which gave it access to a growing amount of user and usage data and insight into dominant viewing habits and content preferences.<sup>450</sup> Two, while Get Glue's user base represented only a small portion of the overall television audience, it targeted viewers who actively watched, discussed, and shared their programming experience online. As such, the company's user base consisted of high-volume television viewers and dedicated fans who, by documenting their viewing online, functioned as promotional ambassadors for the industry. And, three, Get Glue's check-in software application provided legacy media companies with a digital mechanism to boost content distribution and develop audience loyalty, using the company's website and user base as an incremental amplification device to highlight its programming slate. Indeed, Time Warner leveraged Get Glue as a distribution add-on for shows from its cable subsidiary HBO, helping the premium cable channel amplify its audience during Sunday night prime time hours.<sup>451</sup> In this sense, broadcast and cable networks effectively implemented Get Glue as an incremental data-

---

<sup>450</sup> Warren, Christina, "GetGlue Gets Slick New Dashboard for TV Marketers," *Mashable*, September 8, 2011, accessed June 6, 2020, <https://mashable.com/2011/09/08/getglue-business/>

<sup>451</sup> See, Reuters Press Article, "GetGlue to boost live TV with Facebook check-ins," *Reuters*, March 23, 2011, accessed June 6, 2020, <https://www.reuters.com/article/us-getglue/getglue-to-boost-live-tv-with-facebook-check-ins-idUSTRE72M92720110324>



driven distribution channel, a mechanism that drew on the viewing information of a small set of highly engaged viewers to drive tune-in among a broader television audience.

In a similar vein, Atom Tickets, an online ticketing company, brought the concept of data-driven distribution to the film business. In 2014, the company raised Series A capital from Lionsgate to bring its technology to a small number of regional markets in the United States, partnering with small theater chains on an early-stage test-and-learn initiative.<sup>452</sup> Atom Tickets subsequently launched as a social network and e-commerce platform that enabled movie-goers to search, find, and purchase tickets for entire groups by polling every group member's viewing behavior and content preferences to recommend ideal screenings. Atom Tickets drew on two principal sources of data, namely the user and usage data on its own app as well as external information sourced from digital platforms (e.g., likes on social media). Effectively, the company used its software application to streamline the movie-going experience, making it easier for groups to find a movie, pay for tickets, and organize an outing at the theater. Atom Tickets deployed the data from its user base to algorithmically select a movie, a theater, and a showing that appealed to all members of a movie-going group, thereby limiting the logistics involved in organizing a movie night. Furthermore, the company used data from theaters to discover undersold showings and offer viewers group discounts, enabling studios and theater owners to sell out screenings, thereby improving the industry's revenue per screening average and elevating overall box office performance.<sup>453</sup>

---

<sup>452</sup> See, Crunchbase, *Atom Tickets* entry

<sup>453</sup> Atom Tickets noted that some of its partners experienced “double-digit percentage growth” in ticket and concession sales during the early rollout. See, Barnes, Brooks, “A Movie Ticketing Start-Up Hopes to Fill Empty Seats,” *New York Times*, December 4, 2016, accessed June 6, 2020, <https://www.nytimes.com/2016/12/04/business/media/movie-ticketing-start-up-atom-tickets-hopes-to-fill-theaters.html>

Atom Tickets actively positioned its service offering as a data-driven optimization layer for the theatrical business. Indeed, the company argued that theatrical distribution followed a manual and linear approach that did not incorporate audience behavioral patterns and preferences. Atom Tickets, by contrast, applied a wide range of data to the value chain in order to customize the experience for consumers and optimize the economics for stakeholders involved (i.e., studios and theaters). As one of the co-founders noted accordingly (my emphasis):

"There's a significant data angle to our product. **We collect several thousand data points on a per-transaction basis; including profiles, previous movies seen, prior groups of friends who hang out together, and sentiment analysis from signals across social networks.** We believe there's been no significant tech innovation in the ticketing space for the last decade until we came along."<sup>454</sup>

Indeed, Lionsgate rationalized its early investment in Atom Tickets as an opportunity to close the data gap in film distribution, enabling studios to gain a deeper understanding of audiences and distribute their films more efficiently. As Lionsgate's vice chairman explained, "[w]e believe that this venture continues to bring us closer to our consumers, and it has the potential to revolutionize an industry."<sup>455</sup>

Atom Tickets later raised capital from The Walt Disney Company and 21st Century Fox as well as prominent Hollywood personnel such as J.J. Abrams and The Rock to finance its mainstream rollout.<sup>456</sup> For the legacy media industry, the investment in Atom Tickets offered a

---

<sup>454</sup> Faughnder, Ryan, "This movie-ticket startup says it can fill empty theaters across the country," *Los Angeles Times*, June 28, 2016, accessed June 6, 2020, <https://www.latimes.com/entertainment/envelope/cotown/la-et-ct-atom-tickets-20160620-snap-story.html>

<sup>455</sup> Ibid

<sup>456</sup> See, Shu, Catherine, "Atom Tickets raises \$60M Series C for its movie booking app," *TechCrunch*, March 8, 2018, accessed June 6, 2020, <https://techcrunch.com/2018/03/08/atom-tickets-raises-60m-series-c-for-its-movie-booking-app/>

suite of strategic advantages. One, the company offered a repository of data on movie-goers, their viewing behavior and preferences, which had previously been limited to theater owners. As such, it enabled studios to build a deeper understanding of the theatrical audience. Two, the company hosted a social networking environment for movie-goers who regularly go to the movies, thus representing a core segment of the movie-going audience that drives a disproportionately high percentage of theatrical revenue each year. In this sense, Atom Tickets gave the legacy media industry a direct line to active movie-goers, offering a highly addressable audience segment. And, three, the company facilitated more efficient distribution by directly surfacing movie showings to active movie-goers, matching viewers with the films they want to see, and identifying undersold screenings to increase ticket sales across the theatrical value chain. Indeed, Atom Tickets' initial investor pitch led with the fact that five billion movie tickets went unsold each year, offering a massive incremental revenue opportunity for studios and theater owners. In effect, Atom Tickets used its data-driven ticketing platform to bring avid movie-goers to theaters and drive turnout at the movies.

Get Glue and Atom Tickets share key commonalities in their approach to the legacy media industry. First, both companies built vertically-focused social platforms targeted at avid film and television viewers, gathering information on their viewing behavior and content preferences. Second, they segmented their user base into addressable audience segments, organized by viewing behavior and programming interests. And, third, they implemented algorithms that enabled users to engage with film and television programming. Get Glue enabled users to check in to a television show, interact with friends, and engage in co-viewing activities. Atom Tickets enabled users to discover, select, and buy tickets for specific film screenings. As such, they streamlined and promoted the traditional film and television distribution process.

Get Glue and Atom Tickets represent a set of data application companies that leveraged data-driven, algorithmic software applications to optimize the legacy media industry's distribution mechanisms.<sup>457</sup> In the television business, data application companies used data-driven algorithms to drive tune-in as a way to increase audience viewership of television shows and thereby drive ratings. In the film business, they focused on generating turnout as a way to increase ticket sales and thereby drive box office. While other digital media companies focused on redefining the viewing experience by increasingly shifting distribution online, effectively shifting beyond the established legacy media system with investments in video-on-demand (VOD) and direct-to-consumer streaming initiatives, data application companies prioritized the strengthening of the traditional distribution window, operating as add-on services in the legacy media industry value chain. In this sense, data application enabled the legacy media industry to sustain and supercharge the existing approach to film and television distribution.

The legacy media industry's approach to distribution has traditionally centered around windowing, the process of managing the release of content across multiple window stages to maximize associated revenues.

- *Television*: The television business was organized around time-based distribution, with the first window comprising the week-by-week release of new episodes. Networks made content available incrementally to manage audience development over time and monetize viewership over the course of a season, either through ratings-based advertising sales or subscription-based cable bundles. The second window involved the syndication of

---

<sup>457</sup> On media distribution, see, Ulin, Jeffrey C., *The Business of Media Distribution: Monetizing Film, TV and Video Content in an Online World* (New York: Focal Press, 2010).

successful shows across the television ecosystem,<sup>458</sup> licensing content to other broadcast and cable channels for a specific period of time. Another stage of downstream value creation through windowing included home entertainment, with networks releasing physical and digital copies of their shows through retailers and digital platforms.

- *Film*: The film business was built around a rigid system of linear window sequencing that served to generate incremental revenues over time. The theatrical release constituted the first window, with films competing for ticket sales at the box office. Beyond that, films generally circulated through an interwoven network of monetization channels, including home entertainment (i.e., sales from physical home video like DVDs and Blu-Rays as well as digital video like electronic sell-through and video-on-demand rentals), airline and hospitality licensing (i.e., licensing fees from airlines and hotels), the first Pay-TV window (i.e., licensing fees from subscription video-on-demand services and premium Pay-TV channels), the second Pay-TV window (i.e., licensing fees from premium and standard Pay-TV channels), and the advertising window (i.e., licensing fees from broadcast TV channels and advertising video-on-demand services).

Effectively, legacy media companies deployed windowing as a distribution strategy to manage demand, audience scale, and revenue generation over time. In both film and television, the success during the primary window - the theatrical release and the week-by-week airings - largely determined the profitability of downstream windows.

The development of digital technologies has changed, complicated, and challenged the underlying tenets of the windowing process over the course of the new millennium. In particular,

---

<sup>458</sup> TV shows generally need 100 episodes to enter syndication. See, Allen, Robert Clyde & Hill, Annette, eds., *The Television Studies Reader* (New York: Psychology Press, 2004).

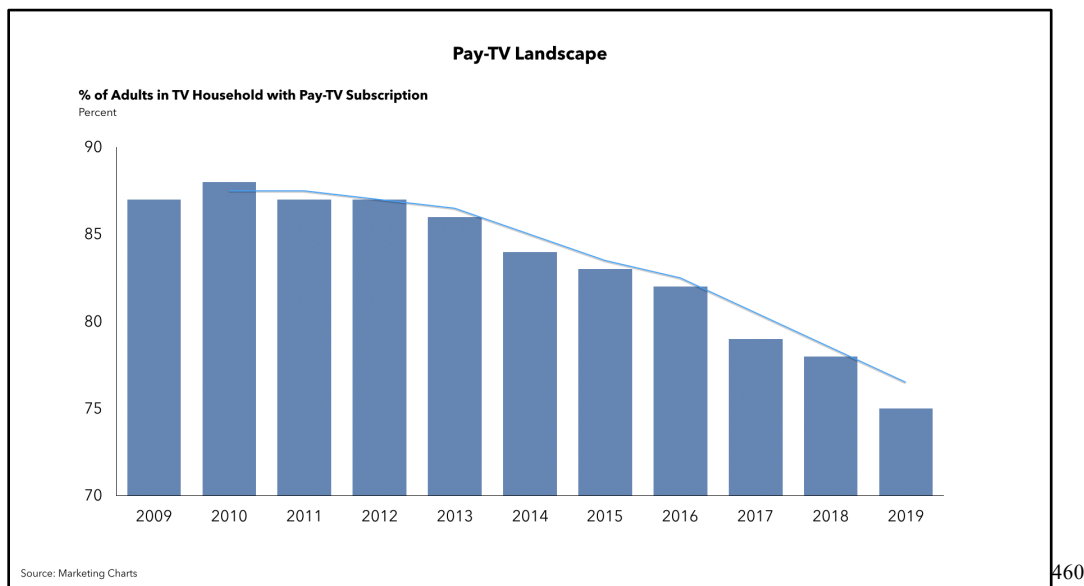
the rise of alternative distribution platforms, shifting viewing behavior, and evolving release mechanisms has given rise to a new era that goes beyond traditional windowing.<sup>459</sup>

- *Alternative Distribution:* The rise of new digital platforms enabled a non-linear distribution environment, with content becoming available across a variety of outlets, from websites to digital platforms, set-top boxes, and mobile apps. As a result, the tenets of theatrical and television distribution increasingly moved to the periphery.
- *Consumer Behavior:* Audiences increasingly migrated to new distribution outlets and adopted new digital viewing practices. In particular, viewers accessed film and television programming through a suite of emerging devices, such as Internet-connected television, computers, smartphones, and tablets. Additionally, audiences signed up for digital streaming services at an accelerating rate. For example, Netflix grew its U.S. subscriber base from 15 million at the beginning of 2010 to 61 million by the end of 2019.
- *Release Mechanisms:* The majors experimented with emerging release strategies to account for the shift to a digital environment. Studios premiered day and date releases, debuting new films in theaters and digital on the same day. Meanwhile, networks made television episodes available on websites the day after airing, enabling audiences to catch up on shows online, and transferred cable and satellite programming to the digital landscape with the launch virtual multi-video programming mechanisms.

---

<sup>459</sup> See, Wasko, Janet, "Hollywood in the 21st Century," *Economia della Cultura*, Issue 4, 2017 and Ball, Matthew, "Letting it Go: The End of Windowing (and What Comes Next)," *Redef*, August 24, 2016, accessed October 5, 2019, <https://redef.com/original/letting-it-go-the-end-of-windows-and-what-comes-next>

In this environment, the legacy media industry’s existing distribution dynamics and associated revenues have increasingly come under pressure. The television business has seen a steady decline in viewership, ratings, and subscriptions, with a growing number of viewers migrating to digital viewing options (e.g., downloads, video-on-demand, streaming), cutting the proverbial cable cord for alternative subscription models, and embracing new screens across the digital landscape (e.g., laptops, smartphones, tablets). For example, the number of domestic households with a Pay-TV subscription has declined over the past decade (figure 3.12).



**Table 3.12:** The overall percentage of adults with a Pay-TV subscription has declined over the past decade, which directly impacted the bottom line of the majors as cable revenues make up a large portion of media capital.

Similarly, the film business has experienced an ongoing downturn across traditional channels, with box office revenues, ticket sales, and attendance declining among younger audience segments, with overall economics in ongoing fluctuation. While overall box office

<sup>460</sup> Article, “U.S. Pay-TV Penetration Rate, 2009-2019,” *MarketingCharts.com* (Leichtman Research Group, “Pay-TV in the U.S.,” 2019), January 10, 2020, accessed June 6, 2020, <https://www.marketingcharts.com/charts/us-pay-tv-penetration-rate-2009-2019>

revenues have shown incremental growth, due to raised ticket prices, tentpole releases, and a shift to platform filmmaking (i.e., the release of interconnected franchise films over time), ticket sales and admissions have been in decline since peaking in the early 2000s.<sup>461</sup> In particular, the number of frequent movie-goers, has seen intermittent periods of decline.

At the same time, the shift beyond traditional windowing has enabled a set of distribution dynamics previously unavailable to the legacy media industry, making content available directly to consumers anytime and anywhere across a range of digital channels. While the release of films and television shows generally followed a sequential logic, the new media environment pressured the temporal parameters of windows. Networks made television shows available on websites and mobile apps for a limited time after they aired in their scheduled broadcasting slots in order to accommodate new consumer viewing behaviors and preferences. Studios contracted, curtailed, and collapsed release windows, with films spending considerably less time in theaters and appearing across multiple distribution windows within shorter periods of time, as a way to lean into new audience dynamics.<sup>462</sup> As such, the legacy media industry has taken steps to negotiate and adapt to the dynamics of the new media windowing environment.

Yet, moving beyond windowing proved challenging and cumbersome for the legacy media industry, for several reasons. First, the majors effectively had to invest in cultivating a new distribution approach, deploying capital to license and build direct-to-consumer technologies, bring in external talent to devise new release strategies for the digital landscape, and negotiate licensing and rights fees with a new set of digital partners. At the same time, the

---

<sup>461</sup> Statista, “Number of movie tickets sold in the U.S. and Canada from 1980 to 2019,” April 24, 2020, accessed June 6, 2020, <https://www.statista.com/statistics/187073/tickets-sold-at-the-north-american-box-office-since-1980/>

<sup>462</sup> See, Fahey, Mark, “Why movies are sometimes here and gone in theaters,” CNBC, November 17, 2015, accessed June 6, 2020, <https://www.cnn.com/2015/11/17/why-movies-are-sometimes-here-and-gone-in-theaters.html>



majors had to balance investments in the legacy media business which, while challenged and pressured following the shift beyond windowing, remained profitable and valuable to the industry's bottom line. Second, the majors had to adjust well-honed and long-standing processes for a new era, adapting established economic, cultural, and technological dynamics that defined operations for decades. And, third, the majors had to manage change with their existing set of legacy media partnerships in the distribution landscape, primarily advertisers and theater owners, whose own business value was predicated on the legacy media industry's windowing model.<sup>463</sup>

Accordingly, the legacy media industry remained invested in the traditional windowing approach as a way to sustain its existing business model, partnerships, and associated revenues. Data application companies emerged as viable incremental mechanisms to manage the legacy media window in a new digital era. Indeed, companies like Get Glue and Atom Tickets were designed to vertically integrate with the legacy media industry's distribution value chain to maximize the value of the existing window by increasing audience tune-in and turnout. To this end, the majors cultivated partnerships with data application companies across the distribution ecosystem (table 3.5).

---

<sup>463</sup> The film industry, in particular, has undergone a wide range of negotiations with theater owners about adapting the window for the digital era. See, Wasko, 2017

**Table 3.5**

<b>Data Application in Distribution</b>				
<b>Company</b>	<b>Industry</b>	<b>Founding Year</b>	<b>Model</b>	<b>Select Partnerships</b>
Get Glue	Television	2010	Social check-in	USA Network, SyFy, MSNBC, ABC, ESPN, HBO, CNN
Miso	Television	2010	Social check-in	Starz, Fox, Showtime, Food Network, USA, Comedy Central, TNT
Viggle	Television	2012	Social check-in	DirectTV, HGTV, Hulu
Fandango	Film	2000	Online Ticketing	All major film studios
Movio	Film	2010	Audience Targeting	Viacom, STX Entertainment
Tugg.com	Film	2011	Cinema-on-demand	Filmmakers, distributors, exhibitors
Atom Tickets	Film	2014	Online Ticketing	Lionsgate, Disney, 21st Century Fox
MoviePass	Film	2016	Cinema subscription	Landmark Theater Group

Data application companies, in concept, enabled the major studios and networks to distribute programming more efficiently during the primary window, following an integrated process of data-driven mechanics. They built vertical social platforms that promoted an interactive exchange on media programming among audiences, specifically targeting active television viewers and movie-goers with a high probability of seeing content during the first window. They then accessed and analyzed a wide range of data on user behavior and preferences and developed addressable audience segments, organized by their interest in film and television programming. Subsequently, they built custom algorithms to action against those audience segments, bringing viewers with similar behaviors and interests together. At the same time, their reliance on algorithmic programming played into the larger network effect of digital user bases, with audiences promoting, sharing, and broadcasting their viewing across the digital landscape, effectively exposing the legacy media release window to a wider audience. In short, they

surfaced specific films and television shows to highly interested audience segments during the primary release window, effectively aligning custom audiences for the legacy media industry's distribution channels. They applied data-driven algorithms to promote the traditional media experience, communal viewing structured around the primary media screens (i.e., the television and the movie theatre).

Data application thus enabled targeted distribution, channeling the release of film and television programming to active viewers who, in turn, promoted the viewing experience to their audience networks, thereby increasing the exposure of the primary window. As such, data application companies deployed data-driven algorithms to select audience segments that would maximize viewership and revenues during the traditional window.

Notably, this approach to data-driven distribution did not represent a new method for the legacy media industry, but effectively built on an existing model routinely enacted by studios and networks. For example, studios have long experimented with preview screenings and platform releases to target fans and open a film through word-of-mouth. Similarly, networks have consistently premiered new shows in high-traffic time slots, pairing them with popular shows in order to capitalize on existing viewership. The integration of data application reconfigured this approach by introducing automation into the mix, facilitating broader scale and advanced targeting to conceptually increase efficiencies within legacy media distribution.

While unified in their overall approach to data-driven distribution, selecting and targeting high-value audience segments to maximize viewership and revenue during the primary window, data application companies pursued a variety of strategic executions across the distribution value chain. For instance, the television business saw an emerging trend toward social TV, with social media platforms introducing new features to facilitate online conversations about specific

television shows and thereby drive appointment viewing, effectively enabling an expansive co-viewing environment where viewers could simultaneously watch and discuss TV in real time. The micro-blogging platform Twitter, in particular, emerged as a hub for second-screen television engagement, enabling users to follow television-related conversations in real time. Indeed, Twitter made a case that live usage of its platform during show airings could increase audience engagement, cultivate fan loyalty, and boost ratings,<sup>464</sup> frequently pointing to ABC's *Scandal* (2012-2018) as a prime example of a show that excelled at "must-tweet TV."<sup>465</sup> In addition to partnerships, networks further set out to develop their own social TV applications with the prospect of deepening audience engagement with live TV viewing.<sup>466</sup>

Meanwhile, in the film business, studios increasingly invested in "curating audiences"<sup>467</sup> to optimize the economics of the theatrical window, partnering with data application companies to segment, select, and target active movie-goers. For example, Paramount partnered with Eventful, a crowdsourcing company, to enable regional audiences to digitally demand the release of *Paranormal Activity* (2009), with Eventful measuring demand and identifying regions with high viewer engagement to roll out the film's distribution. Furthermore, studios cultivated partnerships with data application companies specializing in audience targeting. For instance, New Zealand-based company Movio developed a repository of 15 million active movie-goers

---

<sup>464</sup> See, Wagner, Kurt, "Twitter Says It's Time for Must-Tweet TV," *Recode*, September 18, 2014, accessed June 6, 2020, <https://www.vox.com/2014/9/18/11631024/twitter-to-tv-networks-live-tweeting-may-boost-your-ratings>

<sup>465</sup> McNamara, Mary, "'Scandal' has become must-tweet TV," *Los Angeles Times*, May 11, 2013, accessed June 6, 2020, <https://www.latimes.com/entertainment/tv/la-xpm-2013-may-11-la-et-st-scandal-abc-social-media-20130511-story.html>

<sup>466</sup> See, Shaw, Lucas, "Showtime Goes Social, Launches New 'Second Screen' App," *The Wrap*, September 26, 2011, accessed June 6, 2020, <https://www.thewrap.com/showtime-goes-social-launches-new-second-screen-app-31308/>

<sup>467</sup> Tryon, 2013

from theater loyalty programs to help studios identify addressable audience segments for specific films. Theater owners followed a similar approach with the launch of proprietary ticketing apps and loyalty programs, using audience information to optimize ticket sales and revenue per screening (e.g., AMC Stubs). Finally, companies like MoviePass moved the industry from a transactional to an on-demand distribution model, enabling users to see a number of films in theaters for a monthly subscription fee, using viewing behavior to recommend new titles.

By partnering with data application companies across the transactional distribution value chain, the majors made use of data-driven, algorithmic strategies to maximize the traditional release window in film and television. Specifically, studios and networks relied on data application software to select (by way of analysis), target (by way of distribution marketing), and engage (by way of distribution transactions) audience segments whose viewing behavior and content preferences matched new programming releases. In this sense, the legacy media industry's integration of data application worked to amplify the return on investment of the window, moving from a broad distribution approach to a more customized model that automatically adjusts rollout patterns based on audience information input. While data-driven distribution enabled the majors to drive increased value for incremental programming releases, the overall business model ultimately proved too selective, niche, and fickle to impact overall industry economics. Indeed, the phenomenon of social TV and cinema subscriptions experienced obstacles in overcoming the ongoing shift to on-demand viewing. Social TV effectively gave way to an increase in focused video-on-demand viewing and while online ticketing has continued to grow in popularity, the vast majority of ticket transactions still occur in-person at the box office.<sup>468</sup>

---

<sup>468</sup> See, Fritz, Ben & Schwartzel, Erich, "Online Tickets Seek Bigger Role at the Movies," *Wall Street Journal*, February 12, 2015, accessed June 6, 2020, <https://www.wsj.com/articles/online-tickets-seek-bigger-role> and White

--- --- ---

In navigating the dynamics of a new media environment, Hollywood has made incremental efforts to apply data to the cultural production value chain as a way to preserve and optimize the economics of the legacy media industry. In particular, film studios and television networks have implemented data-driven, algorithmic strategies across the processes of production, marketing, and distribution in an overarching attempt to minimize uncertainty, boost performance, and increase return on investment, thereby preserving the pressured tenets of Hollywood's traditional business model. Yet, while the majors' tolerance toward data application experienced a notable uptick, driven by an ongoing economic, cultural, and industrial convergence with the technology industry that saw new talent and technologies enter the media and entertainment industry, the integration of data and algorithms primarily functioned as a way to reconfigure, rather than redefine, the legacy media business. Indeed, the majors drew on data application companies as an additional layer in cultural production in an effort to enhance traditional practices of producing, marketing, and distributing content.

- *Data-Driven Production*: The majors developed partnerships with a suite of data application companies to automate the key elements of the production process. Specifically, they worked to select ideas, stories, and talent with a high probability of increasing the profitability of film and television programming.
- *Data-Driven Marketing*: The majors partnered with data application companies to enhance their marketing campaign outreach, implementing data-driven, algorithmic

---

Paper, "From Discovery to Purchase - The Moviegoing Experience Begins Online - A Box Office Profile," *Webedia & Vertigo*, 2019, quoted in *Forbes*, 2019.

marketing strategies to identify highly interested audience segments to drive engagement and increase profits.

- *Data-Driven Distribution*: The majors engaged with data application companies in order to maximize television tune-in and movie-going turnout during the traditional content release window.

The integration of data application reflected Hollywood's ongoing negotiation of shifting media dynamics. The majors relied on data-driven, algorithmic automation as a way to steer the legacy media business by attempting to mitigate risk, course-correct established processes, and optimize outcomes. Yet, the majors' investment in data application to preserve the status quo was neither static or inflexible. Rather, it formed part of a multi-dimensional process of adapting to the ever-expanding complexities of the digital economy. While the legacy media industry has been described as slow in its response to the exponential rise of digital media, the majors have consistently undertaken experiments to operate in a more data-driven way. This ongoing period of experimentation was marked by varying degrees of success that are difficult to account for holistically. Overall, the majors incubated, implemented, and iterated new forms of decision-making, operational practices, and executional strategies that more or less balanced the demands of the legacy media business with the dynamics of the digital economy.

Ultimately, Hollywood's experimental application of data across cultural production may not have succeeded in preserving the underlying economics of the legacy media business. Yet, the majors' exploration of data-driven production, marketing, and distribution worked to reconfigure established legacy processes, effectively leveraging data as an operating logic to produce, market, and distribute content in alternative and advanced ways. The industry's

investment in data application thereby signaled an industrial shift toward holistic data integration, a fully-integrated practice of accessing, analyzing, and applying data across the industrial value chain. As the legacy media industry ventured forward into a new era of Internet-distributed programming and digital business dynamics, incubating, launching, and managing their own platform ecosystems, the majors entered a new stage of cultural production, where data not only dictates, but defines, drives, and determines the way the industry operates.



## Conclusion

### The Iterative Industry

Over the course of the early 21st century, from the late 1990s to the late 2010s, the legacy media industry integrated data as an industrial operating logic that informed corporate decision-making and institutional practice in cultural production across film studios, television networks, and their conglomerate structures. This process of industrial integration did not result in a holistic, fully-integrated system, but a decentralized network of interconnected data centers that work independently and interdependently of one another, synthesizing external resources with internal developments to manage a data-driven industrial mode of operation. Indeed, while the majors adopted a set of data-driven practices that increasingly informed the way the industry operates, the overall lingering reliance on the dynamics of the legacy media business (e.g., the primacy of theatrical distribution, ad-supported linear broadcasting, and cable television programming) effectively positioned data as a supplementary industry driver, rather than a central driving force. As such, the legacy media industry effectively negotiated established and emerging modes of operation in engineering a more data-driven business practice.

Yet, more recently, the legacy media industry has taken additional steps to expand its business logic, laying a foundation for what might become a long-term industrial pivot. Indeed, leading media conglomerates executed large-scale investments in streaming platforms, positioning direct-to-consumer video services, previously largely deemed complementary and experimental, as the future business focus of the industry.<sup>469</sup> In the fall of 2019, the Walt Disney

---

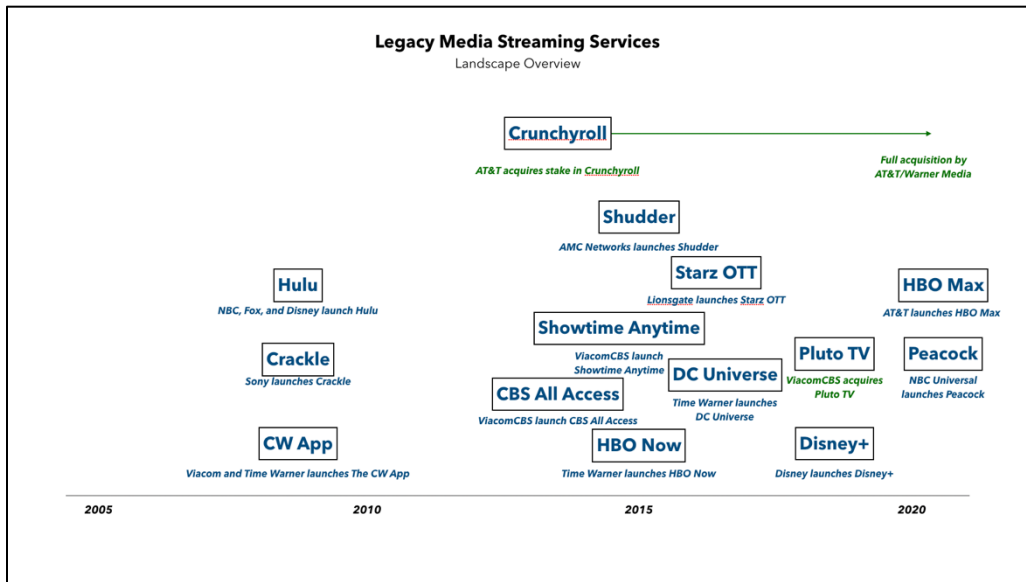
<sup>469</sup> See, Faughnder, Ryan, "Who will win the streaming wars?," *Los Angeles Times*, October 10, 2019, accessed June 6, 2020, <https://www.latimes.com/entertainment-arts/business/story/2019-10-10/streaming-wars-winners-and-losers-disney-plus-netflix-hbo-max-peacock-quirbi-apple-tv>

Company launched Disney+, a global subscription video-on-demand service designed to host all of its licensed programming as well as new original content.<sup>470</sup> In the spring of 2020, WarnerMedia, following its acquisition by telecommunications giant AT&T, rolled out HBO Max, a comprehensive subscription offering that includes the company's broad set of programming brands. And, that same summer, NBCUniversal released its advertising-supported video service Peacock. The companies also created dedicated direct-to-consumer business units, holistically focused on managing the new media business. Effectively, the legacy media industry built a system to scale the development of data-driven platform ecosystems<sup>471</sup> (figure 4.1).

---

<sup>470</sup> The company further bundled the offering with its Hulu streaming service following the acquisition of 21st Century Fox. See, Lucas Shaw & Christopher Palmeri, "Disney takes tighter control of Hulu," *Los Angeles Times*, February 3, 2020, accessed June 6, 2020, <https://www.latimes.com/entertainment-arts/business/story/2020-02-03/disney-takes-tighter-control-of-hulu>

<sup>471</sup> Given the slow pace of transformational change in the legacy media industry, the majors initially positioned direct-to-consumer as a long-term investment with a years-long development and integration roadmap, yet the unexpected implications of the COVID-19 pandemic that emerged in early 2020 revised the strategic timeline to effect increasingly short-term changes. The virus arguably accelerated and increased the pressure on the bottom-line economics of the legacy media business, with theaters closing and television advertising spend declining, prompting the majors to consider a shift to direct-to-consumer options, including premiering new movies on digital. See, Wallenstein, Andrew, "'Trolls World Tour' Marks Real Progress For Premium VOD. But It's Just The Beginning," *Variety*, April 30, 2020, accessed June 6, 2020, <https://variety.com/vip/trolls-world-tour-marks-real-progress-for-premium-vod-but-its-just-the-beginning-1234591645/>



**Figure 4.1:** The legacy media industry created a network of streaming services, with major conglomerates consolidating their offering into large-scale services to aggregate scale and control over data.

The shift to streaming effectively engineered, encoded, and established a digital interface with the audience through a centralized data infrastructure. Indeed, the majors expanded their traditional distribution infrastructure, from delivering content through analog intermediaries across legacy channels (i.e., theaters, television, home entertainment) to programming content directly for consumers through a visual interface, while simultaneously managing the legacy media business to avoid cannibalizing existing revenue streams. As such, film studios and television networks effectively merged into a generative content network for a digital platform system. In effect, the legacy media industry re-positioned its business model as a platform-enabled, data-driven direct-to-consumer economy.

In building out a dedicated platform business, the legacy media industry took initial steps to vertically integrate with the data value chain of Silicon Valley, with the majors developing a centralized infrastructure for accessing, analyzing, and applying data to consistently iterate the dynamics of the media business (table 4.1). As such, the majors effectively expanded beyond the core industry focus of pushing out content, making pulling in data a key priority of their

business. With the shift to streaming, Hollywood thus entered a new phase of operating the media business, one where data is no longer an incremental factor in the industrial logic, but a potentially existential differentiator. As emerging platforms, the major media conglomerates have come to operate in a non-linear digital environment where business dynamics are holistically defined, driven, and determined by data.

**Table 4.1**

<b>The Data Value Chain in Streaming</b>			
<b>Stage</b>	<b>Access</b>	<b>Analysis</b>	<b>Application</b>
<b>Process</b>	Access data on consumer interaction with content on the streaming service	Analyze consumer behavior and identify preferences	Apply algorithmic procedures to program the streaming service

With this project, I set out to document Hollywood’s complex convergence with Silicon Valley’s platform economy as a non-linear process of software-based industrial strategies to access, analyze, and apply data in an overarching effort to negotiate the dynamics of a new media environment, sketching the emergence of data as an industrial operating logic in the legacy media industry. In particular, I examined how the major film studios and television networks adopted and adapted data-driven practices to manage change at the economic, cultural-organizational, and technological level. In doing so, I demonstrated that the industry actively developed targeted strategies and responded to surrounding developments in order to leverage data as an industry driver across the media and entertainment value chain.

Over the course of the digital era, data has effectively become a core logic of Hollywood. My research has traced a suite of industrial scenarios where software enables, encodes, and enacts data-driven strategies that inform the way the industry operates across production, marketing, and distribution. At the same time, I highlighted some of the limitations that shape the

implementation of data at the economic, cultural, and technological level, which can be attributed to various factors, from entrenched operational dynamics to organizational bottlenecks and gatekeeper resistance. Overall, I argue that the idea of data as logic is an evolving site of industrial negotiation that manifests across a complex spectrum of corporate actions and use cases. It is clear that for Hollywood's legacy media industry, data has never been the *sine qua non* that it is for Silicon Valley, a defining existential condition that drives the entire industrial sphere. Instead, I see it as an incremental function in an expansive digital operating system, gaining more prominence, functionality, and impact over time.

Looking ahead, as the legacy media industry increasingly appears to transform into a subset of the digital platform economy, data no longer informs, but increasingly *iterates* the way Hollywood operates. By embedding streaming services as core business units within their larger conglomerate structures, the majors have effectively created a centralized and holistic industrial data center, at least in concept, with the potential to automate data-driven procedures and processes across the entire organizational sphere. With total data access, the industry is likely to evolve the dynamics of data analysis and application outlined in this project.

Accordingly, future research into media industries can productively look at the development of Hollywood as an iterative industry and the implications of the data value chain in institutionalizing, rather than merely informing, new modes of operation, decision-making dynamics, and forms of cultural production, including new forms of cultural advocacy and resistance, economic models, software iterations, and organizational manifestations. Studying data as a part of media industries may no longer be an incremental inquiry. Rather, it needs to take the form of a key focus area that reconsiders, reexamines, and recasts established patterns of media industries scholarship, building on emerging analytical frameworks and disciplinary

models that dissect data's varied role in media and entertainment.<sup>472</sup> I hope this project can serve as a foundational layer in this work. The industry's ongoing industrial transformation under the impact of data will chart new economic, cultural, and technological pathways, while reflecting the conceptual frameworks, strategic practices, and industrial use cases observed in this project.

Data is no longer an industrial phenomenon that will define the future of Hollywood. Neither is it merely an extension of its past practice of audience research. It has become a central iteration of its present operating logic.

---

<sup>472</sup> See, Arsenault, Amelia H., "The datafication of media: Big data and the media industries," *International Journal of Media and Cultural Politics*, 13, 1, 2017 and Napoli, Philip M. & Roepnack, Axel, eds., *Big Data and Media Management* (New York: Routledge, 2018).

## Bibliography

a16z Podcast, "The Internet of Taste, Streaming Content to Culture," February 18, 2018, accessed October 5, 2019, <https://a16z.com/2018/02/18/content-culture-digital-sarandos-summit/>

Adalian, Josef, "Inside the Binge Factory," *Vulture*, June 10, 2018, accessed May 4, 2020, <https://www.theverge.com/2014/6/13/5808424/netflix-will-close-its-public-api-to-some-developers-in-november>

Adams, Dan, "Can this Boston startup predict films' success?," *Boston Globe*, October 22, 2016, accessed June 6, 2020, <https://www.bostonglobe.com/business/2016/10/21/can-this-boston-startup-predict-film-grosses/ib6zL7GR0Qj6FoKc7OpCRN/story.html>

Albarran, Alan B., ed., *The Social Media Industries* (New York: Routledge, 2013)

Allen, Robert Clyde & Annette Hill, eds., *The Television Studies Reader* (New York: Psychology Press, 2004)

Andrejevic, Mark, *Infoglut: How Too Much Information Is Changing the Way We Think and Know* (New York, Routledge, 2013)

Angwin, Hulia & Martin Peers, "AOL-Time Warner Megamerger Creates a Web, Media Behemoth," December 15, 2000, accessed June 5, 2020, <https://www.wsj.com/articles/SB976835137406328524>

Arsenault, Amelia H., "The datafication of media: Big data and the media industries," *International Journal of Media and Cultural Politics*, 13, 1, 2017

Aspen Institute Podcast, "The New Golden Age of Television," *Aspen Ideas to Go*, September 10, 2015, accessed October 5, 2019, <https://www.aspeninstitute.org/podcasts/aspen-ideas-to-go-podcast-katie-couric-netflix-new-age-television/>

Auletta, Ken, "Outside the Box," *The New Yorker*, January 27, 2014, accessed June 5, 2020, <https://www.newyorker.com/magazine/2014/02/03/outside-the-box-2>

Avirgan, Jody, "Podcast: The Guy Who Predicts Whether A Movie Will Bomb, Months Before It's Made," *Five Thirty Eight*, September 17, 2015, accessed June 6, 2020, <https://fivethirtyeight.com/features/podcast-the-guy-who-predicts-whether-a-movie-will-bomb-months-before-its-made/>

Baker, Chris, "Q&A: Movie Exec Thomas Tull's Journey From Wall Street to Hollywood," *Wired*, October 19, 2009, accessed June 6, 2020, <https://www.wired.com/2009/10/ff-qa-tull/>

Bakker, Gerben, "Building knowledge about the consumer: The emergence of market research in the motion picture industry," *Business History*, 45, 1, 2003

Balio, Tino, *Hollywood in the New Millennium* (London: British Film Institute, 2013)

“” “”, “Adjusting to the New Global Economy: Hollywood in the 1990s.” In: *Film Policy: International National, and Regional Perspectives*, edited by Albert Moran (New York: Routledge, 1996)

“” “”, *The American Film Industry* (Madison: University of Wisconsin Press, 1985)

Ball, Matthew, “Letting it Go: The End of Windowing (and What Comes Next),” *Redef*, August 24, 2016, accessed October 5, 2019, <https://redef.com/original/letting-it-go-the-end-of-windows-and-what-comes-next>

“” “”, “How YouTube MCNs are Conquering Hollywood,” *Redef*, December 15, 2014, accessed October 5, 2019, <https://redef.com/original/how-youtube-mcns-are-conquering-hollywood>

Bannister, Kristian, “Understanding Sentiment Analysis: What It Is & Why It’s Used”, *Brandwatch*, February 26, 2018, accessed June 6, 2020, <https://www.brandwatch.com/blog/understanding-sentiment-analysis/>

Barnes, Brooks, “”Attacked By Rotten Tomatoes,” *New York Times*, September 7, 2017, accessed June 6, 2020, <https://www.nytimes.com/2017/09/07/business/media/rotten-tomatoes-box-office.html>

“” “”, “A Movie Ticketing Start-Up Hopes to Fill Empty Seats,” *New York Times*, December 4, 2016, accessed June 6, 2020, <https://www.nytimes.com/2016/12/04/business/media/movie-ticketing-start-up-atom-tickets-hopes-to-fill-theaters.html>

“” “”, “Hollywood Tracks Social Media Chatter to Target Hit Films,” *New York Times*, December 7, 2014, accessed June 6, 2020, <https://www.nytimes.com/2014/12/08/business/media/hollywood-tracks-social-media-chatter-to-target-hit-films.html>

“” “”, “Disney Buys Maker Studios, Video Supplier for YouTube,” *New York Times*, March 24, 2014, accessed June 5, 2020, <https://www.nytimes.com/2014/03/25/business/media/disney-buys-maker-studios-video-supplier-for-youtube.html>

“” “”, “Save My Blockbuster,” *New York Times*, June 28, 2013, accessed June 6, 2020, <https://archive.nytimes.com/www.nytimes.com/interactive/2013/06/28/movies/BLOCKBUSTER.html/>

“” “”, “Solving Equation for a Hit Film Script, with Data,” *New York Times*, May 5, 2013, accessed June 6, 2020, <https://www.nytimes.com/2013/05/06/business/media/solving-equation-of-a-hit-film-script-with-data.html>



“””, “A-Listers, Meet Your Online Megaphone,” *New York Times*, November 10, 2012, accessed June 6, 2020, <https://www.nytimes.com/2012/11/11/business/oliver-luckett-of-theaudience-building-online-fan-bases.html>

Basin, Ken, *The Business of Television* (New York: Routledge, 2018)

Belloni, Matthew, “Bob Iger Talks Disney's Streaming Service,” *Hollywood Reporter*, September 20, 2018, accessed June 5, 2020, <https://www.hollywoodreporter.com/news/bob-iger-disneys-streaming-service-james-gunn-star-wars-slowdown-1145493>

Bereznak, Alyssa, “The ‘Blair Witch’ Extended Universe: How a Tiny Indie Film Became a Horror Sensation—and Invented Modern Movie Marketing,” *The Ringer*, March 28, 2019, accessed June 6, 2020, <https://www.theringer.com/movies/2019/3/28/18280988/blair-witch-movie-marketing-1999>

Berry, David M., *The Philosophy of Software: Code and Mediation in the Digital Age* (Landon: Palgrave, 2011)

Bicknell, Craig, “Disney Buys Into Infoseek,” *Wired*, June 18, 1998, accessed June 5, 2020, <https://www.wired.com/1998/06/disney-buys-into-infoseek/>

Bilton, Nick, “Why Hollywood As We Know It Is Already Over,” *Vanity Fair*, January 29, 2017, accessed October 5, 2019, <https://www.vanityfair.com/news/2017/01/why-hollywood-as-we-know-it-is-already-over>

Bilton, Chris, “The Disappearing Product and the New Intermediaries.” In: *Making Media*, edited by Mark Deuze and Mirjam Prenger (Amsterdam: Amsterdam University Press, 2019)

Bishop, Bryan, “From alpha to ‘Betas’: how Amazon is rethinking the way television is made,” *The Verge*, November 26, 2013, accessed June 6, 2020, <https://www.theverge.com/2013/11/26/5147796/betas-how-amazon-is-rethinking-the-way-television-is-made>

Block, Alex Ben, “Ryan Kavanaugh’s secret to success,” *Hollywood Reporter*, September 29, 2010, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/ryan-kavanaughs-secret-success-28540>

Bolter, Jay David & Grusin, Richard, *Remediation: Understanding New Media* (Cambridge: MIT Press, 2000)

Boone, Christopher, “Paint by Numbers? Hollywood Rewriting Scripts Based on Statistical Analysis to Boost Box Office,” *No Film School*, May 7, 2013, accessed June 6, 2020, <https://nofilmschool.com/2013/05/hollywood-rewrite-script-statistical-analysis-box-office>

Borgman, Christine L., *Big data, little data, no data: scholarship in a networked world* (Cambridge: MIT Press, 2015)

Box Office Mojo, "Domestic Yearly Box Office," Date as of June 5, 2020, accessed June 5, 2020, <https://www.boxofficemojo.com/year/>

Boyd, Joshua, "Brand Health and How to Measure It," *Brandwatch*, November 6, 2017, accessed June 6, 2020, <https://www.brandwatch.com/blog/brand-health-how-to-measure/>

Bratton, Benjamin H., *The Stack* (Cambridge: MIT Press, 2016)

Brookey, Robert Alan, *Hollywood Gamers: Digital Convergence in the Film and Video Game Industries* (Bloomington: Indiana University Press, 2010)

Bruell, Alexandra, "Viacom Acquires Whosay to Help Advertisers Create More Branded Content," *Wall Street Journal*, January 8, 2018, accessed June 6, 2020, <https://www.wsj.com/articles/viacom-acquires-whosay-to-help-advertisers-create-more-branded-content-1515420000>

Bruns, Axel, *Blogs, Wikipedia, Second Life, and Beyond: From Production to Producersage* (Frankfurt am Main: Peter Lang, 2008)

Buchanan, Kyle, "How Will Movies Survive The Next 10 Years?," *New York Times*, June 20, 2019, accessed June 6, 2020, <https://www.nytimes.com/interactive/2019/06/20/movies/movie-industry-future.html>

Burroughs, Benjamin, "House of Netflix: Streaming media and digital lore," *Popular Communication, The International Journal of Media and Culture*, Volume 17, Issue 1, 2019

Buzzard, Karen, *Tracking the Audience: The Ratings Industry From Analog to Digital* (New York: Routledge, 2012)

Caldwell, John T., "Para-Industry, Shadow Academy," *Cultural Studies*, Volume 28, Issue 4, 2014

"", *Production Culture: Industrial Reflexivity and Critical Practice in Film and Television* (Durham: Duke University Press, 2008)

"", "Second-Shift Aesthetics: Programming, Interactivity, and User Flows," *New Media: Theories and Practices of Digitextuality*, edited by Ann Everett & John Caldwell (New York: Routledge: 2003)

Calore, Michael, "Flixster: Social Networking for Movie Nuts," *Wired*, November 6, 2006, accessed June 6, 2020, <https://www.wired.com/2006/11/flixster-social-networking-for-movie-nuts>

Campbell-Kelly, Martin, *From Airline Reservations to Sonic the Hedgehog: A History of the Software Industry* (Cambridge: MIT Press, 2004)

Canepa, Steve, “Big Data Is Going To Save The Film Industry, Too,” *Business Insider*, May 12, 2013, accessed June 6, 2020, <https://www.businessinsider.com/the-film-industry-needs-big-data-2013-5>

Carlson, Nicholas, “Facebook Reaches 500 Million Users,” *Business Insider*, May 17, 2020, accessed June 6, 2020, <https://www.businessinsider.com/facebook-reaches-500-million-users-2010-5>

Carr, David, “TV Foresees Its Future. Netflix Is There.,” *New York Times*, July 21, 2013, accessed June 5, 2020, <https://www.nytimes.com/2013/07/22/business/media/tv-foresees-its-future-netflix-is-there.html>

“” “”, “Giving Viewers What They Want,” *New York Times*, February 24, 2013, accessed June 5, 2020, <https://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html>

Christian, Aymar Jean, *Open TV: Innovation Beyond Hollywood and the Rise of Web Television* (New York: New York University, Press, 2018)

Ciamprone, Danny, “Warner Bros.’s Brian Kursar is the Data Knight,” *Sync*, March 15, 2017, accessed June 6, 2020, <https://sync-magazine.com/2017/wb/>

Cieply, Michael, “Hollywood Wants Numbers on the Digital Box Office,” *New York Times*, September 15, 2013, accessed June 6, 2020, <https://www.nytimes.com/2013/09/16/business/media/movie-industry-wants-to-get-a-handle-on-the-digital-box-office.html>

Clifford, Tyler, “Bob Iger on Not Buying Twitter,” *CNBC*, September 24, 2019, accessed June 5, 2020, <https://www.cnbc.com/2019/09/24/disney-bob-iger-on-not-buying-twitter-i-got-cold-feet.html>

CNN Press Room, “Michael Lynton, Chairman & CEO of Sony Pictures Entertainment with CNN’s Fareed Zakaria,” December 19, 2014, accessed June 4, 2020, <https://cnnpressroom.blogs.cnn.com/2014/12/19/michael-lynton-chairman-ceo-of-sony-pictures-entertainment-with-cnns-fareed-zakaria/>

Condliffe, Jamie, “The average American spends 24 hours per week online,” *MIT Technology Review*, January 23, 2018, accessed June 5, 2020, <https://www.technologyreview.com/f/610045/the-average-american-spends-24-hours-a-week-online>

Crucchiola, Jordan, “Charting *The Blair Witch Project*’s Influence Through 10 Horror Films That Followed,” *Vulture*, September 16, 2016, accessed June 5, 2020, <https://www.vulture.com/2016/09/10-horror-movies-inspired-by-the-blair-witch-project.html>

“” “”, “Eli Roth Has a Data-Happy Vision for Cinema's Future,” *Wired*, September 25, 2015, accessed June 6, 2020, <https://www.wired.com/2015/09/eli-roth-data-crunching/>

Cunningham, Stuart, Flew, Terry & Swift, Adam, *Media Economics* (London: Palgrave, 2015)

Cunningham, Todd, “Universal Bets Perky ‘Pitch Perfect’ Will Keep Box Office Beat,” *The Wrap*, October 4, 2012, accessed June 6, 2020, <https://www.thewrap.com/universal-bets-perky-pitch-perfect-keeps-box-office-beat-58856/>

Curtin, Michael, Holt, Jennifer & Sanson, Kevin, *Distribution Revolution: Conversations about the Digital Future of Film and Television* (Berkeley: University of California Press, 2015)

Cusumano, Michael A., Gawer, Annabelle & Yoffie, David B., *The Business of Platforms: Strategy in the Age of Digital Competition, Innovation, and Power* (New York: Harper Business, 2019)

“” “”, “A Study of More Than 250 Platforms Reveals Why Most Fail,” *Harvard Business Review*, May 29, 2019, accessed June 5, 2020, <https://hbr.org/2019/05/a-study-of-more-than-250-platforms-reveals-why-most-fail>

Davenport, Thomas H. & D.J. Patil, “Data Scientist: The Sexiest Job of the 21st Century,” *Harvard Business Review*, October 2012, accessed October 5, 2019, <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

Deighton, John A. & Leora Kornfeld, “Legendary Entertainment: Moneyball for Motion Pictures (Case Study),” *Harvard Business School*, May 2016

De Vany, Arthur S., *Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry* (New York: Routledge, 2003)

Dickey, Josh, “Movie trailers have an effective new strategy you might not have noticed,” *Mashable*, June 19, 2017, accessed June 6, 2020, <https://mashable.com/2017/06/19/movies-trailers-bumpers-online-pre-roll/>

“” “”, “Bizzers: Execs must master data analytics,” *Variety*, November 29, 2012, accessed June 6, 2020, <https://variety.com/2012/film/news/bizzers-execs-must-master-data-analytics-1118062878/>

Dietz, Jason, “House of Cards: Reviews for the Complete 1st Season,” *Metacritic*, February 20, 2013, accessed June 6, 2020, <https://www.metacritic.com/feature/house-of-cards-netflix-full-season-1-reviews>

Digiacomio, Frank, “The Theory of Relativity,” *Vanity Fair*, February 9, 2010, accessed June 6, 2020, <https://www.vanityfair.com/news/2010/03/kavanaugh-201003>

Disney Institute Blog, “A Closer Look at the Disney Data & Analytics Conference,” *Disney Institute*, January 24, 2020, accessed June 6, 2020, <https://www.disneyinstitute.com/blog/a-closer-look-at-the-disney-data-analytics-conference/>

Dodes, Rachel, “Twitter Goes to the Movies,” *The Wall Street Journal*, August 3, 2012, accessed May 5, 2020, <https://www.wsj.com/articles/SB10000872396390443343704577553270169103822>

Dolliver, Mark, “US Time Spent with Media 2019,” eMarketer, May 30, 2019, accessed June 6, 2020, <https://www.emarketer.com/content/us-time-spent-with-media-2019>

Drake, Philip, “Distribution and Marketing in Contemporary Hollywood.” In: *The Contemporary Hollywood Film Industry*, edited by Paul McDonald and Janet Wasko (Hoboken: Wiley-Blackwell, 2008)

Draper, Nora, “Fail Fast: The Value of Studying Unsuccessful Technology Companies,” *Media Industries Journal*, Volume 4, Issue 1, 2017

Duan, Lian & Xiong, Ye “Big data analytics and business analytics,” *Journal of Management Analytics*, Volume 2, Issue 1, 2015

The Economist, “The world’s most valuable resource is no longer oil, but data,” May 6, 2017, accessed June 6, 2020, <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>

“” “”, “Hollywood and its audience look to the Internet,” January 11, 2010, accessed June 6, 2020, <https://www.economist.com/news/2010/01/11/a-happy-ending>

“” “”, “Who’s afraid of AOL Time Warner?,” *Media Giants*, January 24, 2002, accessed June 5, 2020, <https://www.economist.com/business/2002/01/24/whos-afraid-of-aol-time-warner>

Elberse, Anita, *Blockbusters: Hit-making, Risk-taking, and the Big Business of Entertainment* (New York: Henry Holt, 2013)

Epstein, Edward Jay, *The Hollywood Economist: The Hidden Financial Reality Behind The Movies* (New York: Melville House, 2012)

Facebook, “Facebook Unveils Facebook Ads,” November 6, 2007, accessed June 5, 2020, <https://about.fb.com/news/2007/11/facebook-unveils-facebook-ads/>

Fahey, Mark, “Why movies are sometimes here and gone in theaters,” CNBC, November 18, 2015, accessed June 6, 2020, <https://www.cnbc.com/2015/11/17/why-movies-are-sometimes-here-and-gone-in-theaters.html>

Farber, Dan, “Amazon Studios debuts 14 pilots for free viewing,” *Cnet*, April 19, 2013, accessed June 6, 2020, <https://www.cnet.com/news/amazon-studios-debuts-14-pilots-for-free-viewing/>

Faughnder, Ryan, "Who will win the streaming wars?," *Los Angeles Times*, October 10, 2019, accessed June 6, 2020, <https://www.latimes.com/entertainment-arts/business/story/2019-10-10/streaming-wars-winners-and-losers-disney-plus-netflix-hbo-max-peacock-quirbi-apple-tv>

"" "", "This movie-ticket startup says it can fill empty theaters across the country," *Los Angeles Times*, June 28, 2016, accessed June 6, 2020, <https://www.latimes.com/entertainment/envelope/cotown/la-et-ct-atom-tickets-20160620-snap-story.html>

"" "", "Movie trailer makers multiply as online viewing of previews soars," *Los Angeles Times*, July 21, 2015, accessed June 6, 2020, <https://www.latimes.com/entertainment/envelope/cotown/la-et-ct-hollywood-trailers-20150721-story.html>

Flew, Terry, "Social Media and the Cultural and Creative Industries." In: *The SAGE Handbook of Social Media*, edited by Jean Burgess, Alice Marwick, and Thomas Poell (Thousand Oaks: Sage, 2017)

Fritz, Ben, *The Big Picture: The Fight for the Future of Movies* (Boston: Houghton Mifflin Harcourt, 2018)

Fritz, Ben & Erich Schwartzel, "Online Tickets Seek Bigger Role at the Movies," *Wall Street Journal*, February 12, 2015, accessed June 6, 2020, <https://www.wsj.com/articles/online-tickets-seek-bigger-role-at-the-movies-1423782268>

Fuller, Matthew, *Software Studies: A Lexicon* (Cambridge: MIT Press, 2008)

"" "", *Media Ecologies: Materialist Energies in Art and Technoculture* (Cambridge: MIT Press, 2007)

Furrier, John, "Hollywood: Is 'Godzilla' The 'John Carter' Of 2014?," *Forbes*, January 23, 2014, accessed June 6, 2020, <https://www.forbes.com/sites/siliconangle/2014/01/23/hollywood-is-godzilla-the-john-carter-of-2014/#115c1968301c>

Galloway, Alexander, *The Interface Effect* (Cambridge, UK: Polity, 2012)

Gaudiosi, John, "Legendary Pictures CEO talks tech that gave 'Godzilla' its box-office roar," *Fortune*, May 22, 2014, accessed June 6, 2020, <https://fortune.com/2014/05/22/legendary-pictures-ceo-talks-tech-that-gave-godzilla-its-box-office-roar/>

Garrahan, Matthew, "Producer Follows His Own Script," *Fortune*, December 5, 2010, accessed June 6, 2020, <https://www.ft.com/content/96d90b5c-00aa-11e0-aa29-00144feab49a>

Gates, Bill, "Bill Gates' New Rules," *Time*, April 19, 1999, accessed June 6, 2020, <http://content.time.com/time/world/article/0,8599,2053895,00.html>

Gillan, Jennifer, *Must-Click TV* (New York: Routledge, 2010)

Gillespie, Tarleton, “#trendingistrending: When Algorithms Become Culture.” In: *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies*, edited by Robert Seyfert and Jonathan Roberge (London: Routledge, 2016)

“” “”, “The Relevance of Algorithms.” In: *Media Technologies*, edited by Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot (Cambridge: MIT Press, 2014)

“” “”, “Algorithm [draft] [#digitalkeywords],” *Culture Digitally*, June 25, 2014, accessed June 5, 2020, <http://culturedigitally.org/2014/06/algorithm-draft-digitalkeyword/>

Gilmore, James N. & Matthias Stork, *Superhero Synergies: Comic Book Characters Go Digital* (Lanham, MD: Scarecrow Press, 2014)

Gitelman, Lisa, ed., “*Raw Data*” Is An Oxymoron (Cambridge: MIT Press, 2013)

“” “”, *Always Already New: Media, History, and the Data of Culture* (Cambridge: MIT Press, 2008)

Gitlin, Todd, *Inside Prime Time* (Berkeley: University of California Press, 2000)

Godley, Chris, “THR's Social Media Poll: How Facebook and Twitter Impact the Entertainment Industry,” *Hollywood Reporter*, March 21, 2012, accessed June 6, 2020, <https://www.hollywoodreporter.com/gallery/facebook-twitter-social-media-study-302273/1-social-media-as-entertainment>

Goldman, William, *Adventures in the Screen Trade: A Personal View of Hollywood and Screenwriting* (New York: Warner Books/Grand Central Publishing, 1983)

Google, “Our Approach to Search,” accessed June 6, 2020, <https://www.google.com/search/howsearchworks/mission/>

“” “”, “How Search Organizes Information,” accessed June 6, 2020, <https://www.google.com/search/howsearchworks/crawling-indexing/>

“” “”, “Google Agrees to Acquire Urchin,” May 28, 2005, accessed June 5, 2020, [http://googlepress.blogspot.com/2005/03/google-agrees-to-acquire-urchin\\_28.html](http://googlepress.blogspot.com/2005/03/google-agrees-to-acquire-urchin_28.html)

Grainge, Paul & Johnson, Catherine, *Promotional Screen Studies* (New York: Routledge, 2015)

Grainge, Paul, *Brand Hollywood: Selling Entertainment in a Global Media Age* (New York: Routledge, 2007)

Graser, Marc, "Bob Iger Explains Why Disney Bought Maker Studios," *Variety*, May 6, 2014, accessed June 5, 2020, <https://variety.com/2014/biz/news/bob-iger-explains-why-disney-bought-maker-studios-1201173389/>

Gray, Jonathan, "Reviving audience studies," *Critical Studies in Media Communication*, Volume 34, Issue 1, 2017

"" "", *Show Sold Separately: Promos, Spoilers, and Other Media Paratexts* (New York: New York University Press, 2010)

Grover, Ronald, "The Hollywood Numbers Game," *Bloomberg*, August 2, 2001, accessed June 6, 2020, <https://www.bloomberg.com/news/articles/2001-08-02/the-hollywood-numbers-game>

Gunther McGrath, Rita, "15 years later, lessons from the failed AOL-Time Warner merger," *Fortune*, January 10, 2015, accessed June 5, 2020, <https://fortune.com/2015/01/10/15-years-later-lessons-from-the-failed-aol-time-warner-merger/>

Handel, Leo A., "Hollywood Market Research," *The Quarterly of Film Radio and Television*, 7 (3), 1953, 304-310

Harris, Derrick, "Netflix analyzes a lot of data about your viewing habits," *Gigaom*, June 14, 2012, accessed June 5, 2020, <https://gigaom.com/2012/06/14/netflix-analyzes-a-lot-of-data-about-your-viewing-habits/>

Hartley, John, Burgess, Jean & Bruns, Axel, eds., *A Companion to New Media Dynamics* (New York: Wiley-Blackwell, 2015)

Hass, Nancy, "Is Netflix the Next HBO?," *GQ*, January 29, 2013, accessed August 20, 2019, <https://www.gq.com/story/netflix-founder-reed-hastings-house-of-cards-arrested-development>

Havens, Timothy & Amanda Lotz, *Understanding Media Industries* (Oxford: Oxford University Press, 2016)

Havens, Timothy, "Media Programming in an Era of Big Data," *Media Industries Journal*, Volume 1, Issue 2, 2014

Havens, Timothy, Lotz, Amanda D. & Serra Tinic, "Critical Media Industries Studies: A Research Approach," *Communication, Culture & Critique*, 2, 2009

Hayes, Dade, "Legendary Entertainment Selling Majority Stake In Its Applied Analytics Unit," *Deadline*, November 13, 2017, accessed June 6, 2020, <https://deadline.com/2017/11/legendary-entertainment-selling-majority-stake-in-its-applied-analytics-unit-1202206956/>

Hayes, Dade & Jonathan Bing, *Open Wide: How Hollywood Box Office Became a National Obsession* (New York: Miramax Books, 2004)



Hazelton, John, "Big film, big data: how analytics is shaping the business," *ScreenDaily*, June 16, 2016, accessed June 5, 2020, <https://www.screendaily.com/features/big-film-big-data-how-analytics-is-shaping-the-business/5104922.article>

He, Amy, "More Consumers Will Continue to Drop Pay TV Because of Price Hikes," *eMarketer*, August 8, 2019, accessed June 5, 2020, <https://www.emarketer.com/content/more-consumers-will-continue-to-drop-pay-tv-because-of-price-hikes>

Heritage, Stuart, "You've Got Mail: the forgotten world of 90s movie websites," *The Guardian*, March 15, 2017, accessed June 5, 2020, <https://www.theguardian.com/film/filmblog/2017/mar/15/official-film-websites-youve-got-mail-jurassic-park-space-jam>

Herrman, John & Mike Isaac, "The Online Video View: We Can Count It, but Can We Count on It?," *New York Times*, October 2, 2016, accessed June 6, 2020, <https://www.nytimes.com/2016/10/03/business/media/the-online-video-view-we-can-count-it-but-can-we-count-on-it.html>

Hesmondhalgh, David, *The Cultural Industries* (New York: Sage, 2019)

Hill, Annette & Janette Steemers, "Media Industries and Engagement," *Media Industries Journal*, Volume 4, Issue 1, 2017

Hilmes, Michele, *Only Connect: A Cultural History of Broadcasting in the United States* (Boston: Cengage Learning, 2010)

Hoad, Phil, "How we made The Blair Witch Project," *The Guardian*, May 21, 2018, accessed June 5, 2020, <https://www.theguardian.com/culture/2018/may/21/how-we-made-the-blair-witch-project>

Hod, Itay, "How Hollywood Actors' Twitter Followings Have Become as Important as Talent," *The Wrap*, March 10, 2015, accessed June 6, 2020, <https://www.thewrap.com/how-hollywood-actors-twitter-followings-have-become-as-important-as-talent/>

Hollander, J.B., Graves, E., Renski, H., Foster-Karim, C., Wiley, A., Das, D., "A (Short) History of Social Media Sentiment Analysis." In: *Urban Social Listening* (London: Palgrave Macmillan, 2016)

Holson, Laura M., "Warner Venture with Investors," *New York Times*, June 22, 2005, accessed June 6, 2020, <https://www.nytimes.com/2005/06/22/business/media/warner-venture-with-investors.html>

Holson, Laura M. & Lyman, Rick, "In Warner Bros. Strategy, A Movie Is Now a Product Line," *New York Times*, February 11, 2002, accessed June 6, 2020, <https://www.nytimes.com/2002/02/11/business/warner-brothers-strategy-movie-now-product-line-making-franchise-films-that.html>

Holt, Jennifer & Sanson, Kevin, eds., *Connected Viewing: Selling, Streaming, and Sharing Media in the Digital Age* (New York: Routledge, 2013)

Hutsko, Joe, "Behind the Scenes Via Movie Web Sites," *New York Times*, July 10, 2003, accessed June 6, 2020, <https://www.nytimes.com/2003/07/10/technology/behind-the-scenes-via-movie-web-sites.html>

IAB, "IAB Measurement Guidelines," *Internet Advertising Bureau*, accessed October 5, 2019, <https://www.iab.com/guidelines/iab-measurement-guidelines/>

IBM, "The Race to Probe the Twittersphere," *The Atlantic*, accessed June 6, 2020, <https://www.theatlantic.com/sponsored/ibm-transformation-of-business/the-race-to-probe-the-twittersphere/280/>

Ingram, Mathew, "The Real Secret to BuzzFeed's Success Isn't Cat Gifs, It's Data," *Fortune*, February 16, 2016, accessed June 5, 2020, <https://fortune.com/2016/02/16/buzzfeed-data/>

"", "Facebook vs. Twitter: An Infographic," *Gigaom*, December 20, 2010, accessed June 6, 2020, <https://gigaom.com/2010/12/20/facebook-vs-twitter-an-infographic/>

Jarvey, Natalie, "Legendary Hires Bankers to Sell Analytics Business," *Hollywood Reporter*, November 13, 2017, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/legendary-hires-bankers-sell-analytics-business-1057371>

Jeffries, Stuart, "Netflix's Ted Sarandos: the 'evil genius' behind a TV revolution," *The Guardian*, December 30, 2013, accessed June 5, 2020, <https://www.theguardian.com/media/2013/dec/30/netflix-evil-genius-tv-revolution-ted-sarandos>

Jenkins, Henry, Ford, Sam & Joshua Green, eds., *Spreadable Media: Creating Value and Meaning in a Networked Culture* (New York: New York University Press, 2013)

Jenkins, Henry, *Convergence Culture* (New York, New York University Press, 2008)

Johnson, Bradley, "In a New Milestone, the Internet Will Account for Half of Ad Spending in 2020," *AdAge*, December 23, 2019, accessed June 5, 2020, <https://adage.com/article/datacenter/new-milestone-internet-will-account-half-ad-spending-2020/2223511>

Johnson, Derek, ed., *From Networks to Netflix: A Guide to Changing Channels* (New York: Routledge, 2018)

Kafka, Peter, "How Netflix outsmarted everyone else in TV," *Vox*, August 23, 2018, accessed June 5, 2020, <https://www.vox.com/2018/8/23/17770896/netflix-reed-hastings-ted-sarandos-streaming-tv-media-jason-hirschhorn-redef-peter-kafka-podcast>

“” “”, “Full transcript: BTIG Analyst Rich Greenfield on Recode Media,” *Vox*, June 15, 2017, accessed April 3, 2019, <https://www.vox.com/2017/6/30/15904416/transcript-btig-analyst-rich-greenfield-recode-media-peter-kafka-podcast>

Kay, Jeremy, “Relativity Media fires up Gun Hill Road package,” May 11, 2006, accessed June 6, 2020, <https://www.screendaily.com/relativity-media-fires-up-gun-hill-road-package/4027164.article>

Kapko, Matt, “How a small film studio uses Facebook data to compete with Hollywood’s heavyweights,” *CIO*, February 9, 2017, accessed June 6, 2020, <https://www.cio.com/article/3167889/how-a-small-film-studio-uses-facebook-data-to-compete-with-hollywood-s-heavyweights.html>

Keating, Gina, *Netflixed: The Epic Battle for America’s Eyeballs* (New York: Portfolio, 2012)

Keith, Barry, ed., *Film Genre Reader IV* (Austin: University of Texas Press, 2012)

Kelly, Kate, “Defying the Odds, Hedge Funds Bet Billions on Movies,” *Wall Street Journal*, April 29, 2006, accessed June 6, 2020, <https://www.wsj.com/articles/SB114627404745739525>

Kenneally, Tim, “AMC Networks Chief Josh Sapan Says You Can’t ‘Science Your Way’ to Great Content,” *The Wrap*, October 7, 2014, accessed June 5, 2020, <https://www.thewrap.com/the-grill-amc-networks-chief-josh-sapan-says-you-cant-science-your-way-to-great-content/>

Kerrigan, Finola, *Film Marketing* (New York: Routledge, 2009)

Khatchaturian, Maane, “‘Godzilla’ Crushes Box Office With Largest Opening Day of Year, Set for \$98 Million Weekend,” *Variety*, May 17, 2014, accessed June 6, 2020, <https://variety.com/2014/film/box-office/godzilla-98-million-weekend-opening-1201184456/>

Kirkpatrick, Scott, *Introduction to Media Distribution: Film, Television, and New Media* (New York: Routledge, 2018)

Kirsner, Scott, “Making movies the ‘Moneyball’ way,” *Boston Globe*, March 31, 2016, accessed June 6, 2020, <https://www.bostonglobe.com/business/technology/2016/03/31/making-movies-moneyball-way/Uzgwh2cdGthA1N3nZHqz0N/story.html>

Klein, Jason, “Digital Audience Ratings” *Variety*, October 1, 2015, accessed June 6, 2020, <https://variety.com/2015/digital/news/digital-audience-ratings-strong-showing-for-sony-as-family-favorites-top-the-week-1201606163/>

Knowledge@Wharton, “How Viacom Sparked Its Digital and Cultural Transformation,” *Wharton School of Business*, February 28, 2020, accessed June 6, 2020, <https://knowledge.wharton.upenn.edu/article/viacom-sparked-digital-cultural-transformation/>

“” “”, “What’s Driving the Demand for Data Scientists?,” *Wharton School of Business*, March 8, 2019, accessed June 6, 2020, <https://knowledge.wharton.upenn.edu/article/whats-driving-demand-data-scientist/>

“” “”, “Can Hollywood Survive Streaming Services?,” *Wharton School of Business*, May 4, 2018, accessed June 4, 2020, <https://knowledge.wharton.upenn.edu/article/can-hollywood-survive-streaming-services/>

“” “”, “Following Intellectual Curiosity with Thomas Tull,” *The Knowledge Project with Shane Parrish*, May 28, 2019, accessed June 6, 2020, <https://fs.blog/knowledge-project/thomas-tull/>

Krigsman, Michael, “Moneyball for movies: Data science and AI in Hollywood,” *ZDNet*, February 20, 2018, accessed June 6, 2020, <https://www.zdnet.com/article/moneyball-for-movies-data-science-and-ai-in-hollywood/>

Kuntz, William M., *Culture Conglomerates: Consolidation in the Motion Picture and Television Industries* (Lanham: Rowman & Littlefield, 2007)

Kwek, Nick, “Most of big data is 'trash' says Netflix's Todd Yellin,” *BBC News*, April 21, 2016, accessed June 6, 2020, <https://www.bbc.com/news/av/technology-36093007/most-of-big-data-is-trash-says-netflix-s-todd-yellin>

Lang, Brent, “‘Godzilla’ Stomps to \$36 Mil Debut in China,” *Variety*, June 15, 2014, accessed June 6, 2020, <https://variety.com/2014/film/news/godzilla-china-box-office-1201220922/>

“” “”, “Big Media Gambling on Future Digital Riches (Study),” *The Wrap*, January 28, 2014, accessed June 6, 2020, <https://www.thewrap.com/media-entertainment-companies-will-lose-money-set-digital-future-study/>

Laporte, Nicole, “Pitch Perfect: How Universal’s Digital Marketing Helped It Have The Best Year Ever,” *Fast Company*, September 11, 2015, accessed June 6, 2020, <https://www.fastcompany.com/3050984/pitch-perfect-how-universals-digital-marketing-helped-it-have-the-best-year-ever>

“” “”, “Netflix: The Red Menace,” *Fast Company*, January 7, 2014, accessed June 5, 2020, <https://www.fastcompany.com/3024158/netflix-the-red-menace>

LaValle, Steve, Lesser, Eric Shockley, Rebecca, Hopkins, Michael S. & Nina Kruschwitz, “Big Data, Analytics and the Path From Insights to Value,” *MIT Sloan Management Review*, December 21, 2010, accessed June 6, 2020, <https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/>

Leber, Jessica, “‘House of Cards’ and Our Future of Algorithmic Programming,” *MIT Technology Review*, February 26, 2013, accessed June 5, 2020,

<https://www.technologyreview.com/2013/02/26/16555/house-of-cards-and-our-future-of-algorithmic-programming/>

Lee, Edmund, "Netflix Reports a Subscriber Bump," *New York Times*, January 21, 2020, accessed June 6, 2020, <https://www.nytimes.com/2020/01/21/business/media/netflix-q4-2019-earnings-nflx.html>

Leonard, Andrew, "How Netflix is turning viewers into puppets," *Salon*, February 1, 2013, accessed June 5, 2020, [https://www.salon.com/2013/02/01/how\\_netflix\\_is\\_turning\\_viewers\\_into\\_puppets/](https://www.salon.com/2013/02/01/how_netflix_is_turning_viewers_into_puppets/)

Lev-Ram, Michal, "How Netflix Became Hollywood's Frenemy," *Fortune*, June 7, 2016, accessed June 6, 2020, <https://fortune.com/longform/netflix-versus-hollywood/>

Levy, Karyne, "Disney Tried To Buy BuzzFeed, But BuzzFeed Wanted \$1 Billion," April 28, 2014, accessed June 5, 2020, <https://www.businessinsider.com/disney-reportedly-wanted-to-buy-buzzfeed-but-the-deal-fell-apart-2014-4>

Lieberman, David, "'House Of Cards' Was 'A Great Success' Netflix Chief Says," *Deadline*, February 25, 2013, accessed June 5, 2020, <https://deadline.com/2013/02/netflix-reed-hastings-says-house-cards-success-439416/>

Lievrouw, Leah, "Materiality and Media in Communication and Technology Studies." In: *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot (Cambridge: MIT Press, 2014)

LinkedIn Workforce Report, August 10, 2018, accessed October 5, 2019, <https://economicgraph.linkedin.com/resources/linkedin-workforce-report-august-2018>

Littleton, Cynthia, "10 Things We Learned at Variety's Big Data Summit," *Variety*, November 4, 2015, accessed June 6, 2020, <https://variety.com/2015/digital/news/10-things-we-learned-at-variety-s-big-data-summit-1201634065/>

"", "Viacom's Philippe Dauman Talks New Metrics, Creative Tools and Wall Street's 'Short-Term-ism'," *Variety*, November 4, 2015, accessed June 5, 2020, <https://variety.com/2015/biz/news/viacom-philippe-dauman-big-data-summit-1201633402/>

Lobato, Ramon, "The cultural logic of digital intermediaries: YouTube multichannel networks," *Convergence*, Volume 22, Issue 4, 2016

Lohr, Steve, "The Age of Big Data," *New York Times*, February 11, 2012, accessed June 6, 2020, <https://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>

Lowensohn, Josh, "Netflix will close its public API to some developers in November," *The Verge*, June 13, 2014, accessed March 5, 2018,

<https://www.theverge.com/2014/6/13/5808424/netflix-will-close-its-public-api-to-some-developers-in-november>

Lotz, Amanda D., *We Now Disrupt This Broadcast: How Cable Transformed Television and the Internet Revolutionized It All* (Cambridge: MIT Press, 2018)

“” “”, *Portals: A Treatise on Internet-Distributed Television* (Ann Arbor: Michigan Publishing Services, 2017)

“” “”, *The Television Will Revolutionized*, 2nd Edition (New York: New York University Press, 2014)

Luna, Taryn, “Social media play big role in movies,” *Boston Globe*, March 12, 2013, accessed June 6, 2020, <https://www.bostonglobe.com/business/2013/03/12/movies-depend-social-media-support-for-staying-power-box-office/mDRqLV2AaS1xqmLdFV1N5O/story.html>

Lunden, Ingrid, “Disney’s Next Acquisition Could Further Its Digital distribution,” *TechCrunch*, January 7, 2015, accessed June 5, 2020, <https://techcrunch.com/2015/01/07/disney-eyes-more-acquisitions-with-a-focus-on-digital-distribution/>

“” “”, “Big Data Analytics Specialist Tableau Software Raises \$254M In IPO,” *TechCrunch*, May 17, 2013, accessed June 6, 2020, <https://techcrunch.com/2013/05/17/big-data-visualization-goes-public-tableau-software-raises-254m-as-shares-pop-58-while-marketo-raises-85m/>

Lycett, Mark, “Datafication: making sense of (big) data in a complex world,” *European Journal of Information Systems*, 22, 2013

Lynch, Jason, “As Linear Ratings Continue to Slide, Buyers Say Those Viewers Will ‘Never’ Return to TV,” *AdWeek*, September 27, 2019, accessed June 5, 2020, <https://www.adweek.com/tv-video/as-linear-ratings-continue-to-slide-buyers-say-those-viewers-will-never-return-to-tv/>

Madrigal, Alexis C., “How Netflix Reverse-Engineered Hollywood,” *The Atlantic*, January 2, 2014, accessed June 5, 2020, <https://www.theatlantic.com/technology/archive/2014/01/how-netflix-reverse-engineered-hollywood/282679/>

Malinowski, Erik, “‘Space Jam’ Forever: The Website That Wouldn’t Die,” *Rolling Stone*, August 19, 2015, accessed June 5, 2020, <https://www.rollingstone.com/movies/movie-news/space-jam-forever-the-website-that-wouldnt-die-70507/>

Maltby, Richard, *Hollywood Cinema* (New York: Wiley-Blackwell, 2003)

Mamber, Stephen, “Space-Time Mappings as Database Browsing Tools.” In: *Media Computing: Computational Media Aesthetics*, edited by Chitra Dorai and Svetha Venkatesh (Berlin: Springer, 2002)

Mann, Denise, "Introduction: When Television and New Media Work Worlds Collide." In: *Wired TV: Laboring Over an Interactive Future*, edited by Denise Mann (New Brunswick: Rutgers University Press, 2014)

"" "", "Welcome to the Unregulated Wild, Wild, Digital West," *Media Industries Journal*, Volume 1, Issue 2, 2014

Manovich, Lev *Software Takes Command* (London: Bloomsbury, 2013)

"" "", *The Language of New Media* (Boston: MIT Press, 2001)

Marich, Robert, "Netflix Leads Hollywood in Mining Data to Market Films," *Variety*, March 29, 2019, accessed May 15, 2020, <https://variety.com/2019/film/news/entertainment-marketing-netflix-data-1203165944/>

Article, "U.S. Pay-TV Penetration Rate, 2009-2019," *MarketingCharts.com* (Leichtman Research Group, "Pay-TV in the U.S.," 2019), January 10, 2020, accessed June 6, 2020, <https://www.marketingcharts.com/charts/us-pay-tv-penetration-rate-2009-2019>

Marketplace Podcast, "The Business of TV in 2020," *Marketplace*, January 22, 2020, accessed June 12, 2020, <https://www.marketplace.org/shows/marketplace/the-business-of-tv-in-2020/>

Marolda, Matthew, "Changing Hollywood Paradigms with Analytics," *Innovation Enterprise On Demand*, November 12, 2014, accessed June 6, 2020, <https://ieondemand.com/presentations/changing-hollywood-paradigms-with-analytics>

Martinet, Drake, "Almost Famous: Pat Hanrahan of Tableau," *All Things Digital*, February 26, 2010, accessed June 6, 2020, <http://allthingsd.com/20100226/almost-famous-pat-hanrahan-of-tableau>

Marwick, Alice, "Silicon Valley and the Social Media Industry." In: *The SAGE Handbook of Social Media*, edited by Jean Burgess, Alice Marwick, and Thomas Poell (Thousand Oaks: Sage, 2017)

Masters, Kim, "'Battleship' Fallout," *Hollywood Reporter*, May 23, 2012, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/battleship-universal-box-office-taylor-kitsch-327972>

Mayer-Schoenberger, Viktor & Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (New York: Eamon Dolan/Houghton Mifflin Harcourt, 2013)

McAlone, Nathan, "This startup uses artificial intelligence to predict whether a Hollywood film will be a hit or a flop — just by scanning the script," *Business Insider*, July 29, 2015, accessed June 6, 2020, <https://www.businessinsider.com/this-startup-uses-artificial-intelligence-to-tell-whether-a-hollywood-film-will-be-a-hit-or-a-flop-just-from-the-script-2015-7>

McClintock, Pamela, "As Amazon Suffers String of Box Office Flops, Executives Struggle to Find Winning Strategy," *The Hollywood Reporter*, June 26, 2019, accessed June 5, 2020, <https://www.hollywoodreporter.com/news/amazon-studios-film-division-tumult-string-box-office-flops-1220968>

"", "\$200 Million and Rising: Hollywood Struggles With Soaring Marketing Costs," *Hollywood Reporter*, July 31, 2014, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/200-million-rising-hollywood-struggles-721818>

"", "Box Office: 'Godzilla' Opens to Monstrous \$93.2 Million in North America," *Hollywood Reporter*, May 17, 2014, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/box-office-godzilla-opens-monstrous-705168>

McCullough, Brian, *How the Internet Happened: From Netscape to the iPhone* (New York: Liveright, 2018)

McDonald, Paul, *The Star System: Hollywood's Production of Popular Identities* (London: British Film Institute, 2005)

McNamara, Mary, "'Scandal' has become must-tweet TV," *Los Angeles Times*, May 11, 2013, accessed June 6, 2020, <https://www.latimes.com/entertainment/tv/la-xpm-2013-may-11-la-et-st-scandal-abc-social-media-20130511-story.html>

Meehan, Eileen R., "'Holy Commodity Fetish, Batman': The political economy of the commercial intertext." In: *Hollywood: Cultural Dimensions: ideology, identity and culture industry studies*, edited by Thomas Schatz (New York: Routledge, 2004)

Mens, Tom & Demeyer, Serge, *Software Evolution* (New York & London: Springer, 2008)

Mishra, Devendra, "Hollywood's Digital Blind Spots," *Graziadio Business Review*, Volume 18, Issue 2, 2015

Moore, Malcolm, "Film studios select movie-fan data for starring role to drive ticket sales," *Financial Times*, September 6, 2016, accessed June 6, 2020, <https://www.ft.com/content/8489fd08-15bc-11e6-b197-a4af20d5575e>

Morris, David Z., "Netflix says Geography, Age, and Gender are 'Garbage' for Predicting Taste," *Yahoo Finance*, March 27, 2016, accessed June 6, 2020, <https://finance.yahoo.com/news/netflix-says-geography-age-gender-192801834.html>

MPAA, "Theme Report 2019," *Motion Picture Association of America*, March 2020, accessed June 5, 2020, <https://www.motionpictures.org/wp-content/uploads/2020/03/MPA-THEME-2019.pdf>



“” “”, “Theme Report 2018,” *Motion Picture Association of America*, March 2019, accessed June 5, 2020, <https://www.motionpictures.org/research-docs/2018-theatrical-home-entertainment-market-environment-theme-report/>

Murphy, Matt & Steve Sloane, “The Rise of APIs,” *TechCrunch*, May 21, 2016, accessed July 5, 2019, <https://techcrunch.com/2016/05/21/the-rise-of-apis/>

Mustain, Andrea & Mike Osborne, “Gold or Pyrite,” *Raw Data Podcast*, November 29, 2016, accessed June 6, 2020, <https://podcasts.apple.com/us/podcast/gold-or-pyrite/id1042137974?i=1000378399788>

Napoli, Philip M. & Roepnack, Axel, eds., *Big Data and Media Management* (New York: Routledge, 2018)

Napoli, Philip M., “On Automation in Media Industries: Integrating Algorithmic Media Production into Media Industries Scholarship,” *Media Industries Journal*, Volume 1, Issue 1, 2014

“” “”, *Audience Evolution: New Technologies and the Transformation of Media Audiences* (New York: Columbia University Press, 2010)

Netflix, “Prospectus,” *SEC*, May 22, 2002, accessed June 5, 2020, <https://www.sec.gov/Archives/edgar/data/1065280/000101287002002475/d424b4.htm>

Netflix Investor Relations, “Long-Term View,” *Netflix*, September 16, 2013, accessed June 5, 2020, <https://www.netflixinvestor.com/ir-overview/long-term-view/default.aspx>

“” “”, “Shareholder Letter, Q4, 2010,” January 21, 2011, accessed June 5, 2020, [https://s22.q4cdn.com/959853165/files/doc\\_financials/quarterly\\_reports/2010/q4/Q410-Letter-to-shareholders.pdf](https://s22.q4cdn.com/959853165/files/doc_financials/quarterly_reports/2010/q4/Q410-Letter-to-shareholders.pdf)

Netflix Research, “Analytics: Driving Insights from Data,” *Netflix*, accessed October 5, 2019, <https://research.netflix.com/research-area/analytics>

Nielsen, “Nielsen estimates 120.6 million TV homes in the U.S. for the 2019-2020 TV season,” August 27, 2019, accessed June 5, 2020, <https://www.nielsen.com/us/en/insights/article/2019/nielsen-estimates-120-6-million-tv-homes-in-the-u-s-for-the-2019-202-tv-season/>

“” “”, “Nielsen and Twitter Establish Social TV Rating,” December 18, 2012, accessed June 6, 2020, <https://www.nielsen.com/us/en/press-releases/2012/nielsen-and-twitter-establish-social-tv-rating>

Ng, David, “Netflix plays peekaboo with its ratings. Hollywood isn’t amused,” *Los Angeles Times*, January 24, 2019, accessed March 5, 2019, <https://www.latimes.com/business/hollywood/la-fi-ct-netflix-ratings-20190124-story.html>

Nguyen, Nicole, "Netflix Wants To Change The Way You Chill," *BuzzFeed News*, December 13, 2018, accessed June 6, 2020, <https://www.buzzfeednews.com/article/nicolenguyen/netflix-recommendation-algorithm-explained-binge-watching>

OECD, "Mobile broadband subscriptions per 100 inhabitants Q4 2009 – Q4 2019 (Indicator)," accessed June 6, 2020, <https://data.oecd.org/broadband/mobile-broadband-subscriptions.htm>

Ohmer, Susan, *George Gallup in Hollywood* (New York: Columbia University Press, 2006)

Ortner, Sherry B., "Access: Reflections on Studying Up in Hollywood," *Ethnography*, 11, 2, 2010

Osborne, Charlie, "Salesforce acquires Tableau Software in \$15.7 billion deal," *ZDNet*, June 10, 2019, accessed June 6, 2020, <http://zdnet.com/article/salesforce-acquires-tableau-software-in-15-7-billion-deal/>

Otterson, Joe, "TV's Top Marketing Execs Break Down the Power of Data at Variety's Massive Summit," *Variety*, March 21, 2018, accessed June 5, 2020, <https://variety.com/2018/tv/news/tv-marketing-variety-massive-summit-1202733153/>

Packel, Dan & Lee Rainie, "More Online, Doing More," Pew Research, February 18, 2001, accessed June 6, 2020, <https://www.pewresearch.org/internet/2001/02/18/more-online-doing-more>

Pang, Justin, "The Future of News and Publishing," *TechCrunch*, February 17, 2016, accessed June 6, 2020, <https://techcrunch.com/2016/02/17/the-future-of-news-and-publishing/>

Parrot Analytics, Website, "What are your data sources in more detail?," *Parrot Analytics*, accessed June 6, 2020, <https://support.parrotanalytics.com/hc/en-us/articles/222713988-What-are-your-data-sources-in-more-detail>

Patel, Sahil, "How Turner trained 500 employees to sell brand social videos globally," *Digiday*, June 15, 2017, accessed June 6, 2020, <http://digiday.com/media/how-turner-trained-500-employees-to-sell-branded-social-videos>

"", "Inside Disney's troubled \$675 mil. Maker Studios acquisition," *Digiday*, February 22, 2017, accessed June 5, 2020, <https://digiday.com/media/disney-maker-studios/>

Perren, Alisa & Jennifer Holt, eds., *Media Industries Studies: History, Theory, and Method* (Hoboken: John Wiley & Sons, 2011)

Perren, Alisa, "Producing Filmed Entertainment." In: *Managing Media Work*, edited by Mark Deuze (New York: Sage, 2011)

Petruska, Karen, "Amazon Prime Video: Where Information is Entertainment." In: *From Networks to Netflix: A Guide to Changing Channels*, edited by Derek Johnson (New York: Routledge, 2018)

Pew Research, "Internet/Broadband Fact Sheet," June 12, 2019, accessed June 6, 2020, <https://www.pewresearch.org/internet/fact-sheet/internet-broadband/>

"", "Social Media Usage: 2005-2015," October 8, 2015, accessed June 6, 2020, <https://www.pewresearch.org/internet/2015/10/08/social-networking-usage-2005-2015/>

"", "The Internet News Audience Goes Ordinary," January 14, 1999, accessed June 5, 2020, <https://www.people-press.org/1999/01/14/the-internet-news-audience-goes-ordinary/>

Picker, David H., *Musts, Maybes, and Nevers: A Book About The Movies* (Scotts Valley: CreateSpace Independent Publishing, 2013)

Pressman, Aaron & Lashinsky, Adam, "Data Sheet – Why AT&T's Time Warner Deal Brings the Future of Television Closer," *Fortune*, June 13, 2018, accessed June 5, 2020, <https://fortune.com/2018/06/13/data-sheet-att-time-warner-antitrust-analysis/>

Pressberg, Matt & Donnelly, Matt, "Hollywood's Virtual Reality Push: How All 6 Studios Stack Up," *The Wrap*, July 24, 2017, accessed June 6, 2020, <https://www.thewrap.com/hollywood-virtual-reality-push-how-all-6-major-studios-stack-up/>

Proença Santos, Amanda, "Disney to Start Own Streaming Services, Remove Content From Netflix," *NBC News*, August 8, 2017, accessed June 5, 2020, <https://www.nbcnews.com/pop-culture/tv/disney-start-own-streaming-services-remove-content-netflix-n791001>

Proulx, Mike & Stacey Shepatin, *Social TV: How Marketers Can Reach and Engage Audiences by Connecting Television to the Web, Social Media, and Mobile* (Hoboken: Wiley, 2012)

Pullen, John Patrick, "In Hollywood, social media takes a leading role," *Fortune*, August 1, 2014, accessed June 6, 2020, <https://fortune.com/2014/08/01/in-hollywood-social-media-takes-a-leading-role/>

Rafat, Ali, "Disney Buying Storytelling Social Net FanLib; DigiSynd in Process," *CBS News*, June 3, 2008, accessed June 6, 2020, <https://www.cbsnews.com/news/disney-buying-storytelling-social-net-fanlib-digisynd-in-process/>

Rainey, James, "The Perils of Promotion: Pricey TV Campaigns, Fear of Change Shackles Movie Spending," *Variety*, March 8, 2016, accessed June 6, 2020, <https://variety.com/2016/film/features/movie-marketing-advertising-tv-campaigns-1201724468/>

Ramsay, Stephen, *Reading Machines: Towards an Algorithmic Criticism* (Champaign: University of Illinois Press, 2011)

Randle, Keith, "The Organization of Film and Television Production." In: *Managing Media Work*, edited by Mark Deuze (New York: Sage, 2011)

Raviv, Shaun, "Moneyballing the Movies: How the Box Office Became a Sport," *The Ringer*, August 2, 2018, accessed June 6, 2020, <https://www.theringer.com/movies/2018/8/2/17641822/box-office-reporting-mojito-the-numbers-marvel-star-wars>

Reuters, "GetGlue to boost live TV with Facebook check-ins," *Reuters*, March 23, 2011, accessed June 6, 2020, <https://www.reuters.com/article/us-getglue/getglue-to-boost-live-tv-with-facebook-check-ins-idUSTRE72M92720110324>

Robischon, Noah, "How BuzzFeed's Jonah Peretti Is Building A 100-Year Media Company," *Fast Company*, February 16, 2016, accessed June 5, 2020, <https://www.fastcompany.com/3056057/how-buzzfeeds-jonah-peretti-is-building-a-100-year-media-company>

Rodriguez, Ashley, "Netflix didn't make many of the "originals" that made it famous. That's changing.," *Quartz*, February 26, 2019, accessed June 5, 2020, <https://qz.com/1545594/netflix-doesnt-make-most-of-its-originals-now-thats-changing>

Rose, Frank, *The Art of Immersion: How the Digital Generation Is Remaking Hollywood, Madison Avenue, and the Way We Tell Stories* (New York: W. W. Norton & Company, 2012)

Rosenbush, Steven & Michael Totty, "How Big Data Is Changing the Whole Equation for Business," *Wall Street Journal*, March 10, 2013, accessed June 6, 2020, <https://www.wsj.com/articles/SB10001424127887324178904578340071261396666>

Ross, Alexander G., "Creative decision making within the contemporary Hollywood studios," *Journal of Screenwriting*, Volume 2, Number 1, 2011

Ross, Sharon Marie, *Beyond the Box: Television and the Internet* (New York: Wiley-Blackwell, 2008)

Ruparelia, Nayan B., *Cloud Computing* (Cambridge: MIT Press, 2016)

Rutherford, Kevin, "Dwayne Johnson Leads Top Actors Social Media Ranking for 10th Week," *Hollywood Reporter*, September 9, 2017, accessed June 6, 2020, <https://www.hollywoodreporter.com/lists/dwayne-johnson-leads-top-actors-social-media-ranking-10th-week-1036720/item/gal-gadot-mvp-9-13-1036339>

Salesforce, "Salesforce Completes Acquisition of Tableau," *Salesforce*, August 1, 2019, accessed June 6, 2020, <https://www.tableau.com/about/press-releases/2019/salesforce-completes-acquisition-tableau>

Schatz, Thomas, "HBO and Netflix: Getting Back to the Future," *Flow*, January 20, 2014, accessed August 5, 2019, <https://www.flowjournal.org/2014/01/hbo-and-netflix-%E2%80%93-getting-back-to-the-future/>

Schlesinger, Scott, "Using Analytics to Predict Hollywood Blockbusters," *Harvard Business Review*, October 11, 2012, accessed June 5, 2020, <https://hbr.org/2012/10/using-analytics-to-predict-hollywood-blockbusters>

Schomey, Audrey, "Viacom is expanding its streaming distribution to offset ad declines," *Business Insider*, May 13, 2019, accessed June 5, 2020, <https://www.businessinsider.com/viacom-expands-streaming-distribution-to-offset-ad-declines-2019-5>

Seger, Wared, "Parrot Analytics: Understanding Audience Demand for TV Shows - Variety Big Data 2016," *Parrot Analytics*, December 13, 2016

Sellers, Michael D., *John Carter and the Gods of Hollywood* (Lichfield, UK: Universal Media, 2012)

Sharma, Amol, "Amazon Mines Its Data Trove to Bet on TV's Next Hit," *Wall Street Journal*, November 1, 2013, accessed June 6, 2020, <https://www.wsj.com/articles/amazon-mines-its-data-trove-to-bet-on-tv8217s-next-hit-1383361270>

Shaw, Lucas, "Showtime Goes Social, Launches New 'Second Screen' App," *The Wrap*, September 26, 2011, accessed June 6, 2020, <https://www.thewrap.com/showtime-goes-social-launches-new-second-screen-app-31308/>

Shields, Ronan, "U.S. Digital Ad Spend Will Surpass Offline in 2019," *AdWeek*, February 20, 2019, accessed June 6, 2020, <https://www.adweek.com/programmatic/u-s-digital-ad-spend-will-surpass-offline-in-2019/>

Siegel, Tatiana, "Warner Bros. Signs Deal for AI-Driven Film Management System," *Hollywood Reporter*, January 8, 2020, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/warner-bros-signs-deal-ai-driven-film-management-system-1268036>

Smith, Michael D. & Rahul Telang, *Streaming, Sharing, Stealing: Big Data And The Future Of Entertainment* (Cambridge: MIT Press, 2016)

Spangler, Todd, "Edward Norton-Backed TV-Analytics Startup EDO Raises \$12 Million," *Variety*, November 1, 2018, accessed June 6, 2020, <https://variety.com/2018/digital/news/edward-norton-edo-analytics-funding-12-million-1203014461/>

“” “”, “TV Time Launches Social-Analytics Tool to Break Down Fan Reaction to Shows,” *Variety*, July 31, 2018, accessed June 6, 2020, <https://variety.com/2018/digital/news/tv-time-fan-reaction-social-analytics-1202888016/>

“” “”, “Warner Bros. Acquires Full Control of Machinima,” *Variety*, November 17, 2016, accessed June 5, 2020, <https://variety.com/2016/digital/news/warner-bros-acquires-machinima-1201920793/>

Statista, “Number of movie tickets sold in the U.S. and Canada from 1980 to 2019,” April 24, 2020, accessed June 6, 2020, <https://www.statista.com/statistics/187073/tickets-sold-at-the-north-american-box-office-since-1980/>

“” “”, “Number of smartphone users in the United States from 2018 to 2024,” April 21, 2020, accessed June 6, 2020, <https://www.statista.com/statistics/201182/forecast-of-smartphone-users-in-the-us/>

Steel, Emily & Barnes, Brooks, “Disney, Time Warner and Other Media Shares Fall as Investors Worry,” August 5, 2015, accessed October 5, 2019, <https://www.nytimes.com/2015/08/06/business/media/disney-time-warner-media-investors-worry-about-tvs-future.html>

Steel, Emily, “For Its New Shows, Amazon Adds Art to Its Data,” *New York Times*, August 15, 2014, accessed June 5, 2020, <https://www.nytimes.com/2014/08/16/business/media/for-its-new-shows-amazon-adds-art-to-its-data.html>

Shaw, Lucas & Palmeri, Christopher, “Disney takes tighter control of Hulu,” *Los Angeles Times*, February 3, 2020, accessed June 6, 2020, <https://www.latimes.com/entertainment-arts/business/story/2020-02-03/disney-takes-tighter-control-of-hulu>

Shu, Catherine, “Atom Tickets raises \$60M Series C for its movie booking app,” *TechCrunch*, March 8, 2018, accessed June 6, 2020, <https://techcrunch.com/2018/03/08/atom-tickets-raises-60m-series-c-for-its-movie-booking-app/>

Sweeney, Erica, “Digital represents 14% of movie ad budgets but drives 46% of revenue, study finds,” *Marketing Dive*, December 13, 2018, accessed June 6, 2020, <https://www.marketingdive.com/news/digital-represents-14-of-movie-ad-budgets-but-drives-46-of-revenue-study/544253/>

Szalai, Georg, “Warner Bros. to Acquire Rotten Tomatoes Owner Flixster,” *Hollywood Reporter*, May 4, 2011, accessed June 6, 2020, <https://www.hollywoodreporter.com/news/warner-bros-acquire-rotten-tomatoes-185237>

Tableau, “Form S-1 Registration Statement, *SEC*, April 2, 2013, accessed June 6, 2020, <https://www.sec.gov/Archives/edgar/data/1303652/000119312513138700/d469057ds1.htm>

Telotte, J. P., “The *Blair Witch Project* Project: Film and the Internet,” *Film Quarterly*, 54 (3), 2001

Think with Google, “Lights! Camera! Data! How insights help 20th Century Fox Film reach the right audiences,” *Google*, August 2018, accessed June 5, 2020, <https://www.thinkwithgoogle.com/marketing-resources/data-measurement/data-insights-film-marketing>

Thompson, Derek, “The Reason Why Hollywood Makes So Many Boring Superhero Movies,” *The Atlantic*, May 13, 2014, accessed June 5, 2020, <https://www.theatlantic.com/entertainment/archive/2014/05/hollywoods-real-superhero-problem/370785>

Thompson, Kristin, *The Frodo Franchise: The Lord of the Rings and Modern Hollywood* (Berkeley: University of California Press, 2007)

Time Warner, “Time Warner leads \$6M round of funding in GetGlue, the leader in social entertainment.” *Time Warner Group (WarnerMedia)*, December 7, 2010, accessed June 6, 2020, <https://www.warnermediagroup.com/newsroom/press-releases/2010/12/07/time-warner-leads-6m-round-of-funding-in-getglue-the-leader-in>

“” “”, “AOL & Time Warner Will Merge To Create World's First Internet-Age Media & Communications Company,” January 10, 2000, accessed June 5, 2020, <https://www.warnermediagroup.com/newsroom/press-releases/2000/01/10/aol-time-warner-will-merge-to-create-world-s-first-internet-age>

Tryon, Chuck, *On-Demand Culture: Digital Delivery and the Future of Movies* (New Brunswick: Rutgers University Press, 2013)

UCLA Anderson, “Future of Media: Digital, Data & Brands Panel,” *Big Data Conference*, June 28, 2019, accessed June 6, 2020, <https://www.youtube.com/watch?v=BKPSrglFdwg>

Ulin, Jeffrey C., *The Business of Media Distribution: Monetizing Film, TV and Video Content in an Online World* (New York: Focal Press, 2010)

Ungerleider, Neil, ““Pitch Perfect” And How Analytics Are Transforming Movie Marketing,” January 10, 2014, accessed June 6, 2020, <https://www.fastcompany.com/3024655/pitch-perfect-and-how-analytics-are-transforming-movie-marketing>

Van Dijck, Jose, Poell, Thomas & de Waal, Martijn, *The Platform Society: Public Values in a Connective World* (Oxford: Oxford University Press, 2018)

Van Dijck, Jose, *The Culture of Connectivity: A Critical History of Social Media* (Oxford: Oxford University Press, 2013)

Van Dijck, Jose & Poell, Thomas, "Understanding Social Media Logic," *Media and Communication*, Volume 1, Number 1, 2013

Variety, "Turner's Jesse Redniss on How AT&T Changes the Data Game," *Strictly Business Podcast*, October 9, 2018, accessed June 5, 2020, <https://variety.com/2018/digital/news/listen-turners-jesse-redniss-on-how-att-changes-the-data-game-1202973028/>

"", "Big Data Summit: Metrics, Analytics, 'Wild West' Opportunities Parsed by Industry Leaders," *Variety*, November 10, 2016, accessed June 6, 2020, <https://variety.com/2016/tv/news/big-data-summit-facebook-vimeo-twitter-1201915456/>

"", "Variety Launches Vscore to Measure Actors' Value," *Variety*, August 6, 2014, accessed June 6, 2020, <https://variety.com/2014/biz/news/johnny-depp-value-jennifer-lawrence-variety-vsore-1201263164/>

Verhoeven, Beatrice, "How Data Drives Studio Decisions as Much as Gut Instinct in Netflix Era," *The Wrap*, October 2, 2018, accessed June 6, 2020, <https://www.thewrap.com/how-data-is-driving-studio-decisions-as-much-as-gut-instinct-in-netflix-era/>

Vidyardhi, Neil, "Social TV: Water Cooler Effect Gone Online," *AdWeek*, June 19, 2012, accessed June 6, 2020, <https://www.adweek.com/digital/social-tv-water-cooler-effect-gone-online-at-140conf-live-video/>

Vincent, James, "Hollywood Is Quietly Using AI To Help Decide Which Movies To Make," *The Verge*, May 28, 2019, accessed June 6, 2020, <https://www.theverge.com/2019/5/28/18637135/hollywood-ai-film-decision-script-analysis-data-machine-learning>

"", "China has turned Warcraft into the highest-grossing video game film ever," *The Verge*, June 21, 2016, accessed June 6, 2020, <https://www.theverge.com/2016/6/21/11988990/warcraft-highest-grossing-video-game-movie>

Vonderau, Patrick, "The video bubble: Multichannel networks and the transformation of YouTube," *Convergence*, Volume 22, Issue 4, 2016

"", "Industry Proximity," *Media Industries Journal*, Volume 1, Issue 1, 2014

Vonderau, Patrick, Snickars, Pelle & Jean Burgess, eds., *The YouTube Reader* (Stockholm: National Library of Sweden, 2010)

Wagner, Kurt, "Twitter Says It's Time for Must-Tweet TV," *Recode*, September 18, 2014, accessed June 6, 2020, <https://www.vox.com/2014/9/18/11631024/twitter-to-tv-networks-live-tweeting-may-boost-your-ratings>



Wall Street Journal, "Snapchat Spurned \$3 Billion Acquisition Offer from Facebook," November 13, 2013, accessed June 5, 2020, <https://blogs.wsj.com/digits/2013/11/13/snapchat-spurned-3-billion-acquisition-offer-from-facebook/>

Wallenstein, Andrew, "'Trolls World Tour' Marks Real Progress For Premium VOD. But It's Just The Beginning," *Variety*, April 30, 2020, accessed June 6, 2020, <https://variety.com/vip/trolls-world-tour-marks-real-progress-for-premium-vod-but-its-just-the-beginning-1234591645/>

"", "Netflix in Japan: Tepid Demand for Originals, Data Indicates," *Variety*, September 1, 2015, accessed June 6, 2020, <https://variety.com/2015/data/news/tepid-demand-for-originals-on-netflix-in-japan-1201581943/>

"", "Big Data: Media Embracing the Most Detailed Information About You Yet," *Variety*, September 25, 2013, accessed June 5, 2019, <https://variety.com/2013/biz/news/big-data-media-embracing-the-most-detailed-information-about-you-yet-1200665847/>

Ward, Karen, "Hollywood's conception of its audiences in the 1920s." In: *The Classic Hollywood Reader*, Steve Neale, ed. (New York: Routledge, 2012)

Wardrip-Fruin, Noah, *Expressive Processing: Digital Fictions, Computer Games, and Software Studies* (Cambridge: MIT Press, 2009)

Warren, Christina, "GetGlue Gets Slick New Dashboard for TV Marketers," *Mashable*, September 8, 2011, accessed June 6, 2020, <https://mashable.com/2011/09/08/getglue-business/>

Wasko, Janet, "Hollywood in the 21st Century," *Economia della Cultura*, Issue 4, 2017

"", *How Hollywood Works* (New York: Sage, 2003)

"", *Hollywood in the Information Age: Beyond the Silver Screen* (Hoboken: Wiley, 1994)

Watercutter, Angela, "To Predict the Next Box Office Hit, Look at Google Movie Trailer Searches," *Wired*, June 6, 2013, accessed June 6, 2020, <https://www.wired.com/2013/06/google-box-office-prediction/>

Waxman, Sharon, "Ryan Kavanaugh's Relativity Turns 10: The CEO on Succeeding Despite Predictions and What's Next," *The Wrap*, May 19, 2014, accessed June 6, 2020, <https://www.thewrap.com/relativity-ceo-ryan-kavanaugh-on-10th-anniversary/>

Waxman, Sharon & Paul Farhi, "Going Down With The Ship?," *Washington Post*, May 25, 1997, accessed June 6, 2020, <https://www.washingtonpost.com/archive/business/1997/05/25/going-down-with-the-ship/5c0ed22f-9c2a-4420-8c4a-b2b983eea681/>

Weaver, Erik, "Three Ways AI will Solve Hollywood's Data Problem," *Western Digital Blog*, May 11, 2017, accessed June 5, 2020, <https://blog.westerndigital.com/ai-will-solve-hollywoods-data-problem>

Webedia & Vertigo, "From Discovery to Purchase - The Moviegoing Experience Begins Online - A Box Office Profile," *Webedia & Vertigo*, 2019, quoted in Forbes, March 11, 2019, accessed June 6, 2020, <https://www.forbes.com/sites/scottmendelson/2019/03/11/more-moviegoers-buying-more-of-their-tickets-online-study-shows/>

Weiner, Jonah, "The Great Race to Rule Streaming TV," *New York Times*, July 10, 2019, <https://www.nytimes.com/2019/07/10/magazine/streaming-race-netflix-hbo-hulu-amazon.html>

Weinraub, Bernard, "As Problems Delay 'Titanic,' Hollywood Sighs in Relief," *New York Times*, May 29, 1997, accessed June 6, 2020, <https://www.nytimes.com/1997/05/29/movies/as-problems-delay-titanic-hollywood-sighs-in-relief.html>

Willens, Michele, "Putting Films to the Test, Every Time," *New York Times*, June 25, 2000, accessed June 6, 2020 <https://www.nytimes.com/2000/06/25/movies/film-putting-films-to-the-test-every-time.html>

World Economic Forum Report, "Data Science in the New Economy," *World Economic Forum*, July 2019, accessed October 5, 2019, [http://www3.weforum.org/docs/WEF\\_Data\\_Science\\_In\\_the\\_New\\_Economy.pdf](http://www3.weforum.org/docs/WEF_Data_Science_In_the_New_Economy.pdf)

Wright, Richard, "Data Visualizations." In: *Software Studies: A Lexicon*, edited by Matthew Fuller (Cambridge: MIT Press, 2008)

Wu, Tim, *The Attention Merchants: The Epic Scramble to Get Inside Our Heads* (New York: Vintage, 2017)

"" "", "Netflix's Secret Special Algorithm is a Human," *The New Yorker*, January 27, 2015, accessed June 5, 2020, <https://www.newyorker.com/business/currency/hollywoods-big-data-big-deal>

Wyatt, Justin, "Market Research in the Media Industries: On the Strategic Relationship between Client and Supplier." In: Derek Johnson, Derek Kompare & Avi Santo, eds., *Making Media Work: Cultures of Management in the Entertainment Industries* (New York: New York University Press, 2014)

"" "", *High Concept: Movies and Marketing in Hollywood* (Austin: University of Texas Press, 1994)

Zhang, Jack, "Why the Algorithm That Promised to Save Hollywood Destroyed Relativity Media," *No Film School*, April 13, 2016, <https://nofilmschool.com/2016/04/why-algorithm-promised-to-save-hollywood-destroyed-relativity-media>

## Bibliography – Software References

Amazon x-Ray, <https://www.amazon.com/adlp/xray>

Android, Google Play Developer API, <https://developer.android.com/google/play/developer-api>

Apple Developer, App Store Analytics, <https://developer.apple.com/app-store-connect/analytics/>

Facebook Analytics, <https://analytics.facebook.com/>

Facebook for Developers, APIs and SDKs, <https://developers.facebook.com/docs/apis-and-sdks/>

“” “”, Insights API, <https://developers.facebook.com/docs/marketing-api/insights/>

Google Analytics, YouTube, <https://www.youtube.com/user/googleanalytics>

Google Developers, Google Analytics API, <https://developers.google.com/analytics>

Hulu, Advertising Insights, <https://advertising.hulu.com/insights/>

IMDb Developer, API products, <https://developer.imdb.com/>

Netflix Research, <https://research.netflix.com/>

Netflix Tech Blog, <https://netflixtechblog.com/>

Parrot Analytics, Applications, <https://www.parrotanalytics.com/applications>

Rotten Tomatoes Developer Network, [https://developer.fandango.com/rotten\\_tomatoes](https://developer.fandango.com/rotten_tomatoes)

Tableau, Developer Tools, <https://www.tableau.com/developer/tools>

“” “”, Our approach to analytics, <https://www.tableau.com/solutions/analytics>

“” “”, YouTube, <https://www.youtube.com/user/tableausoftware>

Twitter Developer, Use Cases / Analyze, <https://developer.twitter.com/en/use-cases/analyze>

YouTube Analytics, <https://www.youtube.com/analytics>

YouTube / Google Developers, Data API, <https://developers.google.com/youtube/v3>