9-2023

Evaluating Copyright Protection in the Data-Driven Era: Centering on Motion Picture's Past and Future

Chieh-Li Pai

JEROME HALL LAW LIBRARY
INDIANA UNIVERSITY
Maurer School of Law
Bloomington

EVALUATING COPYRIGHT PROTECTION IN THE DATA-DRIVEN ERA: CENTERING ON MOTION PICTURE'S PAST AND FUTURE

Chieh-Li Pai

Submitted to the faculty of Indiana University Maurer School of Law

in partial fulfillment of the requirements for the degree

Doctor of Juridical Science

September 2023

Accepted by the faculty, Indiana University Maurer School of Law, in partial fulfillment of the requirements for the degree of Doctor of Juridical Science.

Doctoral Committee

_____

Marshall A. Leaffer

Professor of Law

Distinguished Scholar in Intellectual Property Law and University Fellow

Indiana University Maurer School of Law

_____

Mark D. Janis

Robert A. Lucas Chair in Law

Director, Center for Intellectual Property Research

Indiana University Maurer School of Law

_____

Michael Mattioli

Professor of Law and Louis F. Niezer Faculty Fellow

Indiana University Maurer School of Law

Date of Defense – September 8, 2023

ii

ACKNOWLEDGMENTS

This dissertation concludes my four years of studies in Bloomington. I have mixed feelings: part sadness that I will leave this beautiful town and many lovely people, part happiness that I completed a significant achievement in my life, and part emptiness that I have not learned enough.

I want to express my deepest appreciation to my advisor, Professor Marshall Leaffer. He was always interested in what I would write for the next chapter. This dissertation could not have been completed without his continuous encouragement and guidance. I also want to thank Professor Mark Janis and Professor Michael Mattioli for their assistance and feedback. They inspired me to explore more about the motion picture's past and future.

Many thanks to Professor Gabrielle Goodwin, Mr. Tom Bitters, and Ms. Nicole Scelta. I learned valuable English writing skills and citation format from them, but I am responsible for any mistakes or errors committed in this dissertation. I am thankful to Dr. Yin-Song Hsu, Dr. Shuimei Liu, and Dr. Chieh Huang for sharing their expertise on dissertation writing and defense experiences. I also appreciate the fellow Toastmasters, Mr. John Pedersen, and Ms. Hwi Won Kim's support in improving my English speech. All of those people contributed to the success of my dissertation defense.

I am grateful that my mother stayed strong these years so I can focus on studying abroad with less worry. Great gratitude to my wife, my son, and my daughter. I will never forget the love, companionship, and care they showed during this time.

Finally, I want to thank all the people, friends or strangers, who gave me a hand or showed goodness and kindness to me for the past four years. I have had a pleasant study abroad experience because of them.

Chieh-Li Pai

EVALUATING COPYRIGHT PROTECTION IN THE DATA-DRIVEN ERA:
CENTERING ON MOTION PICTURE'S PAST AND FUTURE

Since the 1910s, Hollywood has measured audience preferences with rough industry-created methods. In the 1940s, scientific audience research led by George Gallup started to conduct film audience surveys with traditional statistical and psychological methods. However, the quantity, quality, and speed were limited. Things dramatically changed in the internet age. The prevalence of digital data increases the instantaneousness, convenience, width, and depth of collecting audience and content data. Advanced data and AI technologies have also allowed machines to provide filmmakers with ideas or even make human-like expressions. This brings new copyright challenges in the data-driven era.

Massive amounts of text and data are the premise of text and data mining (TDM), as well as the admission ticket to access machine learning technologies. Given the high and uncertain copyright violation risks in the data-driven creation process, whoever controls the copyrighted film materials can monopolize the data and AI technologies to create motion pictures in the data-driven era. Considering that copyright shall not be the gatekeeper to new technological uses that do not impair the original uses of copyrighted works in the existing markets, this study proposes to create a TDM and model training limitations or exceptions to copyrights and recommends the Singapore legislative model.

Motion pictures, as public entertainment media, have inherently limited creative choices. Identifying data-driven works' human original expression components is also challenging. This study proposes establishing a voluntarily negotiated license institution backed up by a compulsory license to enable other filmmakers to reuse film materials in new motion pictures. The film material's degree of human original authorship certified by film artists' guilds shall be a crucial factor in deciding the compulsory license's royalty rate and terms to encourage retaining human artists. This study argues that international and domestic policymakers should enjoy broad discretion to qualify data-driven work's copyright protection because data-driven work is a new category of work. It would be too late to wait until ubiquitous data-driven works block human creative freedom and floods of data-driven work copyright litigations overwhelm the judicial systems.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

I. BACKGROUND

This dissertation addresses the legal challenges of defining appropriate copyright protection for motion pictures in the data-driven era. Data means big data which has characteristics of large volume and variety regardless of its format.[1] In addition to volume and variety, velocity and veracity help determine big data's value.[2] But what is most valuable is the hidden patterns and insights extracted from the big data.[3] Data technology is reshaping the traditional production of commercial motion pictures such as movies and television series. It can help to parse the audience's favorite content and keep up with trending topics through the collection and analysis of big data.[4] The valuable information extracted from big data informs filmmakers to make more concrete, diverse, and audience-oriented content. With a deep understanding of audience opinions over content, filmmakers can make informed creative decisions and target the right audience precisely.[5] Accordingly, data-driven creation has become an irresistible trend in the motion picture industry.[6]

---

[1] *See* Nina Baur et al., *The Quality of Big Data. Development, Problems, and Possibilities of Use of Process-Generated Data in the Digital Age*, 45(3) HIST. SOC. RES. / HIST. SOZIALFORSCHUNG 209, 211 (2020).

[2] *See* Daniel L. Rubinfeld & Michal S. Gal, *Access Barriers to Big Data*, 59 ARIZ. L. REV. 339, 345-47 (2017). The authors explain that volume refers to the size of the data; velocity refers to the refresh frequency of the data; variety refers to the diversity of sources, types, and timeframes of the data; and veracity refers to the reliability of the data.

[3] *Id.* at 347.

[4] *See* Institute for Information Industry (資策會), *Buzhi Xie Juben Zicehui Dashuju Fenxijishu Yanjiu Shehui Baitai* (不只寫劇本 資策會大數據分析技術研究社會百態), INSTITUTE FOR INFORMATION INDUSTRY (Jan. 29, 2021), https://www.iii.org.tw/Press/NewsDtl.aspx?nsp_sqno=2341&fm_sqno=14 [https://perma.cc/5QJ3-X5YX].

[5] *See Big Data and Hollywood: A Love Story*, THE ATLANTIC, https://www.theatlantic.com/sponsored/ibm-transformation-of-business/big-data-and-hollywood-a-love-story/277/ [https://perma.cc/7QV5-PNMZ].

[6] *See* Kal Raustiala & Christopher Jon Sprigman, *The Second Digital Disruption: Streaming and the Dawn of Data-Driven Creativity*, 94 N.Y.U. L. REV. 1555, 1584-86. (2019).

A. TWO DATA-DRIVEN CASES IN TAIWAN

In Taiwan, application of big data in the creative process of motion pictures is still in the beginning stages. The first attempt was the television series *The World Between Us* (2019).[7] It gained the third-highest viewership rating in the Taiwan Public Television Service's (PTS) programming history.[8] Another example is the PTS's television series, *Who Killed the Good Man* (2021) (also known as *The Great Debt Era*). This show was highly admired for touching young-generation viewers' hearts.[9] Why were these melodramas so popular and touching? Big data analytics, offered by the Digital Service Innovation Institute (DSI) of the Institute for Information Industry (Taiwan), was claimed to be a crucial factor.[10]

1. *The World Between Us*

The World Between Us explored society's stigmas and biases towards prisoners on death row and individuals with mental illness. The show portrayed interactions and confrontations between criminal defendants, crime victims, criminal lawyers, news journalists, mental health professionals, and their families.[11] The drama's classic line was,

---

[7] Lin Ruiwan (林瑞婉), *Women yu Dashuju de Juli* (我們與大數據的距離), PTS ACTION Q. (May 31, 2018), https://medium.com/@PTS_quarterly/%E6%88%91%E5%80%91%E8%88%87%E5%A4%A7 %E6%95%B8%E6%93%9A%E7%9A%84%E8%B7%9D%E9%9B%A2-69ecfce2f2dc [https://perma. cc/U7LB-UCMQ].

[8] Digital Transformation Research Institute of Institute for Information Industry (資策會數位轉型研究院), *AI, Dashuju, Juben: Kuaizhi Renkou de Women yu E de Juli Juben Shi Zenme Chansheng d?* (AI、大數據 、劇本：膾炙人口的「我們與惡的距離」劇本是怎麼產生的？) DTRI (Apr. 14, 2020), https://ideas-dtri.iii.org.tw/%E6%9C%8D%E5%89%B5%E6%89%80%E7%A0%94%E7%99%BC%E6%87%89%E7% 94%A8/ai-writed-tv-series-script/ [https://perma.cc/W99V-SDYC].

[9] Yang Wenjun (楊文君), *Dazhaishidai Bei Ping Xieshi dao Jiuxin Dashuju Jiachi de Juben Weihe Neng Ruci Tieqie* (大債時代被評寫實到揪心　大數據加持的劇本為何能如此貼切？) RADIO TAIWAN INT'L (Feb. 15, 2021, 9:35 AM), https://www.rti.org.tw/news/view/id/2091158 [https://perma.cc/CYJ3-C7D4].

[10] *Id.*

[11] Lin, *supra* note 7.

"What are good guys? And what are bad guys? Do you have the right answer for it?"[12] This melodrama production team mined events, character networks, different characters' viewpoints, and conflicts from 40,000 social media posts.[13] They tried to find out those hot topics that most people may have known about parts of them but are interested in learning more about them.[14] The idea was to create a story that could provoke broad discussion.

2. *Who Killed the Good Man*

Who Killed the Good Man depicted the post-graduate lives of three high school classmates.[15] It portrayed the financial difficulties and heavy debt burdens that many young adults in a high-pressure society must confront when trying to start a business, own a house, or have a child. In the initial script, the main characters were around twenty-five to thirty years old. When they were students, their ambition was to participate in social movements after graduation, but instead, because of debt stresses, they became addicted to drugs. Additionally, the initial script crafted the main female character as a bank teller and a part-time tutor. However, after analyzing thousands of social media posts and court judgment databases with the keywords "debt" or "drugs," the script was refined to make the main characters older, and the social movements plot was deleted.[16] Data analysis team also suggested to change the main female character's jobs to debt manager and a part-time food delivery driver. In turn, these data analysis suggestions made the show more reasonable and relatable to real life.

---

[12] Women yu E de Juli (我們與惡的距離), MOVIECOOL, https://moviecool.asia/quote/6-jy04-xc [https://perma.cc/H2AG-44M2].
[13] Digital Transformation Research Institute of Institute for Information Industry, *supra* note 8.
[14] *Id.*
[15] Yang, *supra* note 9.
[16] *Id.*

3. How were They Made?

Neither *The World Between Us* nor *Who Killed the Good Man* were AI-generated works. Instead, both were created by human writers with the help of data analysis. Data analysis does not replace writers. Instead, it allows writers to request relevant information. For example, the writers of *The World Between Us* gave DSI five keywords to parse enormous amounts of data available on the internet.[17] During the process, natural language processing techniques were applied to repeatedly clean and filter out unneeded data.[18] The top ten most popular names and events-related keywords were spotted to explore hot topics sequentially.[19] However, humans were still needed to assist machines in precisely labeling characters and topics.[20]

Data mining was beneficial in uncovering topic knowledge graphs, hidden vocabularies, terminologies, and event context.[21] DSI data team leader Yu-Liang Xu says two types of stories can benefit most from big data: (1) a story with a large amount of open data; and (2) a story plot with diverse and mixed viewpoints.[22] The writer of *The World Between Us* recalls that unnoticed topics, keywords, and facts mined from big data helped inspire her ideas.[23] Similarly, the producer of *Who Killed the Good Man* said that big data

---

[17] Lin, *supra* note 7.

[18] Lian Yiting(連以婷), "Womenyue" Weihe You Suanmin Meiti Juese? Muhou Zicehui Dashuju Tuandui de Gong Zuo Jiemi! (《我們與惡》為何有酸民、媒體角色？幕後資策會大數據團隊的工作揭密！), TECHNEWS (June 24, 2019, 2:00 PM), https://technews.tw/2019/06/24/the-world-between-us/ [https://perma.cc/4D5Q-AG72].

[19] Digital Transformation Research Institute of Institute for Information Industry, *supra* note 8.

[20] *Id.*

[21] Digital Transformation Research Institute of Institute for Information Industry, *supra* note 8.

[22] Ye Guanyin (葉冠吟), *Dashuju Zhugong Taiju Chuangzuo Xiankai Xinxie* (大數據助攻 台劇創作掀開新頁), CENT. NEWS AGENCY (Taiwan) (Mar. 14, 2021, 12:20 PM), https://www.cna.com.tw/news/firstnews/202103145002.aspx [https://perma.cc/W6Q6-M5U3].

[23] Lin, *supra* note 7.

assisted in understanding the whole picture and current trends of the young generation depicted in the television show.[24]

B. DATA-DRIVEN TREND IN THE MOTION PICTURE INDUSTRY

Big data has become a big deal because of "the advent of new and improved technologies for the collection, storage, mining, synthesizing, and analysis of data . . . ."[25] These new and improved technologies form a complete data-value chain.[26] Moreover, data science can uncover valuable information from vast amounts of raw data by using new technologies of natural-language processing, pattern recognition, and machine learning combined with traditional means of statistics.[27] This information can help businesses make informed decisions. Consequently, whoever are better able to access and use big data will have comparative advantages over their competitors.[28]

1. What is Big Data?

Big data means any data with characteristics of large volume and variety.[29] An extensive motion picture library in the form of a DVD collection is the traditional type of data.[30] An apparent shortcoming of traditional-type data is that we cannot know its interrelationship with audiences from the data itself. However, when the content is digitized and streamed via the internet, a new type of data will automatically be produced by users' conduct whenever they watch these movies.[31] Users' watching, sharing, and

---

[24] Ye, *supra* note 22.
[25] Rubinfeld & Gal, *supra* note 2, at 341.
[26] *Id. at* 349.
[27] *Id. at* 342.
[28] *Id.*
[29] Baur et al., *supra* note 1, at 212.
[30] *Id.* at 211.
[31] *Id.*

commenting histories are examples of this new-type big data. This data grows quickly and enormously in the internet age.[32] Though it is just a user-generated byproduct, it leads the way to know the audiences' reactions toward the motion pictures.

Social media platforms, e.g., Facebook, Instagram, Twitter, LinkedIn, YouTube, also aggregate vast amounts of audience data. Today, people share individual opinions and experiences on these platforms anytime and anywhere. They provide an outlet for collecting audiences' opinions of motion pictures.[33] The comments of netizens can also provide inspiration for making motion pictures.[34] Taking *The World Between Us* as an example, DSI unearthed what might interest audience members by analyzing 20,700 posts on Facebook and 24,500 posts on PTT, the largest Taiwanese bulletin board system (BBS).[35] Legendary Entertainment's Applied Analytics Division was also dedicated to collecting audience data from social media.[36] However, the data collected from social media may be unstable or incomplete because of privacy and regulatory issues. Therefore, many motion picture studios have tried to build their own streaming platforms.

2. How Can Big Data Help in Filmmaking?

The actual filmmaking process involves a series of creative decisions. For example, "which scripts to greenlight, who's chosen to direct, which actors to cast in a given role and the diversity of the cast as a whole, where to shoot, the order to shoot scenes and how

---

[32] *Id.*

[33] Shirish Jeble et al., *Role of Big Data in Decision Making*, 11(1) OPERATIONS AND SUPPLY CHAIN MANAGEMENT 36, 38 (2018).

[34] *See* Felix M. Simon & Ralph Schroeder, *Big Data Goes to Hollywood: The Emergence of Big Data as a Tool in the American Film Industry*, SECOND INT'L HANDBOOK OF INT. RSCH. 10 (2019).

[35] Lin, *supra* note 7.

[36] Michael Krigsman, *"Moneyball" for Movies: Data and Analytics at Legendary Entertainment*, CKOTALK (Feb. 09, 2018), https://www.cxotalk.com/episode/moneyball-movies-data-analytics-legendary-entertainment [https://perma.cc/B9T5-6EK9].

to tailor different cuts of the film to different audiences . . ." and more.[37] Data science can help filmmakers make informed decisions in numerous aspects, even in content creation. For example, a team of U.K. scientists categorized six formulas after analyzing 6,147 movie scripts.[38] They concluded "man in a hole – a fall followed by a rise" was the most profitable formula.[39]

Another example is a script evaluation service offered by a former statistics professor, Vinny Bruzzese. Through analyzing similar movies' audience data and surveying 1,500 potential audiences, he found that a targeting demon in horror movies would do better at the box office than a demon being summoned. [40] It is critical to note that Bruzzese's thesis is not uncovered by machines but by statistics and survey results. The last example is *House of Cards* (2013), an original show on Netflix. It was reported that Netflix knew from the beginning that viewers would like the show because of the favorable data collected on David Fincher, the director, Kevin Spacey, the main actor, and the British version of *House of Cards*.[41] Some critics thought "it just isn't true that you can rely on data completely,"[42] but a writer has complained about that a streamer gave up a nearly done

---

[37] Eric Avidon, *Data drives the decisions behind making movies*, TECHTARGET (June 12, 2020) https://searchbusinessanalytics.techtarget.com/feature/Data-drives-the-decisions-behind-making-movies [https://perma.cc/U2R7-GUHP].
[38] Mark Brown, *Scientists uncover formula for box office movie success*, GUARDIAN (July 24 2018), http s://www.theguardian.com/film/2018/jul/24/uk-scientists-uncover-formula-box-office-movie-success [https://perma.cc/XK9G-EQXG].
[39] *Id.*
[40] Brooks Barnes, *Solving Equation of a Hit Film Script, With Data,* N.Y. TIMES, (May 5, 2013) https://www.nytimes.com/2013/05/06/business/media/solving-equation-of-a-hit-film-script-with-data.html [https://perma.cc/HMU4-336B].
[41] Shane Atchison & Jason Burby, *Big data and creativity: What we can learn from 'House of Cards'*, THE NEXT WEB (Mar. 20, 2016) https://thenextweb.com/insider/2016/03/20/data-inspires-creativity/ [https://perma.cc/8HTV-X9Y5].
[42] Tim Wu, *Netflix's Secret Special Algorithm Is a Human*, NEW YORKER (Jan. 27, 2015) https://www.newyorker.com/business/currency/hollywoods-big-data-big-deal [https://perma.cc/DR32-E5JH].

script just because it ran an algorithm, and the data said the genre would not perform well.[43]

3. Making Content by Knowing Audience

In Hollywood, "give the audience what it wants" is the rule of thumb.[44] This is because audience opinions are directly related to revenue. Big data is considered the key to success for the film industry because it can precisely and quickly detect the audience's favorites. Such a task is difficult for humans to do on their own. Data technology can find out the bestselling materials by parsing a vast amount of data to make suitable formulas for target audiences. In turn, formulas can be refined and redefined by investigating more information from the unique culture of a specific formula's audience. The more thoroughly the audience's likes can be understood, the more sophisticated the formula can be developed.

Some have long argued that there are only a small number of fundamental story plots that any storyteller follows.[45] The most basic three-point plot introduced by Aristotle is "beginning, middle, and end."[46] The seven basic plots summed up by Christophe Booker are "overcoming the monster," "rags to riches," "the quest," "voyage and return," "comedy," "tragedy," and "rebirth."[47] John G. Cawelti defined "[a] formula is a combination or synthesis of a number of specific cultural conventions with a more universal story form or archetype."[48] In detail, a formula has two essential parts. First, its elements, such as setting,

---

[43] Writers Guild of America, *Why We Strike*, https://www.wgacontract2023.org/member-voices/why-we-strike [https://perma.cc/ZUG3-QD87].

[44] HORTENSE POWDERMAKER, HOLLYWOOD: THE DREAM FACTORY 36 (1950).

[45] CHRISTOPHER BOOKER, THE SEVEN BASIC PLOTS: WHY WE TELL STORIES 6 (2004).

[46] Piotr Mirowski et al., *Co-Writing Screenplays and Theatre Scripts with Language Models: An Evaluation by Industry Professionals*, ARXIV 2209.14958 3 (2022).

[47] BOOKER, *supra* note 45.

[48] JOHN G. CAWELTI, ADVENTURE, MYSTERY, AND ROMANCE: FORMULA STORIES AS ART AND POPULAR CULTURE 6 (1976).

characters, and plots, must make sense of the specific cultural convention. Second, the formula must have "particular sorts of story patterns and effects . . ." complying with "universal or transcultural concepts of literary structure . . . ."[49] like adventure, romance, mystery, or melodrama. In short, a formula is made with the audience's cultural senses and shared experiences. When a storyteller breaks away from these rules, audiences know something is wrong and feel unsatisfied.

Most audiences watch formulaic movies to seek escape and entertainment.[50] Thus, the motion picture industry has a strong tendency to follow formulas that have proven successful.[51] Indeed, formulas reflect the "characteristic of the people who enjoy these stories . . . ."[52] Different formulas represent different "attitudes and social characteristics of the audiences . . . ."[53] The more experiences that audiences share with a movie, the more they enjoy watching it.[54] Some degree of familiarity gives the audience feelings of "satisfaction and a basic emotional security. . ." and "a sense of what to expect . . . ."[55] With this common ground, filmmakers can effectively convey their ideas and viewpoints to audiences.[56] Some form of standardization is also beneficial to filmmakers to produce new works rapidly and efficiently. This explains why movie screens are often full of reboots, spin-offs, sequels, and prequels.[57]

---

[49] *Id.* at 6-7.
[50] *Id.* at 13.
[51] POWDERMAKER, *supra* note 44, at 40-43.
[52] CAWELTI, *supra* note 48 at 298.
[53] *Id.*
[54] *Id.* at 8-9.
[55] *Id.* at 9.
[56] *Id.* at 8-9.
[57] *See* Thomas Wick, *Demand More Original Films, Not Remakes, From Hollywood*, THE PITT NEWS (Mar. 19, 2017) https://pittnews.com/article/117993/opinions/demand-original-films-not-remakes-hollywood/ [https://perma.cc/XR8F-6VVB]; Stephen Follows, *The prevalence of sequels, remakes and original movies*,

The interrelationship between formula and creativity appears incompatible, but they are complementary. Margaret Boden defines three forms of creativity: (1) combination (making unfamiliar combinations of familiar ideas); (2) exploratory (exploring new ideas in a given conceptual space); and (3) transformation (bringing supposedly impossible ideas into a conceptual space).[58] On the one hand, formulas limit creativity because "[e]ach formula has its own set of limits that determine what kind of new and unique elements are possible without straining the formula to the breaking point."[59] On the other hand, a formulaic movie without adding "some unique or special characteristics of its own . . . ." is boring and mediocre.[60] Audiences welcome creativity that brings in new elements or novel interpretations without destroying the formulas. When the formulas are only slightly altered, audiences' inherent pleasure and definite expectation for conventional patterns are intensified.[61] Some novel attempts may become so successful that they ultimately become new parts of the current formula or even transform a current formula into a new one.[62]

4. Big Data and Streaming Platforms

In the past, to gauge audience opinion, filmmakers could only rely on theater

---

STEPHEN FOLLOWS (Apr. 30, 2018), https://stephenfollows.com/the-prevalence-of-sequels-remakes-and-original-movies/ [https://perma.cc/ED9F-RRNA]; Elle Love, *Why can't Hollywood come up with new ideas for movies?*, GATEWAY (Nov. 1, 2019), http://unothegateway.com/why-cant-hollywood-come-up-with-new-ideas-for-movies/ [https://perma.cc/H55N-4HHH]; Zaynab Zlitni, *Is Hollywood running out of ideas?* UNIV. DAILY KANSAN (Feb. 27, 2023), https://www.kansan.com/opinion/is-hollywood-running-out-of-ideas/article_399774fe-b6ed-11ed-be6a-bfaab38dc00a.html [https://perma.cc/R4J6-D67L].

[58] Margaret A. Boden, *Artificial Intelligence and Creativity: A Contradiction in Terms?*, *in* THE PHILOSOPHY OF CREATIVITY: NEW ESSAYS 224, 228 (Elliot Samuel Paul & Scott Barry Kaufman ed., 2014).

[59] CAWELTI, *supra* note 48 at 10.

[60] *Id.*

[61] *Id.* at 9.

[62] *Id.* at 12.

owners' feedback or audience member surveys,[63] but this information was hardly reliable. Over-the-top (OTT) services have changed the entertainment market worldwide thanks to streaming technology. The years 2019 and 2020 were particularly competitive for the video streaming market. The traditional content providers, Disney, WarnerMedia, and NBCUniversal, all joined the streaming war to challenge the leading status of Netflix.[64] Apple also launched its own streaming platform, Apple TV+.[65] In China, iQIYI and QQ video are the two streaming giants, each having over one hundred million subscribers.[66] Through online interaction with millions of consumers every day, OTT providers can gather vast amounts of audience data first-hand.

For filmmakers, owning a streaming platform ensures steady access to complete audience data and helps them identify the target audience for a niche product. In the days when movies were still novel, entertainment options were few, and audience education level was low; thus, movie audiences were easily pleased.[67] But today, "more important than any other change is the loss of homogeneity in the movie audience."[68] Therefore, instead of making big hits for mass audiences, a new thinking is making, marketing, and distributing films for target audiences.[69] This newer business model is actually based on

---

[63] Wu, *supra* note 42.

[64] Ashley Rodriguez, *How Disney Plus, HBO Max, and NBC's Peacock streaming service stack up in the battle against Netflix*, INSIDER: MARKETS INSIDER (Sept. 17, 2019) https://markets.businessinsider.com/news/stocks/disney-plus-hbo-max-nbc-streaming-comparison-price-release-content-2019-7-1028409686 [https://perma.cc/PEM7-RNS7].

[65] Kaitlin Thomas, *Everything We Know About Apple's Streaming Service, Apple TV+*, TVGUIDE (Mar. 30, 2020) https://www.tvguide.com/news/apple-tv-plus-streaming-service/ [https://perma.cc/2ZWS-EVWF].

[66] Aiqiyi: 2019nian Wangluo Dianying Hangye Baogao (爱奇艺：2019 年网络电影行业报告), 199IT (Jan. 17, 2020), http://www.199it.com/archives/998277.html [https://perma.cc/7X5Q-W3S7].

[67] *See* POWDERMAKER, *supra* note 44, at 297.

[68] *Id.*

[69] *See* JP, *Data-Driven Film Making – Cinema's Future?*, HARVARD BUSINESS SCHOOL: DIGITAL INITIATIVE (Nov. 22, 2015) https://digital.hbs.edu/platform-digit/submission/data-driven-film-making-cinemas-future/ [https://perma.cc/JY4F-7BHD].

the "Long Tail" theory developed by Chris Anderson in 2004, in opposition to the so-called 80-20 rule.[70]

The Long Tail theory suggests that accumulated revenue from selling various niche products to small groups of customers could be more profitable than selling hit products to the general public. The following conditions are necessary to achieve a Long Tail profit: (1) a large volume of diverse products, (2) low storage and distribution costs, and (3) precise marketing to the right customers. Netflix is considered a successful practice of the Long Tail theory by providing a large variety of streaming movies to accommodate the various tastes of numerous subscribers.[71] By using big data to identify the interested audiences and their needs, films from any genre can be made to meet market demands and make a profit.[72]

Moreover, streaming platforms introduce new movie formats, for example, interactive movies. Interactive movies break the mold by segmenting movies into separate sections and transferring some creative decision-making powers from the filmmaker to the audience by providing them with multiple options. The film *Black Mirror: Bandersnatch* (2018) was the first interactive movie on Netflix.[73] It contained five main endings and one trillion permutations.[74] As of May 2023, Netflix has provide twenty-four interactive

---

[70] *See Rethinking the Long Tail Theory: How to Define 'Hits' and 'Niches'*, KNOWLEDGE AT WHARTON (Sept. 16, 2009), https://knowledge.wharton.upenn.edu/article/rethinking-the-long-tail-theory-how-to-define-hits-and-niches/ [https://perma.cc/WL2P-KD9U].
[71] Dillon Berjani, *How Netflix is making use of the "Long tail" theory to attract new subscribers*, MEDIUM (Jan. 8, 2018) https://medium.com/@dillon.berjani/how-netflix-is-making-use-of-the-long-tail-theory-to-attract-new-subscribers-e4a96923ab6d [https://perma.cc/6J6Q-8L6Q].
[72] *See* JP, *supra* note 69.
[73] Zack Sharf, *'Black Mirror: Bandersnatch' Contains 5 Endings and 1 Trillion Story Combos, Changes the Way Netflix Streams*, INDIEWIRE (Dec. 28, 2018) https://www.indiewire.com/2018/12/black-mirror-bandersnatch-endings-one-trillion-story-combinations-netflix-streaming-1202031075/ [https://perma.cc/K4TY-JSZJ].
[74] *Id.*

titles.[75] iQIYI has also released the first Chinese interactive film, *His Smiles* (2019). This film contains five main storylines, twenty-one selection nodes, and seventeen endings.[76] An interactive movie's audience usually has two options waiting for them to choose at some points. The options could be as simple as choosing which cereal brand to eat for breakfast or as difficult as accepting or refusing a job offer.[77] The film will continue with the defaulted option if the audience does not choose within the specified time limit.[78]

Unlike linear motion pictures, interactive motion pictures' audiences are not passively watching movies. Rather, they are guiding the story. Interactive motion pictures have somewhat modified the rule of "giving the audience what it wants" to a new model of "letting the audience make what it wants." Consequently, "the convenient dichotomy between author and user evaporates in an interactive environment."[79] On the one hand, streaming platforms can collect more detailed data, such as audience preferences for plots through interactive movies. On the other hand, since including unlimited options in an interactive motion picture is impractical and wasteful, only potential options that may interest audiences will be considered. Therefore, big data will foreseeably be more critical in guiding filmmakers on what options to include in interactive motion pictures.

---

[75] Netflix de Hudongshi Jiemu yu Dianying (Netflix 的互動式節目與電影), NETFLIX, https://help.netflix.com/zh-tw/node/62526 [https://perma.cc/Q2R7-D83T].

[76] Li Xiazhi (李夏至), *Wuda Zhixian Shiqizhong Jieju Hudongju Guanzhong Maizhangma?* (五大支线十七种结局，互动剧观众买账吗？) BEIJING DAILY (北京日报) (June 25, 2019), http://m.xinhuanet.com/ent/2019-06/25/c_1124665831.htm [https://perma.cc/5PV8-YJ5S].

[77] Hannah Shaw-Williams, *Black Mirror: Bandersnatch Choices Guide - Best Decisions & All Outcomes Explained*, SCREEN RANT (Dec. 28, 2018), https://screenrant.com/black-mirror-bandersnatch-choices-guide-best-decisions/ [https://perma.cc/C2WY-GT4M].

[78] *Id.*

[79] Marshall Leaffer, *Protecting Authors' Rights in a Digital Age*, 27 U. TOL. L. REV. 1, 9 (1995).

II. DEFINITIONS

First and foremost, "data-driven work" and "artificial intelligence (AI) generated work" are distinct concepts in this dissertation, and they have different copyright concerns (see **Figure 1**). Data-driven work focuses on what the work is made of (the human author's subjective creative choices or the objective results of data collection and analysis). AI-generated work focuses on who makes the work (human authors or machines). The basic questions regarding copyright protectability and who should be the author have long been debated for AI-generated work.[80] But data-driven work has not yet attracted much attention from scholars.[81]

Figure 1 Data-Driven Work vs. AI-Generated Work

---
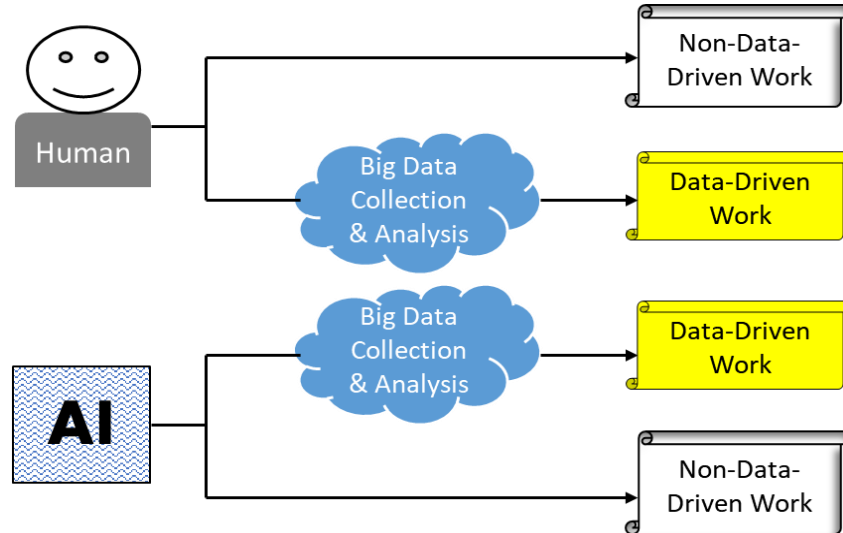
[80] *See* MARSHALL A. LEAFFER, UNDERSTANDING COPYRIGHT LAW 108 (7th ed. 2019); *see also* Andres Guadamuz, *Artificial intelligence and copyright*, WIPO MAGAZINE (Oct. 2017), https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html [https://perma.cc/WRW7-LWCT].
See also Huang Chieh, WHERE IS THE AUTHOR: THE COPYRIGHT PROTECTION FOR AI-GENERATED WORKS, 107 MAURER THESES AND DISSERTATIONS (2022), https://www.repository.law.indiana.edu/etd/107.
[81] *See* Raustiala & Sprigman, *supra* note 6, at 1565,

A. DATA-DRIVEN WORK

Data-driven, an adjective, means "determined by or dependent on the collection or analysis of data."[82] *Data-driven processing*, another name for *forward chaining* in computer science, means:[83]

> A control procedure used in problem solving and rule-based systems. When a data item matches the antecedent part of a rule, then the conditions for that rule have been satisfied and the consequent part of the rule is executed. The process then repeats, usually through some form of conflict-resolution mechanism to prevent the same rules firing continuously.

Law Professor Kal Raustiala and Professor Christopher J. Sprigman name "the use of consumer data to directly shape content creation" as *data-driven creativity*.[84] In this dissertation, what I refer to as *data-driven work* is a work made compellingly by the collection or analysis of data rather than by the maker's personal intuition or artistic judgment.[85] The data can be consumer data or any data available. It is critical to note that a data-driven work can be completed either by a human author(s) or by machine (AI), but part of or all of the creative choices are dominated by data collection and analysis results.

Theoretically, humans are not obligated to follow any suggestions extracted from audience data. But there may only be one reasonable choice in some situations. For example, in a hypothetical romantic scene involving whether Jack will kiss Rose, assuming data finds most audiences prefer that he does, a reasonable writer will follow the audiences' inclinations. In other cases, there might be several creative choices, but the data strongly

---

[82] *data-driven*, OXFORD DICTIONARIES, OXFORD UNIVERSITY PRESS. https://premium.oxforddictionaries.com/us/definition/english/data-driven (accessed via Oxford Dictionaries Online on August 14, 2023).

[83] BUTTERFIELD, A. & GERARD EKEMBE NGONDI, A DICTIONARY OF COMPUTER SCIENCE (7th ed. 2016).

[84] *See* Raustiala & Sprigman, *supra* note 6, at 1560.

[85] See *Data-driven*, WIKIPEDIA, https://en.wikipedia.org/w/index.php?title=Data-driven&oldid=112014906 5 [https://perma.cc/6RLH-66WA].

recommends a specific one. For example, suppose there are four creative possibilities for a romance: (1) a poor boy vs. a poor girl; (2) a poor boy vs. a rich girl; (3) a rich boy vs. a poor girl; and (4) a rich boy vs. a rich girl. If the data reveals that 99% of audiences want to watch the possibility (2) poor boy vs. rich girl, the writer's only reasonable choice is to draft the story based on this possibility.

B. AI-GENERATED WORK

AI-generated work refers to "a work independently created entirely by a computer, without input from any human 'author.'"[86] Not all works generated by AI are data-driven works if data is not used during the generating process. For example, some scholars define AI as machines that can think independently without needing massive data to train on.[87] If such an AI can freely generate a film script without being subject to the data fed into it, then the script is not a data-driven work.

It is worth mentioning that "data-driven work" and "AI-generated work" are not mutually exclusive concepts in this dissertation. The two natures can coexist in the same work if the work is generated by AI and is also made compellingly by the collection or analysis of data. Modern AI models are fed into massive data to generate output. Therefore, current AI-generated work shall have a significant overlap with data-driven work. For example, a large language model like GPT-3 can generate a human-like film script (see *infra* Chapter Three. III. B). Such a script is considered both a data-driven work and AI-

---

[86] LEAFFER, *supra* note 80, at 108.

[87] James Somers, *The Man Who Would Teach Machines to Think*, THE ATLANTIC (Nov. 2013), https://www.theatlantic.com/magazine/archive/2013/11/the-man-who-would-teach-machines-to-think/309529/?fbclid=IwAR0Gvjha5PxS5p47x0pbIEp31KhdS2wmVfCkPpWTuG50wSa84ok96CyizB4 [https://perma.cc/PPQ4-TJXS].

generated work in this dissertation because GPT-3 is trained on massive data to generate text.

III. PROBLEM STATEMENT

This dissertation aims to define an appropriate copyright protection for motion pictures in the data-driven era. Since the conventional filmmaking creative process has changed due to new data technologies, the current approach to determining the copyright scope of motion pictures must be updated. Because a data-driven motion picture's copyrightability and if it infringes on other works are the premise of its scope of copyright. Therefore, data-driven work's copyrightability and copyright infringement issues are studied altogether in this dissertation.

Typically, a film's creative process starts with a theme and goes through five steps: (1) preparation (gathering raw materials); (2) conscious work (studying raw materials); (3) incubation (synthesizing information subconsciously); (4) illumination (suddenly birthing an idea); and (5) verification (implementing the idea).[88] Once it is time for writing, the basic steps of a feature screenplay are treatment, first draft, rewrite, and polish.[89] If an outline is required, the outline will be made before the treatment.[90]

Accordingly, I divide the scriptwriting process into eight steps: (1) theme setting up; (2) raw materials selection; (3) logline; (4) outline; (5) treatment; (6) first draft; (7) rewrite, and (8) polish (see **Table 1-1**). Apparently, different steps require various degrees

---

[88] PETER BLOORE, THE SCREENPLAY BUSINESS: MANAGING CREATIVITY AND SCRIPT DEVELOPMENT IN THE FILM INDUSTRY 147 (2012).
[89] DINA APPLETON & DANIEL YANKELEVITS, HOLLYWOOD DEALMAKING: NEGOTIATING TALENT AGREEMENTS FOR FILM, TV, AND NEW MEDIA 49-50 (2nd ed. 2010).
[90] BLOORE, *supra* note 88, at 16.

of creative input. For the convenience of comparative and interpretative purposes, I assign each writing step a creative score and transform the score into a percentage.[91] Suppose a writer contributes step (1) through step (3) of a script. His creative contribution will be 22% (3%+8%+11%=22%). If another writer contributes to step (4) through step (8), his creative contribution will be 78% (17%+19%+22%+14%+6%=78%).

Table 1-1 The Scriptwriting Process

	Step	Definition	Creative Score	Percent
1	Theme	Choose a theme	★ (1)	3%
2	Raw materials	Select needed raw materials	★★★ (3)	8%
3	Logline	Depict the main ideas	★★★★ (4)	11%
4	Outline	Construct a story framework	★★★★★ (5)	14%
5	Treatment	Build characters and main plots	★★★★★★★ (7)	19%
6	First draft	The first complete draft	★★★★★★★★ (8)	22%
7	Rewrite	Rewrite draft one or more times	★★★★★★ (6)	17%
8	Polish	Minor changes to the rewrite	★★ (2)	6%

In the past, human writers did all the above tasks independently or collaboratively. They could be assisted by human assistants, editors, electronic databases, or grammar checker software. But they independently made the final creative decisions to transform their ideas into language. In contrast, modern data technology can extract audience-favorite story elements from vast amounts of audience and content data. Even more, modern data technology anticipates what the audience will like. It is too hard for humans to reject these

---

[91] The score is not an absolute standard but is determined by the author's personal judgment on possible creative effort input.

insights because they are stressed to be precise, scientific, and objective. Meanwhile, AI writing tools built on large language models can generate decent-quality text word by word and sentence by sentence. In fact, those words and sentences also come from the precise analysis results of massive training data. Since the machines know what to write and how to write, it is possible for machines to perform low-creativity to high-creativity tasks throughout the scriptwriting process (see **Figure 2**).

Figure 2 Data-Driven Creativity Level
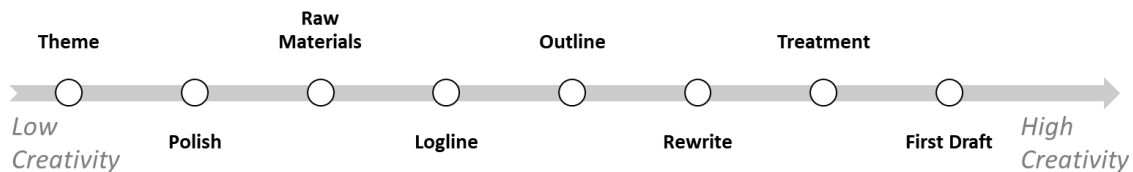
Entirely AI-generated scripts are still rare, short, and imperfect at the current stage because long content generated by machines can easily lose consistency and can be hilarious at times. [92] Nonetheless, human-machine collaborative scripts probably will become (or have become) a new normal. These human-machine collaborative works are copyrightable if the human authors contribute minimum originality input. The tricky question is how to define reasonable copyright protection. The data-driven trend has fundamentally changed the conventional filmmaking creative process; thus, I will review if the current method still works to determine the scope of motion picture copyright. If it does not, I will propose new and necessary approaches.

---

[92] Mirowski et al., *supra* note 46, at 10.

In the following three types, is it fair to grant the same scope of copyright protection over the three types of works? If not, how should the scope of copyright be determined?

[Type 1] Machine first, Human last (Machine → Human)

Example:

A film script is completed by a machine contributing to step (1) through step (3) (creative contribution 22%), and a human writer contributing to step (4) through step (8) (creative contribution 78%).

[Type 2] Human first, Machine last (Human → Machine)

Example:

A film script is completed by a human contributing to step (1) through step (3) (creative contribution 22%), and a machine contributing to step (4) through step (8) (creative contribution 78%).

[Type 3] Human and machine back and forth (Human ⇄ Machine)

Example:

A film script is completed by a human and a machine working together on step (1) through step (8) of the writing process. The machine generates content based on human-created content, and the human creates content based on machine-generated content. Either the human or the machine revises each other's content back and forth. The human and the machine contribute together to each step until all steps

are finished. The human creativity efforts and the machine's outputs are commingled. Overall, humans contribute 30%, and machines contribute 70%.

IV. RESEARCH LIMITATION

This dissertation mainly focuses on the **copyright issues** derived from the **data collection, management, storage** activities, and subsequent **data use** activities during the **motion picture** creation process and in the final motion pictures.

Other data-related legal issues, e.g., data ownership, privacy, free speech, trade secret, antitrust, and unfair competition, are not covered in this study. Also, other subject matters of copyright, e.g., literary works, musical works, dramatic works, pantomimes and choreographic works, pictorial, graphic, and sculptural works, sound recordings, and architectural works, are not in the scope of research unless necessary for this research.

This dissertation tries its best to provide a primary description of data technology's use in filmmaking. Recognizing data science and the filmmaking process involve special knowledge and the actual practices are varied, this dissertation cannot guarantee a completely unmistakable technical description.

V. LITERATURE REVIEW

The copyright system grants authors exclusive rights to their writings for promoting science and useful arts progress.[93] Although copyright limits market entry and public access to copyrighted works, it is believed to be a necessary evil for providing incentives

---

[93] *See* U.S. CONST. art. I, § 8, cl. 8.

for creativity.[94] In line with the framers' intent, the distinction between creative and non-creative activity is important.[95] The scholars argue that non-creative products could never obtain copyright just because they possess a "public good" character. For example, a motion picture shared on the internet is like a "public good" because it is hard to exclude others from copying and distributing. But if it is made without creative effort, it is not copyrightable.

To obtain copyright protection, a motion picture is not required to have materials that are "entirely new, and have never been used before; or even that they have never been used before for the same purpose" as long as its "plan, arrangement and combination of materials" has minimum originality from its author.[96] However, data-driven creativity has nakedly invaded creative decision-making in great detail. Considering this phenomenon is very different from the first digital disruption that only impacted the formats, storage mediums, reproduction, and dissemination of work, scholars name it the "second digital disruption."[97] Copyright is compared to a fair bargain between the author and the public.[98] When new technology changes the bargaining basis of social, economic, and cultural life, it will push copyright law changes.[99] Scholars hold copyright law's constant success depends on precise calibration without overprotection or under-protection.[100]

---

[94] Leaffer, *supra* note 79 at 5.
[95] Wendy J. Gordon, *Authors, Publishers, and Public Goods: Trading Gold for Dross*, 36 LOY. L.A. L. REV. 159, 165 (2002).
[96] Emerson v. Davies, 8 F. Cas. 615, 618 (C.C.D. Mass. 1845).
[97] *See* Raustiala & Sprigman, *supra* note 6, at 1611.
[98] Leaffer, *supra* note 79 at 5.
[99] *See* Peter Jaszi, Craig Joyce, Marshall Leaffer & Tyler Ochoa, *Statute of Anne: Today and Tomorrow*, 47 HOUS. L. REV. 1013, 1014 (2010).
[100] *See id.* at 1021.

In response to data-driven creativity, scholars assert that the need for the incentive provided by current copyright protection has reduced and proposed a shorter duration or a narrower scope of protection.[101] They argue that the streaming platforms have made motion pictures by collecting and analyzing massive quantities of consumer data. That guarantees the box office's success and reduces the investment risk in creativity. They also state that the audience's motivation for piracy has declined due to the flood of all-you-can-see choices on the internet. Conversely, another scholar disagrees with less copyright protection on the theories of product differentiation and attention scarcity.[102] He warns that reducing copyright protection will encourage new entrants to produce too many redundant motion pictures. That will harm the public welfare because human beings' attention is a relatively scarce asset in the information age.

Moreover, data-driven creativity undermines the conventional concept of romantic authors.[103] Increasing users' input has blurred the dichotomy between authors and users. A motion picture studio can only contract with creative persons under work-for-hire relationships to be an author. That gives creative persons substantial leverage to sell their intellectual creativity to the studios.[104] But if studios hire people to be "only the mere executants" of the audience's will,[105] creative people will lose their leverage. The essence of copyright law is to "protect the power of the creator against the power of owners of technologies that earn money exploiting the creations of authors."[106] In other words,

---

[101] *See* Raustiala & Sprigman, *supra* note 6, at 1604.
[102] *See* Jake Linford, *Copyright and Attention Scarcity*, 42 CARDOZO L. REV. 143 (2020).
[103] *See* Leaffer, *supra* note 79 at 9.
[104] *See* Peter Lee, *Reconceptualizing the Role of Intellectual Property Rights in Shaping Industry Structure,* 72 VAND. L. REV. 1197, 1240 (2019).
[105] *See* Jane C. Ginsburg, *The Concept of Authorship in Comparative Copyright Law*, 52 DEPAUL L. REV. 1063, 1072 (2003).
[106] Leaffer, *supra* note 79 at 12.

copyright is intended to reward authors rather than publishers.[107] Observing that data-driven creativity has transformed the filmmaker's role into a data collector and processor more than an originator and creator, scholars hold that the moral justification to protect the filmmaker's authorship had been diminished.[108]

There is more concern that data-driven creativity will enhance market concentration on large streaming platforms with a large amount of content and subscribers.[109] In addition to entry barriers, large platforms may abuse market power to charge higher prices. As a result, economically vulnerable groups' access to motion pictures would be prevented. Scholars allege that various access barriers exist in the collection, storage, synthesis, and analysis of big data.[110] The most complicated one is the legal barriers, e.g., data protection and privacy laws, data ownership, prohibiting storing data overseas, and intellectual property protection. On the contrary, other scholars doubt big data's harm to competition.[111] They say the economic scale, network effects, and barriers to entry are overstated and groundless. Their argument is based on the fact that new online firms constantly replace older ones.

Whoever is at fault, people might set hopes on competition law for safeguarding public access to information if necessary.[112] Unfortunately, antitrust law is thought to be powerless to promote competition by forcing large platforms to share data with competitors.[113] Arguably, piracy might be a useful means against large platforms' market

---

[107] Gordon, *supra* note 95, at 165.
[108] *See* Raustiala & Sprigman, *supra* note 6, at 1616.
[109] *Id.* at 1608.
[110] *See* Rubinfeld & Gal, *supra* note 2.
[111] D. Daniel Sokol & Roisin Comerford, *Antitrust and Regulating Big Data*, 23 GEO. MASON L. REV. 1129 (2016).
[112] *See* Jaszi et al., *supra* note 99, at 1021.
[113] *See* Raustiala & Sprigman, *supra* note 6, at 1611.

power, but that will also cause new problems. Scholars suggest reducing copyright protection in response to data-driven creativity. But reducing copyright protection to what degree is still waiting to be discussed.

VI. RESEARCH METHODOLOGY

This dissertation adopts historical research, documentary analysis, doctrinal research, comparative law, and theoretical analysis.

Historical research is applied to explore the early practices that the film industry employed to collect audience data for filmmaking before the emergence of data science. This research is also applied to explore old copyright legal issues about motion pictures.

Documentary analysis is employed to investigate how data science is applied to filmmaking by the motion picture industry.

Doctrinal research is used to find existing laws for motion picture copyright protection by analyzing and synthesizing valid statutes and precedents.

Comparative law research is adopted to compare the motion picture copyrights in the United States, China, and Taiwan. The United States and China are selected because both are the two biggest motion picture markets globally. Besides, Taiwan's copyright protection is like the United States, but its cultural background is similar to China, so Taiwan is included for comparison.

Theoretical analysis is used to study possible resolutions. Secondary sources are the primary sources for theoretical analysis.

VII. STRUCTURE OF THIS DISSERTATION

The remainder of this dissertation is divided into five chapters as below:

Chapter Two explains why data-driven trends impact motion pictures more than other types of works. This chapter takes Hollywood as an example for analyzing the historical, cultural, and economic reasons.

Chapter Three demonstrates the development of data-driven filmmaking. This chapter presents the data source, data technology, film analytic platforms, content creation tools, and how data is used in filmmaking practices.

Chapter Four investigates current motion picture copyright in the United States, China, and Taiwan. This chapter reviews what approaches are used to determine the motion picture's copyrightability, the scope of protection, and copyright infringement issues.

Chapter Five reviews copyright violation risks and copyrightability issues the motion picture industry confronted in the past and indicates the new challenges they should deal with in the data-driven era.

Chapter Six concludes the findings of this dissertation and proposes to update copyright protection for motion pictures. Mainly: (1) Adding copyright limitations or exceptions to the use of copyrighted film materials for text and data mining and AI model training in the motion picture creation process, and (2) Creating a voluntarily negotiated license mechanism backed up by a compulsory license for reusing copyrighted film materials in new motion pictures.

CHAPTER TWO: FEATURES OF MOTION PICTURE INDUSTRY

> *Technology* itself *does not drive the development of copyright law. Rather, copyright is driven by the changes in social life, economic organization, and cultural outlook that technology inspires.*[114]
>
> — Marshall Leaffer et al., *Statute of Anne: Today and Tomorrow*

We cannot completely understand motion pictures without knowing who makes them and how they are made. This chapter studies the history, culture, and value system of the motion picture industry, particularly Hollywood. I claim motion pictures, especially for Hollywood features, are more likely to use data-driven creativity than other types of works. The three forces motivating this tendency are the film industry's long technology-advocated custom, business-dominated culture, and audience-oriented practice. Therefore, a special concern for the impact on motion pictures' copyright protection is necessary and justified in response to this tendency.

I. TECHNOLOGY-ADVOCATED CUSTOM

A. BORN OF TECHNOLOGY

Queen Anne would never have imagined the invention of motion pictures in 1710, when the Statute of Anne, the world's first Copyright Law, was enacted.[115] On October 17, 1888, Thomas A. Edison first filed his idea for a device that would "do for the eye what

---

[114] *See* Jaszi et al., *supra* note 99, at 1014.
[115] *See id.* at 1017.

the phonograph does for the ear" to the United States Patent Office.[116] On August 24, 1891, Edison filed a patent for the Kinetograph (the camera for making movies) and the Kinetoscope (the viewer for watching movies).[117] On May 9, 1893, the public finally viewed the Kinetoscope at the Brooklyn Institute of Arts and Sciences.[118] However, the Achilles heel of the Kinetoscope was that it could only allow one viewer per time to watch a motion picture through a peephole. That made it difficult to compete with many later-invented projectors that allowed hundreds of viewers to watch a motion picture on a large screen simultaneously. Soon after, even the Edison Company gave up the Kinetoscope and developed a new projector in November 1896.[119]

The motion picture industry has put technology ahead of human creativity from the very beginning. Managing the camera and cutting the film were the most critical filmmaking skills.[120] The early film companies sold motion pictures in reels, and every reel's price was equal regardless of their content.[121] Also, the motion picture industry was criticized for caring more about outside factors than human creative expressions. [122] Hortense Powdermaker, the author of *Hollywood: the dream factory*, had ever sighed that in the Hollywood, the cameraman, sound engineer, and other technicians enjoyed more independence to try their ideas than the writers, directors, and actors could.[123] He also said

---

[116] *Origins of Motion Pictures*, LIBRARY OF CONGRESS, https://www.loc.gov/collections/edison-company-motion-pictures-and-sound-recordings/articles-and-essays/history-of-edison-motion-pictures/origins-of-motion-pictures/ [https://perma.cc/PD7M-272P].

[117] *Id.*

[118] *Id.*

[119] *Shift to Projectors and the Vitoscope*, LIBRARY OF CONGRESS, https://www.loc.gov/collections/edison-company-motion-pictures-and-sound-recordings/articles-and-essays/history-of-edison-motion-pictures/shift-to-projectors-and-the-vitoscope/ [https://perma.cc/QD3F-U6UJ].

[120] POWDERMAKER, *supra* note 44, at 186.

[121] PETER DECHERNEY, HOLLYWOOD A VERY SHORT INTRODUCTION 15 (2016).

[122] POWDERMAKER, *supra* note 44, at 287.

[123] *Id.* at 27.

that many motion pictures' feelings of excitement concentrated on external factors, such as noise, music, props falling, and thunderstorms, rather than the inner human emotions.[124]

Moreover, the emphasis on technology was reflected in the motion picture industry's business philosophy and revealed in Edison's initial strategy to protect his film business by patent instead of copyright.[125] In the early days, motion pictures were only deemed appendants to promote selling film machines. It was prevalent among competing machine manufacturers, including Edison's company, to duplicate each other's films to fit into their manufactured equipment. At first, Edison wanted to monopolize the motion picture business by dominating film technology, so he tried to exclude competitors through patent infringement litigations. He finally added the copyright claims after losing a crucial patent lawsuit. Because of Edison's advocate, the early courts recognized to protect motion pictures as photographs under the copyright law.[126] The oldest film, Fred *Ott's Sneeze* (1894), preserved in the U.S. Copyright Office, was indeed a set of forty-five still pictures.[127] It was not until 1912 that U.S. copyright law officially classified motion pictures as a subject of copyrightable works.[128]

B. CONTINUOUS USE OF NEW TECHNOLOGY

Technologies continue to reshape people's enjoyment and imagination of motion pictures. The sound film was argued to be the most extensive change in motion picture

---

[124] *Id.* at 287.

[125] PETER DECHERNEY, HOLLYWOOD'S COPYRIGHT WARS: FROM EDISON TO THE INTERNET 19-23 (2012).

[126] *See* Edison v. Lubin, 122 F. 240 (3d Cir. 1903); Am. Mutoscope & Biograph Co. v. Edison Mfg. Co., 137 F. 262 (C.C.D.N.J. 1905).

[127] *Early Motion Picture Productions*, LIBRARY OF CONGRESS, https://www.loc.gov/collections/edison-company-motion-pictures-and-sound-recordings/articles-and-essays/history-of-edison-motion-pictures/early-motion-picture-productions/ [https://perma.cc/3QL8-CH7X].

[128] Motion Picture Copyright Amendments, Pub. L. No. 62-303, 37 Stat. 488 (1912).

history.[129] Switching from silent films to sound films took nearly 20 years. Warner Bros. and Fox, then two small studios, were the pioneers in Hollywood to make sound films. Their success prompted major studios to agree with a single sound standard in 1930. Consequently, studios hired writers, musicians, and singers to provide dialogue, music, and sound effects for films. That fundamentally changed motion picture production and exhibition.

In the 1950s, Hollywood introduced color, widescreen, and 3-D technologies to defend it against the competition from television.[130] Soon after, Hollywood started to sell motion picture to television stations and produced TV shows to promote motion pictures. By the mid-1960s, Hollywood had successfully entered television. In the 1970s and 1980s, home video technology, e.g., Betamax, and VHS, provided a new distribution outlet and brought Hollywood more revenue than theaters by 1986.[131] The DVDs invented in 1995 further increased home video consumption for the sake of low-price, high-quality, large storage capacity, and digital format.[132]

Digital technology drove a new wave of the cinematic revolution. Filmmakers began to try computer-generated imagery (CGI) sparsely in the 1970s.[133] After long tries, the stained-glass knight of Steven Spielberg's *Young Sherlock Holmes* (1985) and the T-1000 liquid-metal assassin robot of James Cameron's *Terminator 2: Judgment* (1991) marked the milestones for digital characters.[134] The vivid T-rex of Steven Spielberg's

---

[129] DECHERNEY, *supra* note 121, at 37.
[130] *Id.* at 86.
[131] *Id.* at 101.
[132] *Id.* at 107.
[133] POWDERMAKER, *supra* note 44, at 117-118.
[134] *Id.* at 118.

Jurassic Park (1993) was another impressive progress. But still, only a little CGI was used.[135] The "technical breakthrough" of CGI was James Cameron's *Avatar* (2009).[136] Nearly 60% of the world in *Avatar*, Pandora, was made of CGI.[137] Nowadays, CGI is no longer something added in post-production but a big part of shooting. LED backgrounds allow the directors and actors to see the CGI settings in real-time, needless to imagine the special effects on the green screen anymore.[138]

Historically, the major studios relied on two abilities to maintain market power.[139] One is the ability to effectively manage technical and financial resources essential for filmmaking by economies of scale. However, this power is gradually undermined by those artists who can create professional quality works using cheaper digital camera devices and software.[140] The majors' second traditional advantage is controlling the scarce distribution and promotion channels. This ability makes them able to make the most profit by using a complex release strategy and charging discriminated prices in different release windows.[141] A typical release window is determined by the order from high-value to low-value consumers. This order is (1) theatrical, (2) hotel and airline, (3) DVD sell-through and rental, (4) pay-per-view cable TV, (5) pay cable network and subscription streaming, and

---

[135] Nickalus Rupert, *Jurassic Park: How the Dinosaurs were Created (Is It All CGI?)*, SCREENRANT (Dec. 07, 2020), https://screenrant.com/jurassic-park-movie-dinosaurs-create-cgi-effects-explained/ [https://perma.cc/KDY6-5UY9].

[136] Roger Ebert, *Cameron retains his crown*, ROGEREBERT.COM (Dec. 11, 2009), https://www.rogerebert.com/reviews/avatar-2009 [https://perma.cc/CPJ9-TAM4].

[137] Alex Dudok De Wit, *After A Decade, James Cameron Finally Admits That 'Avatar' Is Animated*, CARTOONBREW (Dec. 19, 2019), https://www.cartoonbrew.com/artist-rights/after-a-decade-james-cameron-finally-admits-that-avatar-is-animated-183990.html [https://perma.cc/LLP7-DQKE].

[138] Renee Laroche-Rheaume, *Goodbye, Green Screen — Here's the Tech That's Changing Film Forever!* GOODNESS-EXCHANGE (Nov. 20, 2020), https://goodness-exchange.com/the-volume-mandalorian-storytelling/ [https://perma.cc/4G3J-LNT3].

[139] MICHAEL D. SMITH & RAHUL TELANG, STREAMING, SHARING, STEALING: BIG DATA AND THE FUTURE OF ENTERTAINMENT 13 (2016).

[140] *Id.* at 103-104.

[141] *Id.* at 40.

(6) advertising-supported broadcasts. Nonetheless, widespread digital media, mobile video devices, and the internet have threatened the old release strategy.[142] The one-to-many broadcast model has changed to a many-to-many interactive model.[143]

Moreover, countless internet piracy and illegal peer-to-peer file sharing increase the customers' choices. This reality makes it difficult for the majors to maintain the old business model and forces them to compete with lower prices and more timely distribution channels.[144] For example, NBCUniversal has adopted a hybrid model. Some films are released simultaneously in theaters and on the Peacock streaming platform, while others are released in Peacock 45 days after theatrical release.[145] New digital content services, such as electronic sell-through (EST), video on demand (VOD), and subscription streaming, have become prevalent. Since 2019, the digital market share (48%) has exceeded the theatrical market share (42%) in the global entertainment market.[146] In 2020, due to the COVID-19 pandemic, the global digital market grew to 76%, whereas the global theatrical market declined to 15%.[147]

Furthermore, the change in distribution modality moves the change in content.[148] The decline in syndication revenue encourages television producers to make cheap reality

---

[142] *Id.* at 48.
[143] DECHERNEY, *supra* note 121, at 123.
[144] Jon M. Garon, Content, *Control, and the Socially Networked Film*, 48 U. LOUISVILLE L. REV. 771, 788-789 (2010).
[145] Lillian Rizzo, *NBCUniversal CEO Jeff Shell says movie business benefiting from new release model*, CNBC (Oct. 4, 2022, 11:56 AM), https://www.cnbc.com/2022/10/04/nbcuniversal-ceo-jeff-shell-says-movie-business-is-better-with-new-release-model.html [https://perma.cc/P6WP-4S6Z].
[146] The digital covers electronic sell-through (EST), video-on-demand (VOD), and subscription streaming. *See* THE MOTION PICTURE ASSOCIATION, INC, THEME REPORT 8 (2019), https://www.motionpictures.org/wp-content/uploads/2020/03/MPA-THEME-2019.pdf.
[147] *See* THE MOTION PICTURE ASSOCIATION, INC, THEME REPORT 43 (2020), https://www.motionpictures.org/wp-content/uploads/2021/03/MPA-2020-THEME-Report.pdf.
[148] *See* Garon, *supra* note 144, at 790-791.

shows instead of expensive scripted shows to be profitable in the initial release window. In addition, new streaming services providers, such as Netflix, Hulu, and Amazon Prime, take advantage of "the internet and data to provide lower-cost, better-tailored programming content directly to consumers."[149] This new competition drove a new wave of vertical integration between content creators like Time Warner, and content distributors like AT&T.[150] Besides, active and participatory consumers built many online communities and shared user-generated content over the internet. The majors have tried hard to develop ideas and promote motion pictures by interacting with these fan communities.[151]

Recalling the history of Hollywood's response to modern technologies as continuous "competition, adjustment, and triumph,"[152] I am optimistic that Hollywood will make the most out of data-driven creativity too.

II. BUSINESS-DOMINATED CULTURE

A. MOTION PICTURE AS A BUSINESS

[M]aking of movies is both a big business and a popular art.[153]

Movies are a mass median and to remain as such in our society they must

make a profit.[154]

— Hortense Powdermaker, *Hollywood: the dream factory*

---

[149] United States v. AT&T Inc., 310 F. Supp. 3d 161, 173 (D.D.C. 2018).
[150] *See id.*
[151] POWDERMAKER, *supra* note 44, at 126.
[152] *Id.* at 123.
[153] *Id.* at 25.
[154] *Id.* at 26.

The 1912 House report for amending the U.S. Copyright Act cannot speak more frankly about the business nature of motion pictures:[155]

> The occasion for this proposed amendment is the fact that the production of motion picture photoplays and motion pictures other than photoplays has become a business of vast proportions. The money invested therein is so great and the property rights so valuable that the committee is of the opinion that the copyright law ought to be so amended as to give to them distinct and definite recognition and protection.

Unlike other categories of work, motion pictures, especially for features and television programs, could hardly be made by an individual. The typical filmmaking participants are producer, writer, director, actor, cinematographer, cutter, musician, make-up man, set designer, and many other technicians.[156] Generally, filmmaking can be divided into three phases.[157] The first phase is preproduction which includes script development, production planning, casting, and hiring. The second phase is the production which involves the actual shooting of the film. The third phase is post-production which includes editing, recording, and synchronizing music and other sound effects.[158] This complicated process was analogized to a factory assembly line.[159] On the one hand, both have similar needs in space, timing, schedule, budget, multiple professionals' collaborations, and various ingredient combinations. Most workers do their parts without knowing the whole picture until the final product is completed. On the other hand, the motion picture industry

---

[155] H.R. Rep. No. 62-756 at 1 (1912).
[156] POWDERMAKER, *supra* note 44, at 30.
[157] Eileen Morley & Andrew Silver, *A Film Director's Approach to Managing Creativity*, HARV. BUS. REV. (Mar. 1977), https://hbr.org/1977/03/a-film-directors-approach-to-managing-creativity [https://perma.cc/9KGY-SSEL].
[158] *See id.*
[159] POWDERMAKER, *supra* note 44, at 30-31.

differs from other industries because motion pictures do not have uniform raw materials, and film artists need creative freedom to do their jobs.[160]

Because of the complexity, filmmaking costs are notoriously high. Generally, a movie made less than US$2 million is considered a low-budget one.[161] In the case of a major studio feature movie, the average production cost could reach US$65 million, and distribution and marketing could cost additional US$35 million.[162] According to the 2020 Theme Report of the Motion Picture Association,[163] 23% of American features cost less than US$1 million, 54% of American features cost between US$1 million and US$15 million, and 23% of American features can cost over US$15 million (see **Table 2-1**).

Table 2-1 U.S. Features Film Entering Production

Film budget (US$)	2016	2017	2018	2019	2020
$15M + budget	135(17%)	162(20%)	171(21%)	178(22%)	101(23%)
$1M-15M budget	376(48%)	387(48%)	405(50%)	423(52%)	243(54%)
<$1M budget	277(35%)	263(32%)	232(29%)	213(26%)	103(23%)
Total Films (all budgets)	788	812	808	814	447

(Source: MPA 2020 Theme Report)

In the case of a blockbuster, it could even cost more than US$300 million (see

---

[160] *Id.* at 288-289.

[161] Erin Pearson, *The Basics to Making a Low Budget Film*, Topsheet (Sept. 27, 2019, https://topsheet.io/blog/basics-to-making-a-low-budget-film [https://perma.cc/FA9F-ES9C].

[162] Annie Mueller, *Why Movies Cost So Much to Make*, Investopedia (July 31, 2022), https://www.inves topedia.com/financial-edge/0611/why-movies-cost-so-much-to-make.aspx#the-bottom-line [https://perma.cc/HZJ4-NGW7].

[163] The Motion Picture Association, *supra* note 147, at 43.

Table 2-2).[164] However, a big budget does not guarantee hefty profit. The studio might easily lose US$100 million over a single flop (see **Table 2-3**).[165]

Table 2-2 Biggest-Budget Films

Release Date	Movie	Production Budget (US$)	Domestic Gross (US$)	Worldwide Gross (US$)
04/26/2019	Avengers: Endgame	$400,000,000	$858,373,000	$2,797,800,564
05/20/2011	Pirates of the Caribbean: On Stranger Tides	$379,000,000	$241,071,802	$1,045,713,802
05/01/2015	Avengers: Age of Ultron	$365,000,000	$459,005,868	$1,395,316,979
12/18/2015	Star Wars Ep. VII: The Force Awakens	$306,000,000	$936,662,225	$2,064,615,817
05/24/2007	Pirates of the Caribbean: At World's End	$300,000,000	$309,420,425	$960,996,492

(Source: https://www.the-numbers.com/movie/budgets)

Table 2-3 Biggest Money Losers

Release Date	Movie	Approx. Income (US$)	Approx. Expense (US$)	Profit (US$)
03/11/2011	Mars Needs Moms	$26,761,641	$170,166,000	-$143,404,359
12/25/2020	Wonder Woman 1984	$117,132,400	$258,940,000	-$141,807,600
09/04/2020	Mulan	$58,615,640	$200,000,000	-$141,384,360
03/06/2020	Onward	$86,237,128	$217,240,000	-$131,002,872
03/09/2012	John Carter	$182,557,164	$295,824,000	-$113,266,836

(Source: https://www.the-numbers.com/movie/budgets)

---

[164] MOVIE BUDGET AND FINANCIAL PERFORMANCE RECORDS, THE NUMBERS, https://www.the-numbers.com/movie/budgets [https://perma.cc/ETL6-KJ4G].
[165] *Id.*

Under heavy financial stress, the first and foremost question for a movie project is whether it can make money after covering the enormous costs.[166] Thus, drawing audiences to the theater is always considered more important than fulfilling the creative desires of the writers, directors, actors, or any other film artists.[167] The Hollywood blockbusters developed since the 1970s reflect such a mindset.[168] To play safe, blockbusters make heavy use of existing well-known works and previously successful ideas. Therefore, many reboots, spinoffs, sequels, prequels, and remakes of popular hits continually appear on the big screen.[169] Some of them have even become franchises that seem endless.

B. BUSINESSPEOPLE ABOVE ARTISTS

The domination of the business executives over the artists in Hollywood is not surprising either in view of the history of the movies or in terms of American culture.[170]

— Hortense Powdermaker, *Hollywood: the dream factory*

Knowing filmmaking people helps us understand Hollywood's filmmaking culture. The most influential people in Hollywood are studio executives and producers, who are businessmen in nature.[171] They believe money can make good movies, search for "sure-fire" formulas,[172] and trust their own instincts and luck more than the artists' creativity.[173]

---

166 *See* POWDERMAKER, *supra* note 44, at 287.
167 *See id.*
168 *See id.* at 97-100.
169 *See* Thomas Wick, *supra* note 57; Stephen Follows, *supra* note 57; Elle Love, *supra* note 57; Zaynab Zlitni, *supra* note 57.
170 POWDERMAKER, *supra* note 44, at 28.
171 *See id.* at 28.
172 *See id.* at 25.
173 *See id.* at 113.

The executives have the paramount authority to decide scripts, castings, cutting, shooting location, budget, and settling disputes.[174] Today, the greenlight authority at Hollywood studios is often exercised by a committee of executives in order to pool opinions, avert risks, and share decision-making responsibilities.[175] The producers are people whom the executives delegate to supervise individual film projects.[176] They have the power to approve the artists' work before submitting it to the executives.[177] Executives and producers are so-called gatekeepers who actually control whether a story idea can be made into a motion pictures.[178]

Conversely, the writers, directors, and actors are artists who want to convey their thoughts to the moviegoing public.[179] However, they are considered cogs in a machine, and the executives and producers often restrain their creative autonomy.[180] The writers lose creative control over the scripts once they deliver the first draft.[181] The directors lose creative control over the films once the shot finishes.[182] In Hollywood's golden age, most artists had long-term contracts with studios. But since *United States v. Paramount Pictures, Inc.* (1948) prevented studios' control of theaters,[183] studios have made fewer films and eliminated long-term contracts with most directors, writers, and actors.[184] Freelance artists

---

[174] *See id.* at 87.

[175] Stephen Galloway, *The New Power of "Yes": Who Actually Has Greenlight Authority at the Movie Studios*, HOLLYWOOD REPORTER (Feb. 15, 2017), https://www.hollywoodreporter.com/news/general-news/new-power-yes-who-actually-has-greenlight-authority-at-movie-studios-976236/ [https://perma.cc/C4K9-BHZT].

[176] POWDERMAKER, *supra* note 44, at 111-112.

[177] *Id.* at 130.

[178] BLOORE, *supra* note 88, at 25, 134.

[179] POWDERMAKER, *supra* note 44, at 52.

[180] *Id.* at 287.

[181] BLOORE, *supra* note 88, at 95.

[182] *Id.* at 104.

[183] United States v. Paramount Pictures, Inc., 334 U.S. 131 (1948).

[184] DECHERNEY, *supra* note 121, at 96.

rely on talent agencies to seek job opportunities on their behalf.[185] The agencies can earn ten percent commission from their clients' income.[186] Large agencies may sell film projects to the studios by packaging multiple talent clients under their roofs, e.g., a writer's script, a director, and several actors.[187] In such cases, the agencies would charge the studios a packaging fee regardless of their clients' earnings.[188] Some largest agencies even form studios to hire their clients.[189] This circumstance worries many film artists that the agents may put their own interests ahead of their clients' interests.

The writers and directors are the most important artists in selecting, coordinating, and arranging the elements of a motion picture. Their attitudes toward data-driven creativity might be like their predecessor Dudley Nichols' view for Gallup's scientific audience research (see *infra* Part III. B). "This new scientific era in movie making [is] sometimes tough on writers and filmmakers who can work with enthusiasm on something they personally happen to like and can grow strangely apathetic when told to work on public demand."[190] In the following parts, I will discuss the roles of the writers and directors in Hollywood. In sum, the artists' attitudes toward data-driven trends might be essential but not determinative for the businessmen's decision-making.

---

[185] Nowadays, the big four agencies are Creative Artists Agency (CAA), William Morris Endeavor (WME), United Talent Agency (UTA), and International Creative Management Partners (ICM). *See* WRITERS GUILD OF AMERICAN WEST, HOW THE MAJOR HOLLYWOOD TALENT AGENCIES PUT THEIR INTERESTS AHEAD OF THEIR CLIENTS' INTERESTS, https://www.wga.org/uploadedfiles/members/member_info/agency_agreement/wga_no_conflict_no_interest_19.pdf [https://perma.cc/FK6D-MBB2].
[186] *See id.*
[187] Christopher Schiller, *What is Film Packaging?* SCRIPT (Feb. 9, 2018), https://scriptmag.com/features/it-depends-what-is-film-packaging [https://perma.cc/VL3X-B7QK].
[188] *See* WRITERS GUILD OF AMERICAN WEST, *supra* note 185.
[189] Writers Guild of American West, *The Truth About Agency Studios*, YOUTUBE (Nov. 11, 2019), https://www.youtube.com/watch?v=EaXQ84Hn6_Y&t=4s.
[190] SUSAN OHMER, GEORGE GALLUP IN HOLLYWOOD 2 (2006).

1. Writer

> *Of all the creative workers in Hollywood, the writers are the most frustrated because they are allowed to function least as artists. They write to dictation, expressing someone else's fantasies, and even this is later changed and mangled by others . . . . the actor and director at least do play their professional roles regardless of the outcome, while the writer rarely even does that.*[191]

> — Hortense Powdermaker, *Hollywood: the dream factory*

Either the producer or the writer may produce a story idea that may be an original one or based on pre-existing works.[192] But the writer is responsible for translating a story idea into a screenplay containing characterization, scenes, and dialogues. Stars and genres have long controlled Hollywood motion pictures' consistency and stability.[193] The writer is supposed to know what the producer wants and can put all the elements (characters, settings, tensions, endings, and subplots) together when given the stars and genres.[194]

Despite the close relationships between producers and writers, the producers rarely respect the writers.[195] The writers are considered expendable and replaceable[196] and their function is to "transfer the ideas and fantasies of the producer onto paper."[197] The producers

---

[191] POWDERMAKER, *supra* note 44, at 292-293.
[192] BLOORE, *supra* note 88, at 11.
[193] DECHERNEY, *supra* note 121, at 30.
[194] *See Id.* at 34.
[195] POWDERMAKER, *supra* note 44, at 167.
[196] Chad Fitzgerald, *Cinema Law: Screenwriter Rights*, MOVIEMAKER (Dec. 14, 2009), https://www.moviemaker.com/cinema-law-screenwriter-rights-20091214/ [https://perma.cc/65DU-2PZY].
[197] POWDERMAKER, *supra* note 44, at 152.

will pay a massive price for a gadget to replace disobedient writers if they can.[198] For example, a Marvel executive and the directors joined with the scriptwriters, Stephen McFeely and Christopher Markus, to develop the script of *Captain America: Civil War (2016)* from the outset.[199] The final says belonged to the producer Kevin Feige even though the writers claimed they enjoyed a fair amount of freedom. They needed to rewrite the script if the characters' use rights changed. They also had to collaborate with other Marvel movie directors and writers to ensure different Marvel movies' internal continuity.

Generally, the writing steps of a feature screenplay are divided into treatment, first draft, rewrite, and polish.[200] A script can be completely reworked in the rewrite stage, the so-called "page one rewrite."[201] Even if a writer initiates his own ideas, the script might need to be rewritten to move the project forward, especially when the script is attached to a director with a different vision or an actor unfitting the character well.[202] In these cases, most producers would rather hire another writer to do so because adaptation usually needs many compromises.[203] Expecting the original authors to rewrite their hard works is usually more challenging. For example,[204] *Jaws* (1975) is adapted from Peter Benchley's same-name novel "Jaws." Unsatisfied with Benchley's three screenplay drafts, director Steven Spielberg turned out his version. Howard Sackler, John Milius, Robert Shaw, and Carl

---

[198] *See id.* at 167.
[199] Peter Suderman, *How Marvel built such an impressive movie universe*, Vox (May 9, 2016), https://www.vox.com/2016/5/9/11595344/marvel-cinematic-universe-captain-america-avengers [https://perma.cc/K5FG-THHC].
[200] *See* Appleton, supra note 89, at 49-50.
[201] *Id.*
[202] Schiller, *supra* note 187.
[203] Adam T. Dean, The Paradox of Creativity and Business in Feature Hollywood Filmmaking: The Relationship Between Motion Picture Production and Budgeting, M.A. thesis 32 (August 2005) (University of North Texas, https://digital.library.unt.edu/ark:/67531/metadc4885/: accessed August 15, 2023).
[204] Sven Mikulec, *'Jaws': The Groundbreaking Summer Blockbuster that Changed Hollywood, and Our Summer Vacations, Forever*, Cinephilia & Beyond, https://cinephiliabeyond.org/jaws-groundbreaking-summer-blockbuster-changed-hollywood-summer-vacations-forever/ [https://perma.cc/M93H-H8FW].

Gottlieb also contributed to the script even though only Benchley and Gottlieb were credited as writers.

If producers think it necessary,[205] multiple writers may be hired to work on a screenplay based on their expertise simultaneously or subsequently.[206] However, not every writer can receive credits on a motion picture until recently.[207] The Writers Guild of America (WGA) strictly regulates the writer credits and writer numbers of a feature.[208] The most common credits are "written by," "story by," and "screenplay by." The story is only a general idea compared to the screenplay (see **Table 2-4**).[209] Only the writer who creates the story and writes the final script can get the "written by" credit. To receive "screenplay by" credit, a writer's material contribution to the final script must exceed 33%. No more than two writers can share the "written by," "story by" or "screenplay by" credits unless the exception is applied.[210] To have all the participating screenwriters get credits, an "additional literary material" credit was added in 2021.[211] It applies to writers who contribute literary material to a movie but do not receive any other credits if the movie's

---

[205] POWDERMAKER, *supra* note 44, at 11.

[206] *Id.* at 154.

[207] Among 213 movies in 2020, 185 writers from 69 films received no credit. *See* Writers Guild of American West, *Screen Credits Referendum 2021*, https://www.wga.org/uploadedfiles/the-guild/elections/screen_credits_explainer.pdf [https://perma.cc/6ZNR-LQKZ]; from 1949 to 2019, about two-thirds of movies credit one writer, 16% credit two, and 6% credit three. *See* Stephen Follows, *How many screenwriters does it take to write a movie?* STEPHEN FOLLOWS (Feb. 24, 2020), https://stephenfollows.com/how-many-screenwriters-does-it-take-to-write-a-movie/ [https://perma.cc/8GW6-DHY4].

[208] WRITERS GUILD OF AMERICAN WEST, SCREEN CREDITS MANUAL, https://www.wga.org/contracts/credits/manuals/screen-credits-manual.

[209] *See id.*

[210] Story credit may not be shared by more than two writers. "Written by" and "screenplay by" credits can be shared with three writers or three writing teams in unusual cases and only after arbitration. Each writing team cannot exceed three writers unless a waiver from the Guild.

[211] Writers Guild of American West, *Members Vote Yes on Screen Credits Rule Change*, WGAW (Nov. 15, 2021), https://www.wga.org/members/membership-information/members-vote-yes-on-screen-credits-rule-change [https://perma.cc/G96Z-38TC].

credits are final after December 3, 2021.

Table 2-4 The Difference between Story and Screenplay[212]

STORY	SCREENPLAY
Idea and basic narrative	Dramatic construction
Outline indicating action	Original and different scenes
Outline indicating character development	Characterization or character relationships
Theme	Dialogue

2. Director

In the early days, the director was the same person as the producer and oversaw unifying a motion picture's creative elements. [213] Since Hollywood developed a hierarchical studio system in the early 1920s, the directors no longer determined the motion picture's essential ingredients, i.e., script, budget, cast, and shooting location. The executives and producers supervised all the motion picture's production except for empowering the directors with somewhat autonomy during the shooting stage. This was mainly for efficiency and unification reasons. The studio needed a powerful director to make instant and intricate decisions on set when the shooting was ongoing. No matter how much autonomy is authorized, the director was always asked to keep his boss's limit in mind, precisely time and money. Once the shooting was over, the executives and producers

---

[212] WRITERS GUILD OF AMERICAN WEST, *supra* note 208.
[213] Ronny Regev, *Hollywood Works: How Creativity Became Labor in the Studio System*, 17(3) ENTER. & SOC'Y 591, 595 (2016).

43

ultimately determined the "final cut."

By the late 1960s, the directors became more influential along with New Hollywood's emergence.[214] Several factors resulted in the transition. First, exhibitors began to play more independent motion pictures after escaping the major studios' vertical control due to the 1948 Paramount antitrust decision.[215] The prosperity of independent motion pictures awoke Hollywood to notice a niche market long ignored. Second, *Joseph Burstyn v. Wilson* (1952) affirmed that the First Amendment protected the motion pictures.[216] Subsequently, the 1968 new rating system replaced the Production Code's censorship and formally emancipated Hollywood's creativity.

The directors with unique personal film styles were called "auteurs."[217] They were also named the "film school generation" because many had film school backgrounds. Aiming to make films bear the "mark of directorial authorship,"[218] the studios gave the directors more autonomy except for reserving "the right to have the first chance to produce or refuse a new project."[219] Reputable directors, such as Francis Ford Coppola, Steven Spielberg, Martin Scorsese, and John Milius, became new stars instead of the conventional genre or star actors.[220] Rather than saying that Hollywood respected these directors' creative autonomy, it was better to say that it loved the audiences and profits these directors attracted.

---

[214] DECHERNEY, *supra* note 121, at 85-10.
[215] United States v. Paramount Pictures, Inc., 334 U.S. 131 (1948).
[216] Joseph Burstyn, Inc. v. Wilson, 343 U.S. 495 (1952).
[217] DECHERNEY, *supra* note 121, at 95.
[218] *Id.* at 94.
[219] *Id.*
[220] *See id.* at 93.

Nonetheless, a director's work of art could be the studio's nightmare. Michael Cimino's notorious western "*Heaven's Gate*" (1980) was described as the biggest disaster in Hollywood though some considered it a masterpiece three decades later.[221] Cimino's five-Oscars dazzling achievement on "*The Deer Hunter*" (1978) lured the United Artists studio to entrust him with almost complete authority to make the "*Heaven's Gate*" (1980). Unfortunately, Cimino's perfectionism in detail caused terrible budget overruns and schedule delays. The final total production cost came to US$44 million, which was much more than the approved budget of US$11.6 million,[222] whereas the global gross was only US$3.5 million.[223] This failure nearly ruined Cimino's career and bankrupted the United Artists. [224] Many believed this tragedy terminated the New Hollywood Wave. [225] Henceforth, an era that "the studios took back control from the artists and box office trumped everything" was reinstated in Hollywood.[226]

In modern Hollywood, the brand franchise model has become the most reliable source of income.[227] A movie in a successful franchise, like *Harry Potter*, *Star Trek*, *Spiderman*, and *The Terminator*, can cover all the other wrong investments.[228] Because of

---

[221] Nicholas Barber, *Heaven's Gate: From Hollywood disaster to masterpiece*, BBC (Dec. 4, 2015), https://www.bbc.com/culture/article/20151120-heavens-gate-from-hollywood-disaster-to-masterpiece [https://perma.cc/33WF-T6Q4].

[222] *Id.*

[223] HEAVEN'S GATE, IMDB, https://www.imdb.com/title/tt0080855/ [https://perma.cc/83CC-HRZE].

[224] *See* Marlon Mosley, *Heaven's Gate: UA and Michael Cimino's demise*, MWMBLOG (Oct. 26, 2019), https://mwmblog.com/2019/10/26/heavens-gate-ua-and-michael-ciminos-demise/ [https://perma.cc/85PR-J8M9]; *see also* Robert W. Welkos, *From the Archives: 'Heaven's Gate': The film flop that reshaped Hollywood*, L.A. TIMES (June 12, 2004), https://www.latimes.com/entertainment/movies/la-et-mn-heavens-gate-flop-archive-20040612-snap-story.html [https://perma.cc/F279-CQ6G].

[225] Chris Heckmann, *What is New Hollywood? The Revolution of 1960s and '70s Hollywood*, STUDIO BINDER (May 17, 2020), https://www.studiobinder.com/blog/what-is-new-hollywood/ [https://perma.cc/FY47-5MXG].

[226] Welkos, *supra* note 224.

[227] *See* Garon, *supra* note 144, at 778.

[228] *Id.* at 777.

the high stakes, an unresolved creative difference for a motion picture in a franchise between directors and executives can lead to film reshoots and even director replacement. For example, *Rogue One: A Star Wars Story* (2016), directed by Gareth Edwards, was reshot because Disney executives wanted more than the first cut.[229] As another example, Phil Lord and Chris Miller were forced to leave the director seats of Disney's *Solo: A Star Wars Story* (2018) even though the shooting had finished three-quarters.[230] It was because the producer Kathleen Kennedy had a different creative vision of the film.[231] Not to mention that most of the Marvel movie directors are thought to have limited creative freedom because the Marvel movies' tones and visual styles are unified.[232] Probably what the major studios seek most, particularly for the movie franchises, is no longer the directors' talent to create unique visions but their ability to be directed.[233] The directors' styles would be appreciated only to fulfill the studios' heavy demand.

---

[229] Andrew Liptak, *Rogue One's reshoots show how Disney saved the first standalone Star Wars movie*, VERGE (Jan 15, 2017), https://www.theverge.com/2017/1/15/14268850/star-wars-rogue-one-reshoots-disney-gareth-edwards-tony-gilroy [https://perma.cc/53TP-9L59].

[230] Borys Kit & Kim Masters, *Ron Howard Steps into Direct Han Solo Movie*, HOLLYWOOD REPORTER (June 22, 2017), https://www.hollywoodreporter.com/movies/movie-news/star-wars-han-solo-movie-ron-howard-steps-direct-1015674/ [https://perma.cc/9LWA-YSFP].

[231] Brent Lang, *'Star Wars' Han Solo Spinoff: Lord & Miller Fired after Clashing with Kathleen Kennedy*, VARIETY (June 20, 2017), https://variety.com/2017/film/news/star-wars-han-solo-kathleen-kennedy-director-fired-1202473919/ [https://perma.cc/JUZ5-YL4Q].

[232] Jillian Unrau, *The One Thing Marvel Should Change to Succeed in the Future*, GAMERANT (Dec. 14, 2021), https://gamerant.com/marvel-dirctor-styles-change-succeed-future/ [https://perma.cc/CE8K-TSJD].

[233] Kwame Opam, *Does Disney want its directors to have creative freedom?*, VERGE (June 23, 2017), https://www.theverge.com/2017/6/23/15861162/disney-marvel-star-wars-lucasfilm-han-solo-director-creative-freedom [https://perma.cc/HVG9-AFTQ].

III. AUDIENCE-ORIENTED PRACTICE

A. CONVENTIONAL INDUSTRY-DEVELOPED METHODS

"We give the public what it wants" is Hollywood's golden rule.[234] The underlying theory is that audiences will pay for their favorite motion pictures to bring large profits to the studios. Since the 1910s, Hollywood had begun to collect audience preference with industry-created methods.[235] In addition to the studio executive's intuition, Hollywood has developed the following methods to measure audience preference.

1. Box Office Figures

Whether a motion picture is booming can be evaluated through box office figures. But it cannot explain why the motion picture could succeed.[236] Too many factors, e.g., quality, title, genre, theme, casting, advertising, publicity, promotion, competitive substitute, timing, and even the weather, can affect ticket sales. The box office figures are unable to isolate each factor's effects.

2. Sneak Previews

A good way to observe the test audiences' spontaneous reactions to a motion picture before its official release is the sneak preview.[237] Preview cards are given to the test

---

[234] POWDERMAKER, *supra* note 44, at 42.
[235] Susan Ohmer, *Measuring Desire: George Gallup and Audience Research in Hollywood*, 43(½) Film & Video J. 3 (1991).
[236] LEO A. HANDEL, HOLLYWOOD LOOKS AT ITS AUDIENCE: A REPORT OF FILM AUDIENCE RESEARCH 6-8 (1950).
[237] *Id.* at 8-9.

audiences to collect their comments about their favorite degree of the motion picture, characters, and scenes. (see **Appendix A**[238] and **Appendix B**[239]). The endings and plots may be changed according to the previewers' responses.[240] But the test audiences are not strictly controlled to represent the cross-section of the moviegoing public, and not all the test audiences would share honest views.

3. Fan Mails

The numbers and quantity change of fan mails received by individual stars are used to evaluate the stars' popularity.[241] But it is only a rough index and has the same problem as the sneak previews in representativeness.

4. Exhibitor's Opinions

The exhibitors sit between the motion picture studios and ticket buyers.[242] A New York film exhibitor described that he could "see, hear, and feel" the audience's reaction by watching the audience's faces.[243] Their opinions were valuable in evaluating the film's market acceptance. But this method can be limited by the exhibitors' localities, biases, and

---

[238] Appendix A is a questionnaire given to the audience at a sneak preview of *Return of the Jedi* at the Northpoint Theater in San Francisco on April 9, 1983. *See* Paul Duncan, FACEBOOK, (Apr. 9, 2021, 12:24 PM), https://www.facebook.com/kershed/photos/a.2625273730832567/4539457882747466/?locale=hi_IN (last visited Aug. 15, 2023).

[239] Appendix B is writer John August's questionnaire for his film, *The Nines* (2007). Download from https://johnaugust.com/downloads_ripley/nines_questions.pdf. *See also* Yuri Baranovsky, *Six Tips for a Successful Test Screening*, FRAME.IO (July 6, 2017), https://blog.frame.io/2017/07/06/successful-test-screening-tips/ [https://perma.cc/E742-GVXX].

[240] Rudie Obias, *11 Movies That Changed Because of Test Audiences*, MENTAL FLOSS (Apr. 4, 2016), https://www.mentalfloss.com/article/64467/11-movies-changed-because-test-audiences [https://perma.cc/PB8J-T9TE].

[241] HANDEL, *supra* note 236, at 10.

[242] *Id.*

[243] OHMER, *supra* note 190, at 1.

the inherent conflict of their interest with the studios in bargaining the trade-in prices.[244]

5. Theatrical Tryouts

This method's basic assumption is that a famous play on the Broadway stage could also become a hit film on the big screen. Motion picture studios might buy film rights from the existing famous stage plays or, in reverse, make their intended films into stage shows first to try out the audiences' reactions.[245] Nonetheless, the audiences between plays and motion picture may be quite different. Besides, the motion picture has a wider distribution than the plays.

B. SCIENTIFIC AUDIENCE RESEARCH FOR GENERAL TREND

Since the 1940s, scientific film audience research institutions have emerged.[246] The Audience Research Institute (ARI), formed by George Gallup in 1940, worked only for RKO.[247] With RKO's permission, it further conducted researches for Selznick International Pictures[248] and Disney.[249] The Motion Pictures Research Bureau (MPRB), headed by Leo A. Handel since 1942, worked exclusively for MGM.[250] Instead of relying on intuition, hearsay, guesswork, and personal impressions, the polling institutions conducted film audience research with modern statistical and psychological tools.[251] They

---

[244] HANDEL, *supra* note 236, at 11.
[245] *Id.*
[246] Leo A. Handel, *Hollywood Market Research*, 7(3) FILM RADIO & TELEVISION Q. 304, 306 (1953).
[247] OHMER, *supra* note 190, at 127.
[248] *Id.* at 184-185.
[249] *Id.* at 198-200.
[250] HANDEL, *supra* note 236, at 4.
[251] Handel, *supra* note 246, at 306.

used interviews, questionnaires, IBM cards, [252] and machines to collect and analyze moviegoers' opinions. The number of respondents in most ARI surveys ranged from 1,000 to 5,000 people[253] and were strictly controlled by sex, age, and socioeconomic status to mirror the cross-section of actual audiences.[254] Consequently, specific audience groups can be targeted more precisely.[255]

This kind of research by outsiders was only welcomed by some. The executives were concerned that the study might imperil their authority.[256] The artists were worried that the survey might restrict their autonomy.[257] Nonetheless, the benefits of scientific audience research eventually let Hollywood cannot bear to let it go. Several factors prompted Hollywood to embrace scientific audience research in the 1940s.[258] First of all, Gallup's precise prediction of the presidential election results persuaded Hollywood that the same approach was also workable in motion pictures. Second, losing overseas markets during and after World War II forced Hollywood to care more about the domestic audience's reactions. Third, increased production costs and the antitrust decrees of 1940, which limited block booking up to five features and banned blind bidding, made each film's merit and profitability more critical. Thus, the studio executives and independent producers needed advanced information to minimize risk and support their judgments. Fourth, the audience research result could be used for motion pictures advertisements and promotions. From

---

[252] "IBM card" is the earliest tool invented by IBM for data processing. It is a punched card with rectangular holes and 80 columns. The size is 7- 3/8 inches by 3- 1/4 inches. *See* THE IBM PUNCHED CARD, IBM100, https://www.ibm.com/ibm/history/ibm100/us/en/icons/punchcard/ [https://perma.cc/XX8S-9WST].
[253] OHMER, *supra* note 190, at 128.
[254] HANDEL, *supra* note 236, at 40.
[255] OHMER, *supra* note 190, at 2.
[256] HANDEL, *supra* note 236, at 4.
[257] Ohmer, *supra* note 235, at 4.
[258] OHMER, *supra* note 190, at 216-217.

1946 to 1951, many major studios and independent producers had used the ARI's survey service.[259]

1. Audience Demographics Analysis

Knowing the extent and nature of the film audience is essential for making motion pictures that draw the most audiences.[260] The researchers surveyed the size of the potential and actual audience, the frequency of attendance, the number of admissions, gender, age, income, education, region, and moviegoing alone or with others.[261] Some findings, e.g., the nearly equal motion picture attendance between men and women,[262] teenagers buying the most tickets,[263] and college-educated people attending motion pictures more often than lower-educated,[264] overturned previous misconceptions.

2. Audience Preference for Content

Content analysis of a motion picture is the basis for studying its effect on audiences. It is the most time-consuming and challenging part.[265] MPRB developed a content analysis chart coding many essential aspects of a motion picture, including primary story type, secondary story elements (36 classifications), geographical designation, sociological designation, age, race, religion, marital status, sex of the lead players, characterization of several occupations, and so on.[266] Researchers read the script and watched the motion

---

[259] *Id.* at 217.
[260] *Id.* at 128.
[261] HANDEL, *supra* note 236, at 93-136.
[262] *Id.* at 99-100.
[263] OHMER, *supra* note 190, at 131-132.
[264] HANDEL, *supra* note 236, at 107.
[265] *Id.* at 172.
[266] *Id.* at 171-173.

picture repeatedly to record the information and punched the relevant information on the IBM cards for further study.

Classification of story types took a lot of work because of the possible overlap.[267] MPRB classified motion pictures into eighteen story types.

A. Musical Comedies	J. Historical Pictures, Biographies
B. Sophisticated Comedies	K. Mystery, Horror Pictures
C. Family Life Comedies	L. Western Pictures
D. Slapstick Comedies	M. Gangster, G-Men Pictures
E. Just Comedies	N. Musicals (Serious)
F. Love Stories	O. Socially Significant Picture
G. War Pictures	P. Child Star Picture
H. Serious Drama	Q. Fantasies
I. Adventure, Action Pictures	R. (Wild) Animal Pictures

In 1942, the MPRB surveyed two thousand people. It concluded that men's most liked story types were war, adventure, and action; women's were love stories, romance, and musical comedies. The most disliked story types of men were love stories, romance, and child stars; women were western, mystery, horror, and gangster. The study also revealed that age and income were relevant to the acceptance of serious drama but had fewer effects on other story types. Many studies conducted by the ARI found that what interested audiences was not the story type but the story itself.[268]

Some studies concerned the characters, subjects, plots, and scenes of motion pictures. For example, two ARI surveys of 1943 showed that Disney characters in order of favorites were Donald Duck, Thumper, Bambi, and Mickey Mouse. *Cinderella* was the

---

[267] *Id.* at 120-127.
[268] *Id.* at 117.

most favorite cartoon subject compared to *Peter Pan*, *Alice in Wonderland*, *Lady and the Tramp*, *Uncle Remus*, and *Hiawatha*.[269] Other ARI studies revealed that most moviegoers had no difficulties distinguishing movie plots from reality, and only a few moviegoers objected to movie scenes of women drinking or smoking.[270] Also, the ARI study showed that most moviegoers sought escapist entertainment from motion pictures, negating those suggestions that Hollywood motion pictures should address more current issues.[271]

3. Audience Preference for Players

Various studies ascertained the cast as the most critical component of a motion picture in drawing power relative to the story types in the second position and the title in the third position.[272] General star popularity ratings were determined by the respondents' star favorite degrees ranging from 100 to zero.[273] The change of a star's popularity index following movie releases indicated the moviegoers' responses to the star's roles in the movie.[274] The data was used to select the best possible movie roles for the stars in the future.

Besides, each star could have different popular degrees among various audience groups according to the audience's sex, income, and attendance frequency.[275] For example, star X may be famous among the elders, while star Y may be viral among youngsters. Accordingly, movie producers can combine the two stars in a motion picture to maximize the volume of potential audiences. Also, studies found that most audiences preferred the

---

[269] OHMER, *supra* note 190, at 205.
[270] *Id.* at 138-139.
[271] *Id.* at 139-141.
[272] HANDEL, *supra* note 236, at 118.
[273] *Id.* at 138.
[274] *Id.* at 139.
[275] *Id.*

stars of their sexes and believed the stars of their sexes had better acting abilities.[276] To obtain more information on specific players, further studies were conducted to probe why a particular player's popularity rating declined or whether the moviegoing public wanted to see a fledgling player in coming motion picture.[277]

C. SCIENTIFIC AUDIENCE RESEARCH FOR INDIVIDUAL FILMS

In 1939, Gallup conducted a nationwide survey for *Gone with the Wind* for the first time.[278] According to the study, fourteen million people had read the novel. Sixty-five percent of the respondents would want to see the film. Scarlett was the character most people were interested in. Thirty-five percents were satisfied with Vivien Leigh playing Scarlett, sixteen percents were dissatisfied with the casting, and twenty percents were undecided. The movie's producer David Selznick repeatedly used the result to defend his casting decision, make report to investors, and promote the film.[279]

For the studios, the audiences' interest in a specific motion picture project may be more important than their general preferences.[280] Scientific and systematic audience research can be applied in a particular motion picture's different filmmaking phase to help the studios make informed decisions for a specific project.

1. Pre-production Phase

a. Story Test

---

[276] *Id.* at 144.
[277] *Id.* at 140.
[278] OHMER, *supra* note 190, at 172.
[279] *Id.* at 174.
[280] *Id.* at 126.

The goal of the story test was to find out the marketability of a potential motion picture project. If a story were based on an existing novel or play, the research would focus on how familiar the moviegoers were with the story, whether they wanted to see a motion picture made from the story, and which part of the story they liked or disliked.[281] In the case of an original story, the interviewee would read its brief outline and listen to the interviewer's reading of it. Then, the interviewee would be asked whether they wanted to see the motion picture (see a sample below).[282] The ARI concluded that a 60-word synopsis was enough to test the interviewee's reactions to a motion picture project after experimenting with summaries ranging from one or two sentences to 5,000 words.[283]

Story test sample

Brief Story Outline

This picture will show in the form of an exciting story how the release of atomic energy will affect you and your family – whether this most important invention of all times will be the beginning of a better world, or the end of everything. The picture will have an all-star cast and its title will be:
 "THE BEGINNING OR THE END"

Questions

1. How interesting would the picture be to you? ☐ Very interesting   ☐ Fairly interesting   ☐ Slightly interesting   ☐ Not at all interesting

2. Do you think you will actually go and see this picture? ☐ Yes   ☐ No   ☐ Undecided

---

(Source: Handel, Leo A. Hollywood, Looks at Its Audience: A Report of Film Audience Research)

The difficulty of this method was that a different narrative of the same film plan

---

[281] HANDEL, *supra* note 236, at 24.
[282] *Id.* at 24.
[283] OHMER, *supra* note 190, at 135.

might lead to a different conclusion.[284] For example, in the film plan about Frederic Chopin's life (*A Song to Remember*, 1945), the respondents were more interested in a version stressing his love for George Sand than other versions.[285] Other criticisms included: audiences usually did not know whether they would like a movie before seeing it, the audience's taste was changeable, and the plot was not the only factor determining the popularity.[286] Even Gallup himself admitted that "the impact of an inventive script, witty dialogue, or imaginative direction" was hardly evaluated by a pre-production test.[287]

b. Casting Test

The casting test aimed to find the best possible cast for a motion picture project.[288] After learning the story and the characters, the interviewees were given a star list and were asked to select the leading star and co-star they liked best to play in the motion picture.

2. During the Production Phase

a. Title Test

The title test can be conducted at any time point before publicity.[289] It was used to test the drawing power of the tentative title. The respondents may be asked if they wanted to see a motion picture with a specific title or selected a motion picture they wanted to see among several titles. For comparison, some controlled titles with average drawing power may be added to the list.

---

[284] HANDEL, *supra* note 236, at 26.
[285] OHMER, *supra* note 190, at 220.
[286] POWDERMAKER, *supra* note 44, at 42-44.
[287] OHMER, *supra* note 190, at 220.
[288] HANDEL, *supra* note 236, at 32-34.
[289] *Id.* at 35-44.

b. Profile Chart

This test helped the filmmakers to know the moviegoers' likes and dislikes of the film before the final release.[290] When the rough cuts were screened, viewers' reactions were recorded synchronously with a viewer-controlled manual device. The viewers pressed the "like" button if they liked a scene and pressed the "dislike" button if they disliked a scene. The two buttons had distinct colors, and each hand of the viewer controlled one button. If no button was pressed, it meant indifference. Similar devices applied identical concepts with somewhat improvement. For example, the Televoting Machine used by ARI had a luminous dial switchable to five positions from "very dull, dull, neutral, like to like very much."[291] It combined the total votes only but did not record individual responses.

This test can also be conducted without any device. By numbering each film scene in advance, the viewers recorded their preferences for each numbered scene by themselves. In any case, a profile chart would be produced to show the viewers' reactions to each part of the film. To reduce deviations and ensure no operation errors, the viewers had to be familiar with the testing process, fill out a questionnaire, and receive interviews after testing. The results were applied in editing the rough cuts.[292] The disliked parts would be deleted or shortened, and the confusing parts would be clarified.

3. Post-Production Phase

a. Analytical Preview

This was an improved version of conventional Hollywood sneak previews.[293] The

---

[290] *Id.* at 44-60.
[291] Ohmer, *supra* note 235, at 18.
[292] OHMER, *supra* note 190, at 222.
[293] HANDEL, *supra* note 236, at 61-67.

test can be conducted through personal interviews or mail questionnaires after screening the motion picture. The preview audiences gave opinions about the picture, actors, scenes, and episodes. One of the critical questions was to rate their favorite degrees to the motion picture. New equipment was designed to monitor the audience's physiological responses to ascertain their honest reactions. The result can be applied in editing the final cut or discovering the liked features of the motion picture for advertising and publicity.

b. Advertising and Publicity Study

This study was used to estimate how many potential audiences were exposed to the marketing efforts and how many of them were likely to see the motion picture.[294] It measured the breadth and effectiveness of various media, trailers, and opinion leaders in promoting the motion picture. The marketing campaign can be adjusted according to the responses.

D. PROBLEMS OF SCIENTIFIC AUDIENCE RESEARCH

1. Too Little or Too Late Information

The scientific audience research can do very little in the preproduction stage as the information obtained after shooting was often considered "post mortems."[295] ARI only used sixty-word synopsis and a few options for the story test in consideration of practicality and efficiency. [296] Thus, it can only provide "a general indication of audience preferences"[297] rather than complete information to guide the scriptwriting and shooting.[298]

---

[294] *Id.* at 74-90.
[295] Handel, *supra* note 246, at 310.
[296] OHMER, *supra* note 190, at 135.
[297] Ohmer, *supra* note 235, at 15.
[298] *See id.* at 16.

By contrast, the profile chart and preview conducted after the shooting did better in offering details. However, reshooting was usually not a practical option. As a result, the study was often limited to the purposes of editing and promoting.[299]

2. Not Good at Creativity or Inhibit Innovations

Scientific audience research was also not good at creativity and even inhibited innovations. First, there shall be some ideas ready for testing. The ideas might come from the studio's developing projects, magazine readership, best-selling novels, or books.[300] The methodology was not designed to create new things but to uncover whether the off-the-shelf ideas would interest the audiences, and which one would interest more audiences.[301]

Even if new ideas were submitted for testing, they hardly prevailed. ARI's studies repeatedly dissented from RKO's new ideas and recommended the already popular subjects in other media.[302] In the 1940s, ARI surveyed the audience's reaction to Disney's innovative animated films, *Fantasias* (1940) and *The Reluctant Dragon* (1941).[303] The studies showed that moviegoers were not interested in new-form or new-subject motion pictures but preferred traditional Disney motion pictures with familiar characters and stories. However, it was argued that audiences needed time to absorb innovative ideas before answering whether they would like them.[304] Under ARI's methodology, innovative ideas had little chance to defeat familiar styles.

---

[299] *See id.* at 18.
[300] OHMER, *supra* note 190, at 136.
[301] *See id.* at 135.
[302] *See id.* at 137.
[303] *See id.* at 202.
[304] *See id.* at 203.

3. Tensions Between Speed and Accuracy

Speed was another critical issue.[305] The nature of scientific audience research was time-consuming because it needed to collect data through nationwide interviews and questionnaires. Not to mention designing surveys and analyzing raw data. However, ARI may be only allowed 24 or 48 hours or a weekend to submit the results.[306] Under the frequent requests, the ARI survey's accuracy was inevitably settled with the pace of film production. Moreover, Gallup's failed prediction of the 1948 presidential election weakened his poll reputation.[307] Some surveys initiated by exhibitors in 1949 and 1951 also challenged the ARI survey's accuracy. Gallup finally left Hollywood in the 1950s.

IV. CONCLUSION

Unlike a novel that a writer could create with a pen or a piece of music that a musician could play with an instrument, a feature film or a television series cannot be made only by an individual with his/her talents. Because filmmaking is so expensive, from the historical perspective, no matter whether motion pictures are technological products, the art of works, commercial goods, or all three, there would be no motion pictures if there were no business. Conversely, if there is business, the motion pictures will always be made even though their artistic values are slight.

As the motion picture industry needs profits to survive, businesspeople are inevitably above the artists and need information to make investment decisions,

---

[305] HANDEL, *supra* note 236, at 12.
[306] OHMER, *supra* note 190, at 22.
[307] Gallup Organization predicted Dewey (50%) would win 5% over Truman (45%) for the 1948 Presidential Election, but an almost exactly reversed election result came out. *See* Everett C. Ladd, *The Trials of Election Polling: Election Poll—1948 and Today*, https://www.trumanlibrary.gov/sites/default/files/1948Campaign_Trials.pdf [https://perma.cc/4EJT-VZGQ].

particularly when the amount involved is significant. Consequently, Hollywood had tried

to import scientific audience research to assist filmmaking since the 1940s, even though

the research's quantity, quality, and speed were limited. Since the motion picture industry

is familiar with adopting innovative technology, considers business superior to art, and has

extensive audience data-collecting practices, I argue that new data-driven trends would

impact motion pictures more than other types of work. The following chapter will introduce

the data-driven trend powered by big data and AI and see how it reshapes the filmmaking

process and content making.

CHAPTER THREE: TOWARDS DATA-DRIVEN FILMMAKING

Continuing with the final part of the former chapter introducing Hollywood's traditional methods to detect audience preferences and use that for making motion pictures, this chapter will explore how new data technology changes modern filmmaking. I will look at where the data comes from, what the data comprises, and what techniques are used to process big data. After that, I will visit some innovative film analytic platforms and content creation models and see what data technology can do to support filmmaking. Finally, I will study some practices in the real world. I will start with the Netflix and see how it transforms from a small content distributor to a powerful original content producer by leveraging the power of data. What do the Hollywood majors do to catch up with data-driven trend? In addition, I will visit some innovative attempts to combine data with AI to generate content and the initial application of data-driven creativity in screen education.

I. OVERVIEW OF DATA TECHNOLOGY AND AI

In a digital world, data is generated by every device and every activity at every moment and everywhere.[308] Data expands enormously, quickly, and diversely such that it is far beyond the handling capacity of human brains or traditional computer systems.[309] However, thanks to the progress of data engineering in the past decades, the limitations of building and maintaining big data systems have been overcome. Data scientists can extract meaningful insights from big data systems with powerful computer models to support business-critical decisions.[310] This is neither to say data scientists can take on the positions

---

[308] *See* LILLIAN PIERSON, DATA SCIENCE FOR DUMMIES 7 (2d ed. 2017).

[309] *See id.* at 18-19.

[310] *See id.* at 7-9.

of decision-makers[311] nor that computers can replace all human judgments.[312] Conversely, data scientists' most important duty is to provide decision-makers the data insights in understandable and effective ways.[313] Interpreting the data insights, making decisions, and taking actions are still humans' jobs.[314]

The scope of data science includes descriptive analytics (what happened), diagnostic analytics (why), predictive analytics (what is next), and prescriptive analytics (how to react).[315] Data scientists often have expertise in the subject matter fields or collaborate with subject matter experts to perform data tasks.[316] Businesses implementing "Data-Driven Development" (DDD) enjoy three benefits:[317] (1) making better-informed decisions based on data prediction; (2) streamlining resources on tasks that machines can perform automatically based on historical data; and (3) adapting to new environments continuously and quickly based on the latest data feedback. Obviously, such businesses are in a better competitive position than those failing to catch up with the data application. The motion picture industry is no exception.

A. DATA SOURCES

Volume, velocity, and variety are the most key features of big data, in addition to veracity and value.[318] Big data may come from internal and external channels of

---

[311] *See id.* at 24.
[312] FIELD CADY, DATA SCIENCE: THE EXECUTIVE SUMMARY: A TECHNICAL BOOK FOR NON-TECHNICAL PROFESSIONALS 3 (2021).
[313] Pierson, *supra* note 308, at 13.
[314] Cady, *supra* note 312, at 3.
[315] Pierson, *supra* note 308, at 35.
[316] *Id.* at 23.
[317] Cady, *supra* note 312, at 4.
[318] DOMINIQUE HAUGHTON ET AL., MOVIE ANALYTICS: A HOLLYWOOD INTRODUCTION TO BIG DATA 3 (2015).

enterprises. [319] Typical internal data is derived from business operations such as "management data, customer service data, sales and marketing data, operation data, and employee performance data."[320] Some data, such as "SCADA data, machine data, or sensor data,"[321] is generated by machines automatically. A type of fast-growing data is social data, e.g., "emails, instant messaging, and social networks such as Twitter, Facebook, LinkedIn, Pinterest, and Instagram."[322] Some businesses subscribe to outside data collection services, such as sale records of various distributors across different countries and regions.[323] There are also many external data sources like blog posts, websites, media, and public data.

Data may exist in structured, unstructured, or semi-structured formats.[324] Audio, video, image, PDF files, and social media data are typically unstructured data.[325] This data cannot be handled by traditional relational database management systems (RDBMS), as which can only manage structured data in rows and columns. [326] Various structured datasets are available online. For example, IMDB Reviews (5,000 movie reviews), IMDB Film Review dataset (50,000 reviews), MovieLens dataset (25 million ratings over 62,000 movies), OMDB API (over 280,000 posters), Film dataset from UCI (over 10,000 film information), Cornell Film Review data (over 220,000 conversations between movie characters),[327] and the Video Genome Project (over 8 million video records).[328] The Video

---

[319] *See* Pierson, *supra* note 308, at 20, 41.
[320] *Id.* at 42.
[321] *Id.*
[322] *Id.*
[323] Stephanie Prange, *Big Data Revolution*, MEDIAPLAYNEWS (Sept. 3, 2018), https://www.mediaplaynews. com/big-data-revolution/ [https://perma.cc/ELY9-TJVX].
[324] *See* Pierson, *supra* note 308, at 8.
[325] *See id.* at 42.
[326] *See id.* at 19, 29.
[327] 13 BEST MOVIE DATA SETS FOR MACHINE LEARNING PROJECTS, IMERIT (July 21, 2021), https://imerit. net/blog/13-best-movie-data-sets-for-machine-learning-projects-all-pbm/ [https://perma.cc/7BSB-AHVR].
[328] THE VIDEO GENOME PROJECT, https://thevideogenomeproject.com/ [https://perma.cc/M9AP-BKE4].

Genome Project (The VGP) aggregates video metadata, analyzes the contextual attributes to determine "genes" of each video, and explores the hidden relationships among videos.[329] In 2016, Hulu acquired the VGP to enhance its recommendation system.

B. DATA TECHNOLOGY

Data engineering and data science are two important branches in the big data field. The roles of data engineers and data scientists are different. Data engineers design, build, and maintain big data platforms, whereas data scientists analyze data on top of the data infrastructures.[330] Therefore, they demand different tools to perform their tasks.

1. Data Engineering

Special computer systems and software are needed to process and store massive amounts of data. Hadoop is one of the popular open-source distributed computing frameworks.[331] It is available on Amazon EMR,[332] Google's Dataproc,[333] and Microsoft's Azure HDInsight.[334] Hadoop ecosystem comprises Hadoop distributed file system (HDFS), MapReduce, Spark, and YARN.[335] Instead of running on one large computer,[336] these programs operate collaboratively on groups or clusters of connected and scalable commodity computers.[337] Servers in a group or cluster of computers are called "nodes."[338]

---

[329] Press Release, Hulu, Hulu acquires the Video Genome Project Technology (Nov. 15, 2016), https://press.hulu.com/hulu-acquires-the-video-genome-project-technology/ [https://perma.cc/Y4LE-YQ6C].

[330] *See* Pierson, *supra* note 308, at 19, 22.

[331] *See id.* at 28.

[332] APACHE HADOOP ON AMAZON EMR, AWS, https://aws.amazon.com/emr/features/hadoop/ [https://perma.cc/65QR-23BV].

[333] DATAPROC, GOOGLE CLOUD, https://cloud.google.com/dataproc [https://perma.cc/4B8A-UBLX].

[334] MANAGE YOUR BIG DATA NEEDS IN AN OPEN-SOURCE PLATFORM, AZURE, https://azure.microsoft.com/en-us/products/hdinsight/#overview [https://perma.cc/RG4R-HYCG].

[335] *See* Pierson, *supra* note 308, at 24.

[336] Apache Hadoop on Amazon EMR, *supra* note 332.

[337] *See* Pierson, *supra* note 308, at 27, 28.

[338] *See id.* at 27.

Thousands of nodes in a cluster are managed by YARN to assign them functions and schedule tasks.[339] Data engineers use MapReduce or Spark to process and transform big data into small datasets.[340] The main difference between MapReduce and Spark is that the former processes big data in batches, whereas the latter processes big data in real time.[341] The processed datasets and backups are stored by blocks on the nodes of HDFS.[342]

An easier and quicker alternative approach for MapReduce is the massively parallel processing (MPP) platform.[343] However, MPP is more expensive, as it runs on custom servers, not commodity servers. Other popular solutions are NoSQL databases, e.g., Apache Cassandra and MongoDB.[344] Unlike RDBMS, which can only handle structured data, NoSQL databases can handle unstructured and semi-structured data.

2. Data Science

A complete data science process includes (1) business understanding and data understanding, (2) preparing data, (3) building a model using algorithms, (4) applying model and performance evaluation, (5) deployment, and (6) knowledge and actions.[345] To derive insights from data, enough correct raw data must be extracted and converted into workable formats for analytics tools.[346] The processes and procedures are called data wrangling, which includes data extraction, data cleaning and transformation, and data quality control. Ideal data scientists not only can query relevant data with Structured Query

---

[339] *See id.* at 28.
[340] *See id.* at 24-27.
[341] *See id.* at 7-9.
[342] *See id.* at 27.
[343] *See id.* at 29.
[344] *See id.* at 30.
[345] Vijay Kotu & Bala Deshpande, Data Science: Concepts and Practice 21 (2d ed. 2019).
[346] *See* Pierson, *supra* note 308, at 36.

Language (SQL), i.e., SQLite, MySQL, and PostgreSQL, from structured databases,[347] but also can write scripts in computer programming languages, such as Python and R, to instruct the computer to process and analyze data.[348]

a. Machine Learning

To discover trends or patterns behind big data, machine learning is widely used.[349] Various statistics and mathematics knowledge, such as "probability, correlation analysis, dimensionality reduction, decision modeling, regression analysis, outlier detection, and time series analysis"[350] are used to design machine learning models. Typically, machine learning can be divided into supervised, unsupervised, and semi-supervised styles:[351]

(1) Supervised Learning

Learning with supervised algorithms can predict incoming unlabeled data based on the features learned from training labeled data. For instance, the Naïve Bayes algorithm can classify countless online motion picture reviews into positive or negative by the defined features.[352]

(2) Unsupervised Learning

Learning with unsupervised algorithms can categorize unlabeled data into groups or clusters based on the data's hidden similar patterns. How many groups will be sorted is unknown before the task is completed.[353] For example, the K-means algorithm can divide

---

[347] *See id.* at 43.
[348] *See id.* at 12.
[349] *See id.* at 58-59.
[350] *See id.* at 61.
[351] *See id.* at 51-54.
[352] *See* Atif Khan et al., *Movie Review Summarization Using Supervised Learning and Graph-Based Ranking Algorithm*, 2020 *Computational Intelligence and Neuroscience* (2020).
[353] *See* Pierson, *supra* note 308, at 95.

audiences into several groups based on their motion picture rating histories. The results can

be used to recommend content to audiences in the same groups.[354]

(3) Semi-supervised Learning

Semi-supervised learning (or reinforcement learning) models can improve outputs

continuously to earn maximum rewards. For example, SARSA and Q-Learning algorithms

are two famous algorithms applied to create content recommender systems based on user

feedback and behaviors.[355] A single recommender system may combine various algorithms

depending on the required tasks, data availability, and its properties.[356]

b. Deep Learning

Deep learning is a subset of machine learning and draws more attention. It is

structured with an input layer, two or more hidden (middle) layers, and an output layer.[357]

A model with more hidden layers can implement more complicated tasks.[358] A computer

vision application can have 150 more hidden layers. Learning patterns from unlabeled and

unstructured data is deep learning's primary strength. [359] Important deep learning

applications are text mining, [360] image processing, and natural language processing

---

[354] M. Tim Jones, *Unsupervised learning for data classification*, IBM (Dec. 4, 2017). https://developer.
ibm.com/articles/cc-unsupervised-learning-data-classification/#:~:text=In%20unsupervised%20learning
%2C%20an%20algorithm,for%20tasks%20like%20anomaly%20detection [https://perma.cc/WS68-
VU2M].

[355] Ehtsham Elahi et al., *Reinforcement Learning for Budget Constrained Recommendations*, NETFLIX
TECHBLOG (Aug. 15, 2022), https://netflixtechblog.com/reinforcement-learning-for-budget-constrained-
recommendations-6cbc5263a32a [https://perma.cc/GXU2-CC4Y].

[356] Harald Steck et al., *Deep Learning for Recommender Systems: A Netflix Case Study*, 42(3) AI MAG. 7, 9
(2021).

[357] *See* Jalal Rezaeenour et al., *Systematic review of content analysis algorithms based on deep neural
networks*, 82 MULTIMEDIA TOOLS & APPLICATIONS 17879, 17883 (2023).

[358] *See* Pierson, *supra* note 308, at 58.

[359] *Id.*

[360] Text mining means "extract[ing] meaningful and valuable information from unstructured textual data
and helps researchers achieve their goals." *See* Rezaeenour et al., *supra* note 357, at 17882.

(NLP).[361] The main deep learning methods are Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM). CNN is superior in processing images and videos, while RNN and LSTM are ideal in processing time series, text, and audio.[362]

C. ARTIFICIAL INTELLIGENCE

Artificial Intelligence (AI) means "built-systems that mimic human behaviors by making insightful decisions that are derived from artificial neural network model outputs."[363] The so-called neural network model is a subset of machine learning, and its advanced form is deep learning. In other words, deep learning is a subfield of machine learning, and machine learning is a subfield of AI.[364] Today, people use the word AI more often than Big Data to describe data science.[365] Google is considered "the world's biggest AI system" because it has the largest "data, data, data."[366] The more data the machines are fed, the better the machines will perform.

Though modern AI is excellent at solving problems based on insights extracted from data, it lacks a conceptual understanding of the underlying problems and common sense.[367] At present, AI's application is still limited to specific fields, far from having a real mind for general purposes.[368] Some filmmakers have attempted to use AI to make motion

---

[361] NLP is "concerned with giving computers the ability to understand the text and spoken words in much the same way human beings can." *See id.*

[362] *See id.* at 17879.

[363] *See* Pierson, *supra* note 308, at 113.

[364] WHAT IS ARTIFICIAL INTELLIGENCE (AI)?, IBM, https://www.ibm.com/cloud/learn/what-is-artificial-intelligence [https://perma.cc/427Y-QK3H].

[365] Cady, *supra* note 312, at 161.

[366] Somers, *supra* note 87.

[367] Cady, *supra* note 312, at 165.

[368] *Id.* at 162.

picture. For example, Tommy Pallotta and Femke Wolting created a robot called CameraBot to direct an interview with them.[369] The result was not satisfying because the interview felt like an interrogation, and the questions lacked connection. I will introduce more AI uses in Part IV.

II. MOTION PICTURE ANALYTICS PLATFORMS

In light of big data trends, several companies, Cinelytic, ScriptBook, LargoAi, etc., offer film analytics and predictive services for filmmakers. They combine data analytics with AI to provide evidence-based perspectives based on vast historical data.[370] These cloud-based platforms advertise quick and convenient data predictive services to the movie industry. Generally, their services include script diagnosis, revenue predictions, talent suggestions, and marketing strategies. I will introduce some big platforms in the following paragraphs.

A. CINELYTIC

Cinelytic is an American company. Its database covers almost 100,000 films and 550,000 more talent profiles.[371] It also tracks P2P downloads worldwide and sharing transactions. Cinelytic's AI-powered platform and machine-learning algorithms offer real-time revenue forecasting, rapid film comping, distribution analysis, talent comparison, casting analysis, and financial planning. But creative sides, such as suggesting how to

---

[369] Nadide Gizem Akgulgil Mutlu, *The future of film-making: Data-driven movie-making techniques*, 10(2) GLOBAL J. OF ARTS EDUC. 167, 169 (2020).

[370] HOW ARTIFICIAL INTELLIGENCE CAN HELP TO PITCH A MOVIE, LARGO, http://largofilms.ch/how-artificial-intelligence-can-help-to-pitch-a-movie/ [https://perma.cc/2TUQ-KU5K].

[371] FREQUENTLY ASKED QUESTIONS, CINELYTIC, https://www.cinelytic.com/faq/[https://perma.cc/9Z5V-M6SC].

change a script or how to film a scene, are not within Cinelytic's services.[372] Its predictive model is asserted to be 85% accurate, hence attracting many major Hollywood studios like Sony Pictures, STX Entertainment, and Warner Bros.[373]

By using Cinelytic platforms, users can focus on high-value creative activities and can easily compare different possibilities in real-time. Cinelytic stresses that it only assists users in making decisions rather than making decisions for them. Cinelytic founder Tobias Queisser said, "What it is good at is crunching numbers and breaking down huge data sets and showing patterns that would not be visible to humans. But for creative decision-making, you still need experience and gut instinct."[374]

B. SCRIPTBOOK

ScriptBook is a Belgium company offering scripting analysis and financial forecasts.[375] Its analytics model is trained on 6,500 existing scripts. It can diagnose a proposed script in five minutes, including the rating prediction, character analysis, target audience, and box office estimate.[376] To build an AI model, it applied machine learning, deep learning, and natural language processing to analyze 77 hits and 80 flops released

---

[372] *See* Chris O'Brien, *How Cinelytic is using AI to help Hollywood reboot for the streaming wars*, VENTUREBEAT (May 26, 2020), https://venturebeat.com/business/how-cinelytic-is-using-ai-to-help-hollywood-reboot-for-the-streaming-wars/ [https://perma.cc/V8R2-P2V3].

[373] *See* Adam Epstein, *No, Warner Bros. is not letting AI decide what movies it makes*, QUARTZ (Jan. 9, 2020), https://sg.news.yahoo.com/no-warner-bros-not-letting-152232931.html [https://perma.cc/7DTP-X2PA].

[374] Tatiana Siegel, *Warner Bros. Signs Deal for AI-Driven Film Management System*, HOLLYWOOD REPORTER (Jan. 8, 2020, 12:23 PM), https://www.hollywoodreporter.com/business/business-news/warner-bros-signs-deal-ai-driven-film-management-system-1268036/ [https://perma.cc/N99Z-QH7U].

[375] *See* Interview with Nadira Azermai, CEO of ScriptBook, *Writing new chapters in the film business: Storytelling meets AI*, EUROPEAN BUSINESS, https://www.european-business.com/interviews/scriptbook/writing-new-chapters-in-the-film-business-storytelling-meets-ai [https://perma.cc/QBV3-9TU8].

[376] *See* Peter Caranicas, *Artificial Intelligence Could One Day Determine Which Films Get Made*, VARIETY (July 5, 2018), https://variety.com/2018/artisans/news/artificial-intelligence-hollywood-1202865540/ [https://perma.cc/C7YG-3WNL].

from 2015 to 2017.[377] When it applied the model to test the same 157 films, 104 films were predicted to be hits (76 of them indeed were hits), whereas 53 films were predicted to be flops (52 of them indeed were flops). As for 50 films released in 2017 and 2018, ScriptBook had an 86 percent success rate in predicting which films were hits.[378]

ScriptBook data scientist Michel Ruelens said some big Hollywood studios had used their services quietly.[379] ScriptBooks CEO Nadira Azermai stated, "While a computer may not be able to capture why we find a certain movie enthralling, it can compare the script of said movie to a very large dataset of other scripts and draw parallels to story structures in past movies that have been associated with successes or failures."[380] Asked whether ScriptBook would harm creativity, she did not think so. Instead, she said, "ScriptBook's AI will just kick out movies that follow certain formulas. It is very good at picking out artistic movies that do well financially."[381]

C. LargoAi

LargoAi is a Swiss-based company. Its AI-assisted filmmaking tools help all sizes of motion picture producers to make decisions across the filmmaking process.[382] Its data analytics system comprises 7,845 lines of code.[383] Through analyzing video, audio, and text to study motion picture components and ingredients, it suggests script modification,

---

[377] *See* Interview with Nadira Azermai, *supra* note 375.
[378] *See* James Vincent, *Hollywood is quietly using AI to help decide which movies to make*, VERGE (May 28, 2019, 11:10 AM EDT), https://www.theverge.com/2019/5/28/18637135/hollywood-ai-film-decision-script-analysis-data-machine-learning [https://perma.cc/2LLK-4WBF].
[379] *See id.*
[380] *See* Interview with Nadira Azermai, *supra* note 375.
[381] Caranicas, *supra* note 376.
[382] *See* Aran Davies, *How to use Largo.ai: Part 1 – Introduction*, SOFY (2021), https://sofy.tv/blog/use-largo-ai-part-1-introduction/ [https://perma.cc/6PYR-2XLS].
[383] *See* Aran Davies, *Analytics and Movies: Predicting Movie Gross*, SOFY (2019), https://sofy.tv/blog/analytics-movies-predicting-movie-gross/ [https://perma.cc/J59Z-V87H].

recommends casting, and identifies the target audience. Some script suggestions that LargoAi can make include identifying if there are gender inequality issues (too few women's roles or screen time) and if a single subplot stays too long.[384] But it may be able to do more because its website says, "LargoAi and our data scientists can ascertain which part of the film/script needs to be altered in order to increase the market appeal of the film."[385] Like Cinelytic and ScriptBook, LargoAi can also predict audience reactions on a country basis[386] and movie revenue.[387] In 2019, LargoAi predicted an Italian movie, *Domani e Un Altro Giorno*, might earn between €1.6 million and €3.9 million. The actual result turned out to be €1.96 million. It also forecasted Sony Entertainment's feature, *Venom*, a gross of $201 million, and the actual gross was $213 million.

Rather than creating successful formulas for specific types of films, LargoAi focuses on finding a pattern that can apply to all films. LargoAi's Founder and CEO Sami Arp described such a pattern as the DNA of a story, just like the musical notes or sheets for music.[388] He said, "Our system does not tell the writers or directors what type of patterns they should create, but it rather shows the patterns of what they created. And these patterns then act like magnifiers that identify strong and weak points of a story."[389] He

---

[384] HOW AI IS BECOMING A SCRIPTWRITER'S BEST FRIEND, LARGO, http://largofilms.ch/how-ai-is-becoming-a-scriptwriters-best-friend/ [https://perma.cc/7RU3-LYQV].

[385] DATA-ASISTED MOVIEMAKING WITH LARGOAI, LARGO, http://largofilms.ch/data-driven-moviemaking/ [https://perma.cc/38DK-EXHW].

[386] Aran Davies, *Largo Films Enters Data-Driven Moviemaking*, SOFY (2019), https://sofy.tv/blog/largo-films-enters-data-driven-moviemaking/ [https://perma.cc/L2RH-MZA7].

[387] Davies, *supra* note 383.

[388] Fotis Georgiadis, *The Future Is Now: Sami Arp of Largo AI on How Their Technological Innovation Will Shake up the Tech Scene*, MEDUIUM (Apr. 27, 2022), https://medium.com/authority-magazine/the-future-is-now-sami-arp-of-largo-ai-on-how-their-technological-innovation-will-shake-up-the-c5fa3cd3a755 [https://perma.cc/DQ2E-G8CQ].

[389] *Id.*

considered LargoAi only "a tool to help in the decision-making process but does not replace human creativity or decision making."[390]

III. MOTION PICTURE CONTENT CREATION MODELS

Not only can data be used to predict a motion picture's success probability, but also it can be used to train machine learning models to generate new content. Almost all good motion picture start with a good script.[391] Would not it be amazing if the machine could produce scripts corresponding to audience preference? Before that, we must know if a machine can write like a human. We cannot expect a machine to write a script if it cannot even write like a human. This is the subject area that Natural Language Processing (NLP), which contains Natural Language Understanding (NLU) and Natural Language Generation (NLG), is concerned about.

The first step for the machine to process natural language is to split paragraphs or sentences into small units as a single word or a phrase (the process called tokenization)[392] and convert each token into a numerical format (the process called embedding).[393] A same word in human eyes is converted to different numbers depending on its location in a sentence and if it is capitalized. For example, in GPT-3, the sentence [My favorite color is red.] is converted to [3666, 4004, 3124, 318, 2266, 13], and [My favorite color is Red.] is converted to [3666, 4004, 3124, 318, 2297, 13], whereas [Red is my favorite color.] is

---

[390] *Id.*

[391] BERTRAND MOULLIER, RIGHTS, CAMERA, ACTION! INTELLECTUAL PROPERTY RIGHTS AND THE FILMMAKING PROCESS 18 (2022).

[392] Mehul Gupta, *Tokenization algorithms in Natural Language Processing (NLP)*, MEDIUM (Mar. 31, 2020), https://medium.com/data-science-in-your-pocket/tokenization-algorithms-in-natural-language-processing-nlp-1fceab8454af [https://perma.cc/PZ4M-KX2N].

[393] Cady, *supra* note 312, at 171.

converted to [7738, 318, 616, 4004, 3124, 13].[394] After training on massive amounts of text data, machines can identify the syntax and semantics of a sentence based on the pattern and relationship between words and phrases.[395] Continuous efforts are made to detect the sentiment and emotion within the text.[396]

Instead of processing word by word in sequence as previous Recurrent Neural Networks (RNN), new Transformer models, such as OpenAI's GPT (Generative Pre-Training Transformer) and Google's BERT (Bidirectional Encoder Representations from Transformers), can ingest all words in a sentence at once and retain better long text's contextual relationships.[397] Powered by NLP, machines can accept text input and generate human-like text output like translation, answering questions, or summarization.[398] The following paragraphs will introduce how the language models are applied in developing film scripts.

A. DISNEY RESEARCH

Backed by the Walt Disney Company, Disney Research is dedicated to leveraging technological innovations in the creative filmmaking process.[399] It developed three scriptwriter-assisting models in 2018. One model maintains the overall consistency of

---

[394] 100 tokens ~= 75 words. *See* Raf, *What are tokens and how to count them?*, OPENAI, https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them[https://perma.cc/3X6Z-TEWT].

[395] Eda Kavlakoglu, *NLP vs. NLU vs. NLG: the differences between three natural language processing concepts*, IBM BLOG (Nov. 12, 2020), https://www.ibm.com/blogs/watson/2020/11/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/ [https://perma.cc/K8P9-JK42].

[396] Pansy Nandwani & Rupali Verma, *A review on sentiment analysis and emotion detection from text*, 11(1):81 Soc. Network Analysis & Mining (2020).

[397] Demi Ajayi, *How BERT and GPT models change the game for NLP*, IBM BLOG (Dec. 3, 2020), https://www.ibm.com/blogs/watson/2020/12/how-bert-and-gpt-models-change-the-game-for-nlp/ [https://perma.cc/HUV3-T8RL].

[398] Kavlakoglu, *supra* note 395.

[399] ABOUT US, DISNEYRESEARCH STUDIOS, https://studios.disneyresearch.com/about-us/ [https://perma.cc/5EC5-MKAN].

characters and events across a long movie script, particularly when the story is part of a large and complex story world.[400] Another model named CARDINAL enables the writers to view their stories through various perspectives by converting the narrative into 2D and 3D visualizations.[401] The last model inspires writers, particularly when they encounter writer's block in the creative process.[402]

It is challenging for a writer to create a continuous story in a pre-existing large story world, such as *Star Wars*, or *Toy Story*.[403] This is because the networks and relationships between numerous characters are already complicated. The historical events bearing complex cause-and-effect relationships are countless. Moreover, the fictional world has its own rules or limitations that must be followed. Disney Research solves this problem by establishing three knowledge bases (story, character, and story world) for storing the relevant knowledge bytes. The knowledge byte is the smallest unit of information parsed from previous scripts, background information, and the writer's input with Natural Language Understanding. It contains the subjects, objects, actions, beliefs, desires, locations, time points, and relationships. Meanwhile, a Cross-Knowledge Base Reasoning system is built to gather information from the knowledge bases and answer the writer's inquiries by applying Logical Reasoning methods. The writer can interact with the model in the human language during the writing process. The model continuously extracts knowledge bytes from new script input and responds to the writer's inquiries in real-time.

---

[400] Rushit Sanghrajka et al., *Computer-Assisted Authoring for Natural Language Story Scripts*, 32 (1) AAAI CONF. ON ARTIFICIAL INTELLIGENCE 7811 (2018).

[401] *See* Marcel Marti et al., *CARDINAL: Computer Assisted Authoring of Movie Scripts*, 23RD INT'L CONF. ON INTELLIGENT USER INTERFACES 509 (2018).

[402] *See* Vincent Fortuin et al., *InspireMe: Learning Sequence Models for Stories*, 32 (1) AAAI CONF. ON ARTIFICIAL INTELLIGENCE 7747 (2018).

[403] Sanghrajka, *supra* note 400 at 7817.

The model can improve writing efficiency and prevent contradictions within a long story or a story series.

Another challenge for scriptwriters is they can only view their scripts in a textual format during the writing process.[404] Subject to the single perspective limitation, writers may miss crucial details or lack a complete overview of their works. To help writers better understand and organize their scripts, Disney Research created a computer-assisted authoring tool called CARDINAL to provide writers with various perspectives to view their ongoing works. A remote natural language processing (NLP) parser is used to extract the subject (nouns), relation (verbs), object (nouns), and relation modifiers (adverbs) from the script text and translate it into a form of subject-verb-object triples (affordances). The occurrence of affordances is counted, and the importance of affordances is assigned. Each parallel line with a distinct color represents an actor in an interactive view model. Each actor's action toward an object (affordances) is shown chronologically as a graphical element on their corresponding lines. The parallel lines will move closer when the actors have interactions. The writer can interact with this view by editing the script or lines. CARDINAL can also generate 2D and 3D previews (animations) from the affordances, giving the writer a concrete vision of each scene and solving the potential conflicts of actors' positions.

The most challenging part during the writing process might be writer's block, which means the writer feels stuck and cannot move forward with writing.[405] To solve this problem, Disney Research created a model to generate suggestions to writers when needed.

---

[404] *See* Marti et al., *supra* note 401.

[405] Fortuin, *supra* note 402, at 7747.

The research team defines stories as sequences of events or scenes. An event is defined as "an action that is performed by some character in the story, potentially with respect to some object."[406] Scene is defined as "a story part with a constant set of characters at a constant location."[407] If a model can learn the pattern of event or scene sequences, it can predict the next one based on the previous one. Hence, it can generate suggestions for writers based on the scene or event already completed by the writers. The team used 1,054 movie scripts and 8,459 novels for training except for keeping 64 stories for validation and 80 stories for testing. Natural language processing was used to extract 2,928 actions, 7,997 objects, and 47,557 locations from the training data and used to encode the story features. A long short-term memory (LSTM) network was then used to learn the pattern and predict the story sequence. They found certain features and correlations were applicable to most motion picture. The study revealed the suggestions generated by the model even got higher ratings from the test users than the actual sequences of real works.

B. GPT-3

As GPT-2's successor, GPT-3 (Generative Pre-trained Transformer 3) was announced by OpenAI in June 2020.[408] OpenAI is an AI research laboratory mainly funded by Microsoft in 2015.[409] GPT-3 is an autoregressive language model handling Natural Language Processing (NLP) tasks.[410] Unlike other language models requiring not only a

---

[406] *Id. at* 7748.
[407] *Id.*
[408] *See* Vanessa Bates Ramirez, *OpenAI's New Text Generator Writes Even More Like a Human*, SINGULARITY HUB (June 18, 2020), https://singularityhub.com/2020/06/18/openais-new-text-generator-writes-even-more-like-a-human/ [https://perma.cc/8V2L-8K95].
[409] Luciano Floridi & Massimo Chiriatti, *GPT-3: Its Nature, Scope, Limits, and Consequences*, 30 MINDS & MACHINES 681, 684 (2020).
[410] Tom B Brown et al., *Language Models are Few-Shot Learners*, ARXIV: 2005.14165 (2020).

large corpus of text to pre-train but also thousands to hundreds of thousands of task-specific examples to fine-tune, GPT-3 can better perform novel NLP tasks given with few (typically 10 to 100), one, or even zero examples for the specific desired task.[411] This is because GPT-3 uses 175 billion parameters for training, which is ten times more than its predecessor GPT-2, which had only 1.5 billion parameters.[412]

GPT-3 is an excellent text-processing tool. It can translate, answer questions, fill a cloze test, unscramble words, make a sentence with a novel word, perform 3-digit arithmetic, or produce news articles comparable to human authors.[413] Users can even access GPT-3 by using an API (Application Programming Interface) to generate stories automatically with simple steps and instructions.[414] For example, "ShortlyAI" is a downstream AI writing tool derived from GPT-3. Its website claims, "We are using advanced AI in a way that no one else ever has - to extend your creativity, to write for you. Just click the button, and the AI will continue writing for you."[415]

In 2020, some film students tried to use "ShortlyAI" to develop a three-and-half-minute screenplay named *Solicitors*.[416] Though not specifically trained for the script-writing purpose, "Shortly" completed everything except for the open scene ("Barb's reading a book. A knock on the door. She stands and opens it. Rudy, goofy-looking, stands on the other side") and the first two dialogue lines given by the students. The story turned

---

[411] *Id.*

[412] Ramirez, *supra* note 408.

[413] Brown, *supra* note 410.

[414] Shruti Saravanan & K. Sudha, *GPT-3 Powered System for Content Generation and Transformation*, 5TH INT'L CONF. ON COMPUTATIONAL INTELLIGENCE AND COMM. TECHS (CCICT) 514 (2022).

[415] FREQUENTLY ASKED QUESTIONS, SHORTLYAI, https://www.shortlyai.com/pricing#faqs[https://perma.cc/ 93R2-8SG4].

[416] Vanessa Bates Ramirez, *OpenAI's GPT-3 Wrote This Short Film—Even the Twist at the End*, SINGULARITY HUB (Oct. 23, 2020), https://singularityhub.com/2020/10/23/an-ai-wrote-this-short-film-and-its-sort-of-fascinating/ [https://perma.cc/5TU2-P977].

out to be a thriller when it finished with Barb killing the visitor. The film was commented,

"It's short, and weird, and honestly not that good. But…it's also not all that bad, especially

given that it was written by a machine."[417]

Nonetheless, being able to generate human-like text does not mean GPT-3 knows

what it is doing. A research reveals that GPT-3 performs poorly in advanced mathematics,

semantics, and ethics tests at this stage.[418] The GPT-3 development team also clearly points

out, "GPT-3 samples still sometimes repeat themselves semantically at the document level,

start to lose coherence over sufficiently long passages, contradict themselves, and

occasionally contain non-sequitur sentences or paragraphs."[419] Besides, like common

problems seen in most deep learning systems, GPT-3 generated content is not easily

interpretable, performs poorly on novel instructions, and is subjected to the bias that existed

in the training data.[420]

It is worth mentioning that two new GPT-3 family models, "text-davinci-003" and

"ChatGPT," were released at the end of 2022. The former can follow more complex

instructions to generate higher quality and longer content.[421] The latter can interact with

people in a dialogue format like humans.[422] With the machine's improved capability to

produce human-like text following short instructions, many text-producing jobs that handle

basic facts or information might disappear. It is predicted the most wanted people's core

---

[417] *Id.*
[418] Floridi, *supra* note 409.
[419] Brown, *supra* note 410 at 33.
[420] *Id.* at 34.
[421] Ralph, *How do text-davinci-002 and text-davinci-003 differ?* OPENAI, https://help.openai.com/en/ articles/6779149-how-do-text-davinci-002-and-text-davinci-003-differ [https://perma.cc/77D6-ZCW9].
[422] INTRODUCING CHATGPT, OPENAI (NOV. 30, 2022), https://openai.com/blog/chatgpt/ [https://perma.cc/ 8XAZ-MYMP].

skills will be "prompt & collate."[423] We will need people who are good at giving machine source inputs (prompts) to obtain the best outputs and can collect and combine (collate) valuable results to create coherent and meaningful stories.

C. GPT-4

As GPT-3's successor, OpenAI published the GPT-4 Technical Report in March 2023.[424] GPT-4 can accept both image prompts and text prompts. But its knowledge is still limited by the pretraining data cutting off in September 2021, and it does not learn from experience. According to the report, GPT-4 outperformed GPT-3.5 in terms of English language performance, following users' intent and refusing unsafe instructions. In a simulated bar exam, its scores fell in the top 10% of test takers, whereas GPT-3.5 fell in the bottom 10% of all test takers.

Nonetheless, GPT-4 shares the same limitations as its predecessors GPT-3 and humans in reliability, correctness, and biases, even though its performance is improved. Users are still recommended to take proper cautious measures. GPT-4 can accept a longer message and generate a longer output because the respective token limits for GPT-4 and GPT-4-32k are 8192 tokens and 32768 tokens, which are much more than GPT-35-turbo's token limit, 4096 tokens.[425]

---

[423] Floridi, *supra* note 409 at 691.

[424] OpenAI, *GPT-4 Technical Report*, ARXIV:2303.08774 (2023).

[425] 100 tokens ~= 75 words. *See* Raf, *supra* note 394; *see also* Michael Mrbullwinkle & Eric Urban, *Learn how to work with the ChatGPT and GPT-4 models*, AZURE (May 05, 2023), https://learn.microsoft.com/en-us/azure/cognitive-services/openai/how-to/chatgpt?pivots=programming-language-chat-completions [https://perma.cc/U4G5-8W25].

D. DRAMATRON

In 2022, Deepmind (backed by Google) released a script-writing tool named Dramatron.[426] It was developed on top of an in-house large language model called Chinchilla.[427] Though the amount of Chinchilla's parameters (70 billion) is fewer than GPT-3's (175 billion), Chinchilla is said to outperform GPT-3 because its training tokens (1.3 trillion) are much more than GPT-3's (300 million tokens). Nonetheless, as Chinchilla has not been released yet,[428] users have to connect to alternative large language models like GPT-3 for using Dramatron.[429] It is worth mentioning that the Dramatron website declares:[430] (1) Dramatron is a co-writing system that has to collaborate with human authors, (2) Dramatron's writing structure doesn't meet every writer's writing process, (3) the output may plagiarize the text of used language model's training data, and (4) the output may contain bias, stereotypes, and offensive text.

Dramatron is defined as an "interactive co-creative system."[431] It starts to generate a long and coherent script based on simple log lines entered by the user. The log lines are

---

[426] *See* Matthias Bastian, *Deepmind's "Dramatron" Can Write Film and Theater Scripts*, DECODER (Dec. 10, 2022), https://the-decoder.com/deepminds-dramatron-can-write-film-and-theater-scripts/ [https://perma.cc/73QV-BQTF].

[427] *See* Maximilian Schreiner, *Deepmind: Artificial Intelligence is far from being fed up*, DECODER (Apr. 23, 2022), https://the-decoder.com/deepmind-artificial-intelligence-is-far-from-being-fed-up/ [https://perma.cc/HP7J-Q8FQ].

[428] *See* Alberto Romero, *A New AI Trend: Chinchilla (70B) Greatly Outperforms GPT-3 (175B) and Gopher (280B)*, TOWARDS DATA SCIENCE (Apr. 11, 2022), https://towardsdatascience.com/a-new-ai-trend-chinchilla-70b-greatly-outperforms-gpt-3-175b-and-gopher-280b-408b9b4510 [https://perma.cc/UQ88-BVZU].

[429] Kris Holt, *Deepmind Created An AI Tool That Can Help Generate Rough Film and Stage Scripts*, ENGADGET (Dec. 9, 2022 4:23 PM), https://www.engadget.com/deepmind-ai-tool-film-theater-scripts-dramatron-212328597.html [https://perma.cc/GV7G-V6VE].

[430] ABOUT DRAMATRON, https://deepmind.github.io/dramatron/details.html [https://perma.cc/YLW4-D9AL].

[431] Mirowski et al., *supra* note 46, at 1.

a few simple sentences summarizing the story setting, characters, conflicts, or incidents.

For example,[432]

> A science - fiction fantasy about a naive but ambitious farm boy from a backwater desert who discovers powers he never knew he had when he teams up with a feisty princess, a mercenary space pilot and an old wizard warrior to lead a ragtag rebellion against the sinister forces of the evil Galactic Empire.

Based on the log lines, Dramatron will automatically generate a title such as "The Death Star's Menace."[433] Moreover, Dramatron can create characters one by one, plot scenes one by one, and locations one by one, as well as add dialogue to each scene in order. At any point in the generating process, the user can let the system work alone or jump in to edit the content until he or she is satisfied. The user can also go back to edit the log lines.

Noting that a long script generated by large language models (LLMs) will lose coherence when it reaches the limitation of 2048 tokens or about 1500 words, Dramatron adopts a method called *hierarchical story generation* to overcome the bottleneck.[434] It chains the log line, characters, plots, locations, and dialogue together in a three hierarchical layers structure. Each part can directly refer to the log line and the previous output to maintain long-term coherence and reincorporation. Such a design removes the current need for a human to select LLM's outputs to keep a machine-generated lengthy article coherent.[435] Deepmind engaged fifteen professional scriptwriters to test Dramatron and received positive feedback.[436] A comment states, "You know, with a bit of editing, I could

---

[432] *Id.* at 31.
[433] *Id.*
[434] *Id.* at 1.
[435] *Id.* at 4.
[436] *Id.* at 9.

take that to Netflix: just need to finesse it a little bit."[437] The Edmonton International Fringe Theatre Festival staged five scripts co-written with Dramatron in August 2022. One reviewer comments that the public performance "proves that artificial intelligence can in fact write a hit Fringe play."[438]

IV. DATA-DRIVEN FILMMAKING PRACTICES

Data science has been widely used in the motion picture industry, e.g., greenlighting movie projects, packaging the talents, hiring the actors and actresses, managing shooting location and sequence, editing the trailers, identifying the target audience, and determining whom, when, where, and how to distribute and market the films.[439] Just as Jim Wuthrich, president of Warner Bros. Worldwide Home Entertainment and Games, says, "From deciding which content to make, how to price, where and what to market, data influences everything we do."[440] Big streaming platforms like Netflix have their own predictive models by taking advantage of possessing huge amounts of audience data. Hollywood majors have either stepped into the streaming service market to harvest data or cooperated with data predictive platforms to catch the data-driven trend. Data-driven creativity has even become part of screen production education.[441]

---

[437] *Id.* at 9.
[438] *Id.* at 13.
[439] *See* Anurag Kapoor, *Datawatch: lights, camera, analytics – the secret role data plays in movie making*, SMART CUBE (Nov. 3, 2021), https://www.thesmartcube.com/resources/blog/datawatch-lights-camera-analytics-the-secret-role-data-plays-in-movie-making/ [https://perma.cc/XHQ3-RGR2]; *See also* Song Minzheong, *A Study on the Predictive Analytics Powered by the Artificial Intelligence in the Movie Industry*, 10(4) INT'L J. OF ADVANCED SMART CONVERGENCE 72 (2021).
[440] Prange, *supra* note 323.
[441] Stuart Bender & Billy Sung, *Data-driven creativity for screen production students: developing and testing learning materials involving audience biometrics*, 31(2) DIGITAL CREATIVITY 98 (2020).

A. NETFLIX

Growing up to be a giant streaming service platform from a small DVD-rental business, Netflix is well-known for its powerful recommender system that serves 238.3 million subscribers worldwide with an extensive film library.[442] Not long later, Netflix is unsatisfied with staying in the downstream entertainment industry as a pure motion picture distributor but wants at least 50% of its library to be original shows.[443] Data technology also helps Netflix predict the next hit that may look risky in networks, such as *House of Cards, Orange is the New Black, The Queen's Gambit*, and *Umbrella Academy*.[444] Asked if machines make decisions at Netflix, Cindy Holland, Netflix's former vice president of original content, says, "Humans are making the decisions about what we choose to invest in, but we're aided by the info we have."[445] The following paragraphs will review how data helps Netflix in various aspects.

1. Films Recommendation

A smart recommender system enhances subscribers' experiences and increases subscriber retention, especially when they visit Netflix's website without planning what to watch.[446] Before Netflix launched its streaming service in 2007, the only data available for

---

[442] Mansoor Iqbal, *Netflix Revenue and Usage Statistics (2023)*, BUSINESSOFAPPS (May 24, 2023), https://www.businessofapps.com/data/netflix-statistics/ [https://perma.cc/A2ZS-DGLY].

[443] Prange, *supra* note 323.

[444] Elizabeth Mixson, *Data Science at Netflix: How Advanced Data & Analytics Helps Netflix Generate Billions*, AI, DATA & ANALYTICS NETWORK (Mar. 30, 2021), https://www.aidataanalytics.network/data-science-ai/articles/data-science-at-netflix-how-advanced-data-analytics-helped-netflix-generate-billion s#:~:text=Using [https://perma.cc/BC9H-8WL6].

[445] Manori Ravindran, *Netflix: Humans, not machines, greenlight our shows*, TBI (Mar. 11, 2019), https://tbivision.com/2019/03/11/netflix-humans-not-machines-greenlight-our-shows/ [https://perma.cc /5SSE-BQPW].

[446] *See* Steck, *supra* note 356 at 8.

analysis were customer ID, movie ID, rating, and the date the movie was watched.[447] But today, Netflix knows exactly "what each subscriber had viewed, when, how long, and on what device."[448] Each customer's implicit feedback like plays, clicks, or searches, as well as explicit feedback like ratings, i.e., "not for me," "I like this," or "love this!" are all collected.[449] With so much data in hand and the application of machine learning models, Netflix defines nearly 80,000 micro-genres of movies[450] and creates a personalized homepage for each subscriber.[451]

2. Originals Greenlight

Investing in an original film or series, also called a title, is an expensive decision full of uncertainty. Up to August 2022, Netflix originals have occupied 50.017% of its library (3,104 originals of the total 6,206 titles) in the United States market.[452] Netflix's content budget has also grown to US$17 billion in 2022.[453] The most expensive original series, *Stranger Things*, cost $30 million per episode.[454] Thus, before greenlighting a title, content decision-makers always ask two questions: "Which existing titles are comparable

---

[447] *See* BERNARD MARR, BIG DATA IN PRACTICE: HOW 45 SUCCESSFUL COMPANIES USED BIG DATA ANALYTICS TO DELIVER EXTRAORDINARY RESULTS 18 (2016).
[448] SMITH, *supra* note 139, at 8.
[449] *See* Steck, *supra* note 356 at 8.
[450] *See* MARR, *supra* note 447 at 19.
[451] *See* Steck, *supra* note 356 at 10.
[452] Kasey Moore, *Netflix Originals Now Make up 50% of Overall US Library*, WHAT'S ON NETFLIX (Aug. 24, 2022, 9:45 AM EST), https://www.whats-on-netflix.com/news/50-of-netflixs-library-is-now-made-of-netflix-originals/ [https://perma.cc/SNN5-XP4F].
[453] Jennifer Maas, *Netflix's Ted Sarandos Is Feeling 'Better and Better' About $17 Billion Content Budget: 'We're Spending at About the Right Level'*, VARIETY (Oct. 18, 2022, 3:53 PM PT), https://variety.com/2022/tv/news/netflix-content-spend-17-billion-subscriber-growth-1235407818/ [https://perma.cc/3XTZ-GNU2].
[454] Julia Stoll, *Most expensive Netflix original series 2022, by production cost per episode*, STATISTA (Oct 19, 2022), https://www.statista.com/statistics/1249573/most-expensive-netflix-original-series-production-cost-per-episode/ [https://perma.cc/DRX8-VGEM].

and in what ways?" and "What audience size can we expect and in which regions?"[455]

To answer the above two questions, Netflix applies machine learning and statistical modeling to extract parameters from enormous amounts of historical data, including title metadata (e.g., genre, runtime, film, or series), tags, and text summaries describing thematic or plot elements.[456] A downstream model uses the extracted parameters for identifying similar titles and estimating audience sizing. Each title is embedded in a similarity map to compare similar titles easily. The more similar the titles, the closer the titles will locate. To estimate audience sizing, a supervised model makes predictions based on the input of title embedding, audience sizes of similar titles embedding, and country embedding. Various data tools and models, such as Bert, LSTM, GRU, Knowledge graphs, and TransE, are applied to produce embeddings of diverse types of source data.

With the aid of data prediction, Netflix dramatically changes the entertainment industry's practice to greenlight a show.[457] It abandons conventional pilot practice.[458] The traditional way to produce a television series is to persuade a television network to fund a 30- or 60-minute pilot episode. The networks rarely commit to a series before seeing a completed pilot episode. Needless to say, giving a whole vision of a story and characters in a pilot episode is difficult for a writer. Such custom is also a massive waste because many pilots spending US$5 million to US$6 million ultimately do not enter a series. While the networks hesitated to fund the *House of Cards* pilot episode based on the conventional

---

[455] Melody Dye et al., *Supporting content decision makers with machine learning*, NETFLIX TECH. BLOG (Dec. 10, 2020), https://netflixtechblog.com/supporting-content-decision-makers-with-machine-learning-995b7b76006f [https://perma.cc/N6BB-DQ7T].
[456] *Id.*
[457] *See* SMITH, *supra* note 139, at 3-5.
[458] Peter White, Netflix Comedy Chief Tracey Pakosta on Netflix Is A Joke Festival, Multicamera Conundrum & Laughter Pipeline, DEADLINE (Apr. 28, 2022, 1:28 PM), https://deadline.com/2022/04/netflix-comedy-tracey-pakosta-netflix-is-a-joke-1235010699/ [https://perma.cc/9GFX-SGCH].

wisdom that political dramas would not appeal to the audience, Netflix saw an opportunity.[459] Based on historical data, Netflix knew its audiences loved David Fincher-directed movies and Kevin Spacey-led movies, as well as many customers rented DVDs of the original BBC *House of Cards*. Knowing the existence of audiences, Netflix bypassed a pilot and committed straight to a two-season order of 26 *House of Cards* episodes at once.

Knowing the audience's specific tastes, the writer of *House of Cards* even enjoyed more creative freedom at Netflix than working for the national television network. As an on-demand streaming platform providing a large film library without commercials, Netflix can provide alternative choices for its subscribers to meet diverse tastes simultaneously with no commercial-selling pressures.[460] Therefore, Beau Willimon, the writer of *House of Cards*, said he indeed enjoyed more creative freedom to include some controversial scenes that might be unacceptable for some people. He also had more flexibility to structure his story in a continuous time framework without a standard time limit. Such freedom was unthinkable when writing scripts for television networks because networks had to take care of public opinions and advertisers' needs.

3. Target Marketing

In addition to producing personalized recommendations and content development analytics, data science is essential in targeting audiences. By knowing its customers individually, Netflix creates multiple trailers to draw customers' attention based on their preferences.[461] Computer vision algorithms also help Netflix to annotate almost 9 million

---

[459] SMITH, *supra* note 139, at 5.
[460] *See id.* at 6-8.
[461] SMITH, *supra* note 139, at 8.

video frames of a TV show for ranking algorithms to identify attractive thumbnail images to individual audiences.[462]

4. Content Generation

We have known about data aiding Netflix to decide which shows to greenlight. But it is unclear if data helps Netflix to create the shows' scripts and how much data is involved. Netflix gives us a reason to believe data can do so. In October 2021, Netflix released a four-minute horror movie titled *Mr. Puzzles wants you to be less alive*.[463] "Netflix Is A Joke" YouTube channel annotated this film "The First Horror Movie Written Entirely By Bots."[464] The short animation depicts a devil named Mr. Puzzle kidnaps a young lady and forces her to answer death puzzles. The lady is ultimately killed because she failed to answer the fifth puzzle. Netflix completes the film by collaborating with Keaton Pattit[465] to train a bot on 400,000 more hours of horror movies.[466] It extracts various elements from at least five popular horror films.[467] Netflix described the film as "quite twisted and pokes fun at various horror tropes" and the dialogue as "sardonic, flat and hilarious, uttered in a

---

[462] Madeline Riley et al., *AVA: The Art and Science of Image Discovery at Netflix*, NETFLIX TECH. BLOG (Feb. 7, 2018), https://netflixtechblog.com/ava-the-art-and-science-of-image-discovery-at-netflix-a442f163af6 [https://perma.cc/5FM6-QDN5].

[463] Netflix Is A Joke, *The First Horror Movie Written Entirely By Bots*, YOUTUBE (Oct. 6, 2021), https://www.youtube.com/watch?v=WZzbxNoMjGM.

[464] *Id.*

[465] Keaton Pattit is a writer and comedian, who is also the author of "I Forced a Bot to Write This Book: A.I. Meets B.S." *See* I FORCED A BOT TO WRITE THIS BOOK: A.I. MEETS B.S., AMAZON, https://www.amazon.com/Forced-Bot-Write-This-Book-ebook/dp/B08L5TRWR4?ref_=ast_sto_dp

[466] *See* Netflix Is A Joke, *supra* 462.

[467] *See* Debadrita Sur, *Netflix releases its first-ever bot-written horror film*, BEST OF NETFLIX (Oct. 8, 2021), https://best-of-netflix.com/netflix-releases-its-first-ever-bot-written-horror-film/ [https://perma.cc/RF9B-BC37].

monotone robotic voice." [468] But it was excellent at "mak[ing] the viewers ride a rollercoaster of emotions."[469]

B. 20TH CENTURY FOX

Working with Google Cloud, 20th Century Fox has built a deep learning model trained with customer data and movie scripts to predict the pattern of audiences' preferences toward diverse types of films. [470] This prediction has become a routine consideration "to support greenlighting decisions, movie positioning studies, and marketing and distribution."[471] I will introduce how 20th Century Fox employs data technology to predict audience attendance and make movie trailers in the following.

1. Audience Prediction

Merlin is a movie prediction and recommendation system developed by 20th Century Fox and Google's Advanced Solutions Lab.[472] It is a hybrid collaborative filtering pipeline for analyzing hundreds of movies and millions of user attendance records.[473] Each movie and each user are represented as a separate vector so that correlation of two vectors can be analyzed. Merlin applies the logistic regression method to predict movie attendance probability by evaluating the interrelation between the subject movie and the customer's

---

[468] *Id.*

[469] *Id.*

[470] *See* Miguel Campo-Rembado & Sona Oakley, *How 20th Century Fox uses ML to predict a movie audience*, GOOGLE CLOUD (Oct. 29, 2018), https://cloud.google.com/blog/products/ai-machine-learning/how-20th-century-fox-uses-ml-to-predict-a-movie-audience [https://perma.cc/HG8J-YJR7].

[471] Hsieh Chengkang et al., *Convolutional Collaborative Filter Network for Video Based Recommendation Systems*, ARXIV:1810.08189 1 (2018).

[472] *See* Rembado, *supra* note 470.

[473] *See* Hsieh, *supra* note 471.

moviegoing history, frequency, and recency. The prediction is based on the hypothesis that customers will likely watch a movie similar to what they have watched.

A movie trailer contains the unique elements of the story and the commonality of the same genre, so Merlin applies a Video Convolutional Neural Network to exploit the movie trailer data.[474] It divides the first 120 seconds of a movie trailer into 120 video frames per second. 1024 image features are extracted from each frame and are processed by a layer containing 1024 convolutional filters. Each filter can handle 8 frames X 1024 image features. In addition, one residual layer with 1024 convolutional filters capable of handling one frame X 1024 image features is applied to increase the model capacity. And an average pooling layer is applied to summarize the trailer signals. The network can learn objects in a single frame (e.g., car and face), temporal sequences of objects across multiple frames (e.g., a car chase followed by a car flipping), and inter-correlation of two object-sequences (e.g., gun fighting and an explosion). The output is fed into a multi-layer perceptron system to find the most predictive signals. Through training and testing, the research finds predictions based on a movie trailer analyzed by the convolutional network outperform a prediction based on a movie synopsis.

To verify Merlin's predictive accuracy, the 20th Century Fox data team applied the model to analyze the movie trailer of *Logan* (2017), an action and adventure movie led by Hugh Jackman. Merlin summed up *Logan's* top 10 labels by frequency: tree, facial hair, car, man, vehicle, atmosphere, mammal, beard, forest, and light. By measuring the label temporal sequences and numerous elements in the trailer, Merlin discovered the top 20

---

[474] *See id.*

movie trailers depicting similar features. It predicted the 20 movies' previous audiences would also be *Logan's* potential audiences. After *Logan* was released, they investigated its actual audiences' moviegoing history. As a result, the top five movies that *Logan's* actual audiences have seen, e.g., *X-men apocalypse, John Wick: Chapter 2, Doctor Strange, Batman vs. Superman Dawn of Justice,* and *Suicide Squad*, all were predicted in the list of predicted 20 movies. Merline found the already-known core superhero audience and discovered new audiences who liked the type of "rugged male action lead" movies. Since the release of *The Greatest Showman* in 2017, 20th Century Fox has been using Merline to predict audience intent.

2. Movie Trailer Editing

In addition, 20th Century Fox leveraged big data and machine learning to make a movie trailer for its horror movie "*Morgan*" in 2016.[475] Partnering with IBM Watson, the research team trained a model to learn the basic elements and common structures of horror movie trailers by dividing a hundred horror movie trailers into moments and analyzing the emotional value within individual scene's image, sound, and composition. After being fed the ninety-two-minute "*Morgan,*" the model selected six-minute clips without specific instructions. The model spent around twenty-four hours selecting footage, which was significantly less than the traditional material selection time ranging from ten to thirty days. However, the model could not make a movie trailer entirely. Hence, a human editor had to rearrange the model-selected footage and add black titles and background music to

---

[475] *See* John R. Smith, *IBM Research Takes Watson to Hollywood with the First "Cognitive Movie Trailer,"* IBM THINK BLOG (Aug. 31, 2016), https://www.ibm.com/blogs/think/2016/08/cognitive-movie-trailer/ [https://perma.cc/RX6A-ZCRE].

complete the trailer. Not every model-selected moment was predicted by the research team, and one moment was not included in the final trailer.

C. BENJAMIN

Benjamin is a deep learning model structured with Long Short-Term Memory (LSTM).[476] LSTM can recognize text and learn the connections of letters, words, and phrases. It can better understand a whole paragraph and generate original sentences than general RNN. Benjamin's builder, Ross Goodwin, partnering with the director Oscar Sharp, trained Benjamin to write screenplays by feeding it dozens of the 1980s and 1990s sci-fi screenplays available online. Eventually, Benjamin can imitate a screenplay format, including stage directions and character lines.

1. *Sunspring*

Sunspring is a nine-minute experimental movie.[477] It was written by Benjamin for attending the 2016 Sci-Fi London 48-Hour Film Challenge and was placed in the top ten of hundreds of contesting films. [478] Note that its movie script is the first one written entirely by Artificial Intelligence. Some stage directions, such as "taking his eyes from his mouth" and "He is standing in the stars and sitting on the floor," are weird.[479] The dialogue between the two male characters and one female character is often incoherent and difficult to understand. A reviewer comments, "The experiment shows AI's potential and also

---

[476] *See* Annalee Newitz, *Movie written by algorithm turns out to be hilarious and intense*, ARSTECHNICA (May 30, 2021, 11:55 AM), https://arstechnica.com/gaming/2021/05/an-ai-wrote-this-movie-and-its-strangely-moving/ [https://perma.cc/R9BK-EDEB].
[477] Ars Technica, *Sunspring | A Sci-Fi Short Film Starring Thomas Middleditch*, YOUTUBE (June 9, 2016), https://www.youtube.com/watch?v=LY7x2Ihqjmc.
[478] *See id.*
[479] *See* Newitz, *supra* note 476.

demonstrates, often to hilarious effect, its shortcomings." [480] A Guardian news comments on the script, "dark, ominous atmosphere and gibberish."[481]

2. *It's No Game*

It's No Game, is another movie written by Benjamin, teamed with two human writers in 2017, depicting a story about how AI replaces the screenwriter's job.[482] By collaborating with human writers, a commenter found this movie "reinforces the notion that the current AI technology can work more efficiently in conjunction with humans rather than being left to its own devices."[483] A YouTube viewer comment that got many thumbs-ups stated, "That one actually made much more sense, than the previous one! I liked how they explored the AI creator phenomenon itself."[484] Note that the dialogue of the eight-minute Sci-Fi movie is made by six different algorithms trained with diverse materials. That made *It's No Game* present multiple styles.[485] Goodwin admitted, "Having all the different models definitely made Ben feel more like a tool, or more disbursed anyway."[486]

3. *Zone Out*

Moreover, Benjamin tried to direct a film named *Zone Out* in the Sci-Fi-London 48-Hour Challenge of 2018. The seven-minute short Sci-Fi film combines footage acted

[480] Dylan Brethour, *Sunspring Review*, FILM & TELEVISION (June 23, 2016), https://www.curatormagazine. com/dylan-brethour/sunspring-review/ [https://perma.cc/8CPN-3THC].

[481] Spam filter, *This is what happens when an AI-written screenplay is made into a film*, GUARDIAN (June 10, 2016), https://www.theguardian.com/technology/2016/jun/10/artificial-intelligence-screenplay-sunspring-silicon-valley-thomas-middleditch-ai?CMP=share_btn_tw [https://perma.cc/XA5C-FN5J].

[482] IT'S NO GAME, IMDB, https://www.imdb.com/title/tt7115406/ [https://perma.cc/34AU-RLMB].

[483] Nantheera Anantrasirichai & David Bull, *Artificial Intelligence in the Creative Industries: A Review*, 55 ARTIF. INTELL. REV. 589, 608 (2022).

[484] Anastasia Egorova (@anastasiaegorova440), YOUTUBE (Apr. 26, 2017, 7:28 AM), https://www.youtube. com/watch?v=5qPgG98_CQ8&lc=UghZ5ElNx8N6y3gCoAEC [https://perma.cc/L8VP-SW53].

[485] Annalee Newitz, *An AI wrote all of David Hasselhoff's lines in this bizarre short film*, ARSTECHNICA (Apr 25, 2017, 12:00 PM), https://arstechnica.com/gaming/2017/04/an-ai-wrote-all-of-david-hasselhoffs-lines-in-this-demented-short-film/ [https://perma.cc/CCZ3-BQKM].

[486] *Id.*

by professional actors in front of the green screen[487] and numerous clips taken from two public-domain movies, *The Last Man on Earth* and *The Brain That Wouldn't Die*.[488] Benjamin used face-swapping to replace the old film's actors' faces with the new actors' faces. This technology was powered by a Generative Adversarial Network (GAN) that could generate synthetic images and judge if they looked realistic. As for the spoken dialogue, Benjamin generated robotic voices via voice-generating technology and plugged them into appropriate places. Moreover, Benjamin scored the film's piano soundtrack based on the content emotion itself.[489] Subject to the two-day time crunch and technology's limitations, *Zone Out* was found, "Blurry faces, computer-generated dialogue, and awkward scene changes fill out this year's Zone Out."[490] Nonetheless, as for a movie wholly made by AI, "the results are horrifyingly encouraging."[491]

D. LEXUS EUROPE

In 2018, Lexus Europe released a one-minute car advertisement, which impressively depicted an automotive designer's emotional fluctuation before and after the car he designed passed a crash and safety test.[492] The creative team used IBM Watson to produce the short advertisement.[493] Watson became known to the world when it defeated

---

[487] Lauren Goode, *AI Made a Movie—and the Results Are Horrifyingly Encouraging*, WIRED (June 11, 2018, 3:00 PM), https://www.wired.com/story/ai-filmmaker-zone-out/ (last visited Aug. 17, 2023).
[488] *See* Sam Machkovech, *This wild, AI-generated film is the next step in "whole-movie puppetry,"* ARSTECHNICA (June 11, 2018, 12:01 PM), https://arstechnica.com/gaming/2018/06/this-wild-ai-generated-film-is-the-next-step-in-whole-movie-puppetry/ [https://perma.cc/E39B-UA9C].
[489] *See id.*
[490] *Id.*
[491] Goode, *supra* note 487.
[492] Lexus Europe, *Lexus ES | Driven by Intuition | A film made by Artificial Intelligence | Lexus Europe*, YOUTUBE (Nov. 19, 2018), https://www.youtube.com/watch?v=-iaBJ5rqOdg&t=10s.
[493] Reece Medway, *Lexus Europe Creates World's Most Intuitive Car Ad with IBM Watson*, IBM THINK BLOG (Nov. 19, 2018), https://www.ibm.com/blogs/think/2018/11/lexus-europe-creates-worlds-most-intuitive-car-ad-with-ibm-watson/ [https://perma.cc/2TJH-PSMU].

two human contestants in a Jeopardy game on a national TV show in 2011.[494] Rob Thomas, Senior Vice President of IBM Cloud and Data Platform describes, "Watson is software capable of making sense of data sets and understanding natural language to provide recommendations, make predictions, and automate work."[495] Today, Watson is a collective term for IBM's data processing, data predictive, and AI-related tools, products, and services.

Several tools, such as Watson Visual Recognition, Watson Tone Analyzer, Watson Text to Speech, and Watson Personality Insights, were used to produce the advert script.[496] Common elements were extracted from the footage, text, and audio of award-winning auto commercials across 15 years, Lexus brand data, and other external data. The creative team used AI-generated script flow and outline to build the story. Sequences of scenes and the soundtrack presented the plot and emotion but no dialogue. Kevin Macdonald, the advert's director, also an Oscar-winner, said, "When I was handed the script, the melodrama of the story convinced me of its potential . . . . The charmingly simplistic way the AI wrote the story was both fascinating in its interpretation of human emotion, and yet still unexpected enough to give the film a clearly non-human edge."[497]

E. DATA-DRIVEN SCREEN EDUCATION

Film data analytics can help with script analysis, casting recommendation, and marketing strategy, but other cinematic skills, such as cinematography, color, lighting,

---

[494] *See* John Markoff, *Computer Wins on 'Jeopardy!': Trivial, It's Not*, N.Y. TIMES (Feb. 16, 2011), https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html.

[495] Rob Thomas, *IBM Watson: Reflections and Projections*, IBM THINK BLOG (Oct. 10, 2019), https://www.ibm.com/blogs/think/2019/10/what-is-watson/ [https://perma.cc/UN3P-XGN2].

[496] *See* Medway, *supra* note 493.

[497] *Id.*

sound, and music editing, are beyond their territories.[498] How to use particular cinematic techniques to impact the audience is usually considered an established filmmaker's intuitive, artistic vision and aesthetic judgment, which is very difficult for a student filmmaker to learn and replicate.[499] To compensate for the insufficiency of traditional screen theory and production practice history teaching, a screen pedagogy called the Data-Driven Creativity Project (DDCP) was developed at Curtin University, Australia.[500] This pedagogy is based on cognitive media theory which studies the audience's response to creative choices of screen images and sounds.

The project used Gazepoint Analysis software and Noldus FaceReader to track and collect the audience's bio-metric data moment by moment while they were watching films. The data included eye movement (for visual attention analysis), skin conductance level (for arousal/tension analysis), and facial expression (for emotion analysis). In one case, they collected data from 117 participants aged 18 to 30. By visualizing the data into simple and understandable graphs, student filmmakers can easily learn how the audience reacts momentarily to the creative decisions of other student-produced short films. The project gained positive feedback from students. The students were encouraged to transfer what they learned from the project to produce the desired effects and avoid making similar mistakes in their films. The project hoped to include more audience response data and became a part of regular screen education.

---

[498] Nadide Gizem Akgulgil Mutlu, *The future of film-making: Data-driven movie-making techniques*, 10(2) GLOBAL J. OF ARTS EDUC. 167, 173 (2020).
[499] Bender, *supra* note 441, at 98.
[500] *Id.*

V. CONCLUSION

Unlike previous scientific audience research that could only investigate sample representative audiences and rely on the audience's honest feedback, the streaming platforms can collect first-hand and actual audience reactions in real time. Besides, novel data processing and mining methods can extract meaningful information from vast amounts of audience data, including their viewing history, demographics, and feedback, as well as enormous amounts of content data, including texts, audio, and videos. By using machine learning algorithms, the hidden patterns regarding the interrelationship between audiences and contents can be uncovered to predict audiences' possible reactions to a new motion picture. Obviously, the quantity, quality, sophistication, and speed of modern data technologies are far better than previous scientific audience research.

However, just like other novel technologies have pros and cons, so does data technology. First, establishing a complete data ecosystem is complex and expensive (see **Figure 3**). To do so requires a large-scale computer system, software, data professionals, and, most importantly, continuous big data input. This is bound to be an unfair war for small players in motion picture industry. Even if cloud-based large language models and data analytic platforms are available, small players still need help to target audiences because they lack direct audience data. Secondly, like previous scientific audience research, data predictive models predict the audience's future behaviors based on past experiences. New-concept scripts may have few chances to be made into a motion picture as no data supports their greenlights. Thus, many worry that data-driven trends will restrain creativity despite the leading big data users' repeated refutation. Needless to say, big data has various concerns about bias, hostilities, privacies, and plagiarism.

Figure 3 Data-Driven Filmmaking Ecosystem
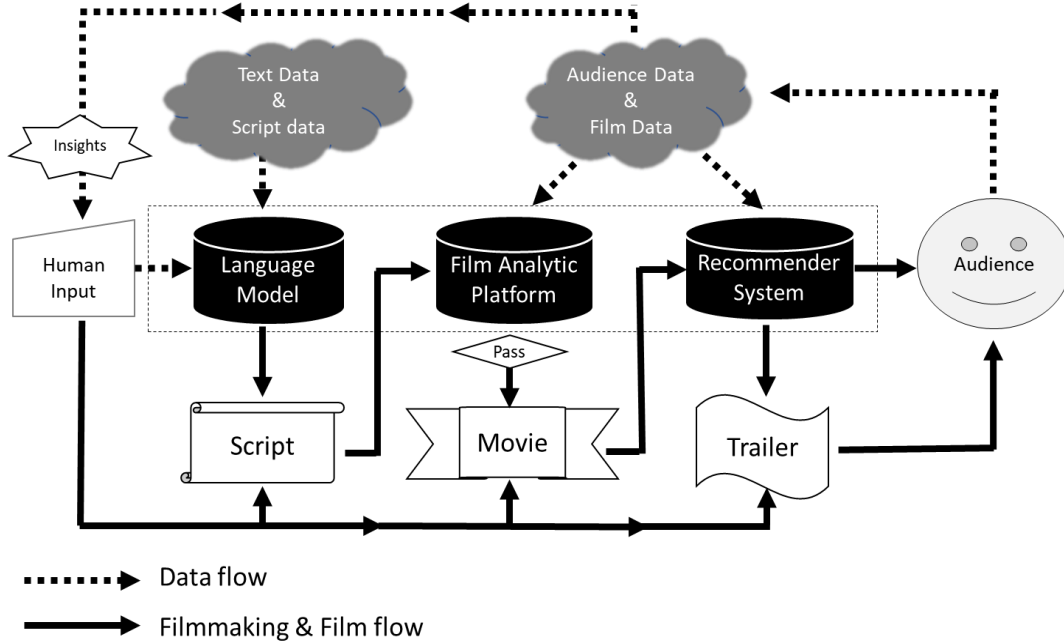
Along with the increasing availability of data sources and data technology, combined with the advancement of large language models, data-driven trends have expanded throughout motion picture pre-production, production, and post-production stages and have even extended to screen education. At present, entirely AI-generated scripts are still in the developing phase because they lack long consistency and are sometimes senseless. But increased human and machine collaborative cases are undergoing and emerging. Viewing the end products alone, it is tough to distinguish which parts are contributed by humans and which parts are contributed by machines. To evaluate whether current copyright system is adequate for the latest technological developments, I will first review current motion picture copyright thoroughly in the next chapter.

CHAPTER FOUR: CURRENT MOTION PICTURE COPYRIGHT

Before evaluating if the current copyright system can take on the recent technological developments stated in the previous chapter, it is necessary to review what approaches are being used to determine the motion picture's protection under current copyright systems. This chapter provides a thorough review of American approaches with respect to motion picture's copyrightability, the scope of copyright protection, and copyright infringement issues. In addition, brief comparisons are made with those approaches of China and Taiwan.

I. MOTION PICTURE'S COPYRIGHTABILITY

Motion pictures are a subset of audiovisual works consisting of a series of related images that, when shown in succession, impart an impression of motion. [501] Sounds are not required. Examples include films, documentaries, television shows, cartoons, videos. [502] Motion pictures have been long recognized as copyrightable cinematographic works of authorship. [503] The copyright owners of motion picture enjoy several rights. In the United States, they have exclusive rights to reproduce, prepare derivative works, distribute to the

---

[501] *See* 17 U.S.C. § 101; Zhonghua Renmin Gongheguo Zhuzuo Quan Fa Shishi Ttiaoli (中华人民共和国著作权法实施条例) [Regulation for the Implementation of the Copyright Law of the People's Republic of China] (2013 Revision) § 4(11): "cinematographic works and works created by a process analogous to cinematography" means works which are recorded on some material, consisting of a series of images, with or without accompanying sound, and which can be projected with the aid of suitable devices or communicated by other means."; Zhuzuoquan Fa Diwutiao Diyixiang Gekuan Zhuzuo Neirong Liqi (著作權法第五條第一項各款著作內容例示) [The Illustrated Contents of Each Kind of Works in Paragraph One, Article 5 of the Copyright Ac] (Taiwan) § 2(7): "Audio-visual works: shall include the images shown in any motion picture, videocassette, videodisc, on computer screen, and the other series of images which could be fixed to any medium with or without sound to be shown by a mechanical device or equipment."

[502] U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES § 503.1(B) (3d ed. 2021).

[503] *See id.* § 618.4(C).

public, and perform publicly.[504] In China, they have the moral rights of publication, authorship, alternation, and integrity, as well as the property rights of reproduction, distribution, lease, projection, broadcasting, dissemination via the internet, adaption, translation, compilation, and others.[505] In Taiwan, they have the moral rights of publication,[506] authorship,[507] alternation, and integrity,[508] as well as the property rights of reproduction,[509] broadcasting,[510] projection,[511] dissemination via the internet,[512] adaption, compilation,[513] distribution,[514] and lease.[515] Though varied terms are used to describe the above rights in the United States, China, and Taiwan, the overall protection is similar except for no moral rights for motion pictures in the United States.[516]

A. THE ORIGINAL AUTHORSHIP REQUIREMENT

1. United States

To be copyrightable, a work "must be original to the author."[517] This means the work is "independently created by the author (as opposed to copied from other works)" and possesses "at least some minimal degree of creativity."[518] The author is "he to whom

---

[504] 17 U.S.C. § 106.

[505] *See* Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment) §10, I.

[506] *See* Zhuzuoquan Fa (著作權法) [Copyright Act)] (Taiwan) § 15.

[507] *See id.* § 16.

[508] *See id.* § 17.

[509] *See id.* § 22.

[510] *See id.* § 24.

[511] *See id.* § 25.

[512] *See id.* § 26-1.

[513] *See id.* § 28.

[514] *See id.* § 28-1.

[515] *See id.* § 29.

[516] Julia Mas-Guindal, *The Motion Picture Industry: Critical Issues Concerning Moral Rights and Authorship*, JOLTDIGEST (Dec.31, 2011), https://jolt.law.harvard.edu/digest/the-motion-picture-industry-critical-issues-concerning-moral-rights-and-authorship [https://perma.cc/3MX8-DC7T].

[517] Feist Publications, Inc. v. Rural Telephone Service Co., Inc., 499 U.S. 340, 345 (1991).

[518] *Id.*

anything owes its origin; originator; maker; one who completes a work of science or literature."[519] There is a fundamental difference between an author and a discoverer in copyright law. The latter is "[t]he first person to find and report a particular fact," who "merely finds and records" but "has not created the fact."[520] In sum, a copyrightable motion picture must "originate from the author of that work" (independent creation) and "contain a sufficient amount of creative expression in the form of a series of sequential images that convey motion" (creative expression).[521]

Typically, a motion picture's production authorship "includes important decision-making about all aspects of the motion picture that affects the outcome of the final motion picture, including writing, directing, camera work, and editing."[522] The expression of a motion picture may extend to "plot, themes, dialogue, mood, setting, pace, characters, and sequence of events"[523] as well as "the individual artistic choices such as a particular montage style, camera angle, framing, hairstyle, jewelry, decor, makeup and background."[524]

A motion picture may be based on a preexisting literary work. However, due to "the difference between literary and graphic expression," no matter how detailed the literal description might be, "[t]he description of a character in prose leaves much to the imagination."[525] Therefore, the filmmaker enjoys fair creative choices for a film apart from the based literary work. "At the very least, the scope of the film copyrights covers all visual

---

[519] Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 58 (1884).
[520] *Feist Publications*, 499 U.S. at 347.
[521] COMPENDIUM (THIRD) § 808.7(A)(B).
[522] *See id.* § 808.4(A).
[523] Kouf v. Walt Disney Pictures & Television, 16 F.3d 1042, 1045 (9th Cir. 1994).
[524] Chuck Blore & Don Richman Inc. v. 20/20 Advert. Inc., 674 F. Supp. 671, 677 (D. Minn. 1987).
[525] Warner Bros. Ent. v. X One X Prods., 644 F.3d 584, 597 (8th Cir. 2011).

depictions of the film characters"[526] But the films' scope of copyright protection "is limited to the increments of character expression in the films that go beyond the character expression in the books on which they were based."[527]

It is not rare that a motion picture contains old elements, just as Judge Sneed says, "in Hollywood, as in the life of men generally, there is only rarely anything new under the sun."[528] However, "whether the materials which are used are entirely new, and have never been used before; or even that they have never been used before for the same purpose" is not a question so long as the author does not copy "the same plan, arrangement and combination of materials."[529]

For example, in *Universal Pictures Co. v. Harold Lloyd Corp.*,[530] defendant Universal Pictures' movie, *So's Your Uncle*, was found infringing the copyright of plaintiff Harold Lloyd's movie, *Movie Crazy*, by copying a sequence. Defendant contended the sequence was uncopyrightable as it was commonplace, merely comic accretion, gags, and stage business. However, the court held, "It is true that the mere motions, voice and postures of actors and mere stage business is not subject of copyright protection"[531] but "such material may be so combined with events as to become subject to copyright protection"[532] if originality is displayed by "taking commonplace material and acts and making them into a new combination and novel arrangement."[533]

---

[526] *Id.* at 598.
[527] *Id.* at 597.
[528] Berkic v. Crichton, 761 F.2d 1289, 1294 (9th Cir. 1985).
[529] Emerson v. Davies, 8 F. Cas. 615, 618 (C.C.D. Mass. 1845).
[530] Universal Pictures Co. v. Harold Lloyd Corp., 162 F.2d 354 (9th Cir. 1947).
[531] *Id.* at 363.
[532] *Id.*
[533] *Id.*

Many people, including the producer, director, writer, camera operator, editor, and others, may be involved in creating a motion picture.[534] However, everyone's creative contribution does not necessarily entitle him or her to an independent copyright on his or her contributing component apart from the copyright to the motion picture as a whole.[535] In some cases, these individuals may be joint authors if they intend to merge their contributions into inseparable or interdependent parts of a unitary whole,[536] and each person contributes a sufficient amount of original authorship to the motion picture.[537]

In practice, most motion pictures are created as "work made for hire."[538] A motion picture can be a "work made for hire" if it is either: (1) prepared by an employee within the scope of his or her employment (no written agreement required) or (2) specially ordered or commissioned by the parties who sign a written agreement stating the motion picture is a "work made for hire.[539] In the two cases, the employer or other person for whom the work is prepared, such as a motion picture studio or a producer, is considered the author for copyright purposes.[540]

---

[534] COMPENDIUM (THIRD) § 808.10(A).

[535] Garcia v. Google, Inc., 786 F.3d 733 (9th Cir. 2015) (stating, "Treating every acting performance as an independent work would not only be a logistical and financial nightmare, it would turn cast of thousands into a new mantra: copyright of thousands." Id at 743).

[536] 17 U.S.C. § 101.

[537] COMPENDIUM (THIRD) § 800.10(A)(3); see also TMTV, Corp. v. Mass Prods., Inc., 645 F.3d 464 (1st Cir. 2011) (holding that an actor participating in the show idea development, approving "the idea of setting a comedy segment in a condominium building," and suggesting "general plot ideas and stock characters" is not an author); Childress v. Taylor, 945 F.2d 500, 509 (2d Cir. 1991) (holding that an actor who furnished "the results of research concerning the life of 'Moms' Mabley, made "some incidental suggestions, contributing ideas about the presentation of the play's subject and possibly some minor bits of expression" is not "more than the helpful advice that might come from the cast, the directors, or the producers of any play. A playwright does not so easily acquire a co-author." Id. at 509).

[538] COMPENDIUM (THIRD) § 800.10(A).

[539] 17 U.S.C. § 101.

[540] Id. at. § 201.

2. China

Like the United States, the originality requirement under China's copyright law is determined by two elements: (1) independent creation, which means that the author conceives and creates the work independently without copying others' works; (2) the expression of the work must present the author's intellectual selection and judgment.[541] It would be sufficient to meet the originality's minimum requirement if the work can somewhat present the author's personality.[542] Unless there is contrary proof, the person whose name is represented on a work as the author is presumed to be the work's author.[543]

Under China Copyright Law, motion pictures, also called cinematographic works and works created in a way similar to cinematography, are a subset of audiovisual works.[544] They are copyrightable works if they are original intellectual achievements in the fields of literature, art, and science that can be presented in a certain form. Works mean "intellectual creations with originality"[545] that are created directly by intellectual activities, which do

---

[541] Beijing Xinlang Hulian Xinxi Fuwu Youxian Gongsi Yu Beijing Tianying Jiuzhou Wangluo Jishu Youxian Gongsi Buzhengdang Jingzheng Jiufen (北京新浪互联信息服务有限公司与北京天盈九州网络技术有限公司不正当竞争纠纷) [Beijing SINA Internet Info. Serv. Co. v. Beijing Tianying Jiuzhou Network Tech. Co.], 2020 Jing Min Zai No. 28 (2020 京民再 128 号) (Beijing Mun. High People's Ct. Sept. 23, 2020) [法宝引证码 CLI.C.115093220].

[542] Yangshi Guoji Wangluo Youxian Gongsi Su Shanghai Juli Chuanmei Jishu Youxian Gongsi Zhuzuo Quan Qinquan Ji Buzhengdang Jingzheng Jiufen (央视国际网络有限公司诉上海聚力传媒技术有限公司著作权侵权及不正当竞争纠纷) [China Cent. Television Int'l Network Co. v. Shanghai SynaCast Media Tech. Co.], 2020 Hu Min Zhong No. 581 (2020 沪 73 民终 581 号) (Shanghai IP Ct. Apr. 20, 2021) [法宝引证码 CLI.C.410250170].

[543] *See* Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment) § 12, III.

[544] "Audiovisual works" were named "cinematographic works and works created in a way similar to cinematography" in the copyright law before the 2020 amendment took effect on June 1, 2021. *See id.* § 3(6).

[545] Zhonghua Renmin Gongheguo Zhuzuoquan Fa Shishi Tiaoli (中华人民共和国著作权法实施条例) [Regulation for the Implementation of the Copyright Law of the People's Republic of China] (2013 Revision) § 2.

not include "organizational activity, consultation, material support or other auxiliary services conducted or offered for another person's creation . . . ."[546]

Unlike the United States, besides audiovisual works, China also protects "video recording."[547] Nonetheless, video recording producers only enjoy neighboring rights that contain fewer rights and a shorter term of protection than audiovisual work authors.[548] The distinction between "audiovisual works" and "video recordings" is based on whether they present originality, rather than the degree of originality.[549] There is no originality requirement for video recordings at all because such protection aims to encourage labor and investment efforts in disseminating works rather than rewarding creating original works. Accordingly, a cinematographic or similar work is a series of images made by cinematography or similar ways that reflects the author's independent conception, transmits certain ideas, and represents the author's personality in cinematography, selection, edition, and production of images.[550]

Conversely, a video recording is a series of consecutive images that mechanically and faithfully records preexisting works, images, or graphs. It is possible that a video recording producer makes individual selections while producing a video recording. Still,

---

[546] *Id.* at § 3.

[547] "A producer of sound recordings or video recordings shall have the right to permit others to reproduce, distribute, lease and disseminate to the public through information network such sound recordings or video recordings and shall have the right to receive remuneration for it. The term of protection of such rights shall be fifty years, expiring on December 31 of the fiftieth year after the production of the recording is firstly completed." Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment) § 44.

[548] *See id.*

[549] *See* Beijing Xinlang Hulian Xinxi Fuwu Youxian Gongsi Yu Beijing Tianying Jiuzhou Wangluo Jishu Youxian Gongsi Buzhengdang Jingzheng Jiufen (北京新浪互联信息服务有限公司与北京天盈九州网络技术有限公司不正当竞争纠纷) [Beijing SINA Internet Info. Serv. Co. v. Beijing Tianying Jiuzhou Network Tech. Co.], 2020 Jing Min Zai No. 28 (2020 京民再 128 号) (Beijing Mun. High People's Ct. Sept. 23, 2020) [法宝引证码 CLI.C.115093220].

[550] *See id.*

the purpose is to improve the images' technical processing, not involving personal selection and arrangement of the image's expression.[551]

In *China Cent. Television Int'l Network Co. v. Shanghai SynaCast Media Tech. Co.*,[552] the court classifies sequential images into four levels. The first two levels are uncopyrightable video recordings, whereas the last two levels are copyrightable audiovisual works, depending on whether the selection, edition, and management of the images reflect the author's personality. The lowest level is simply a mechanical recording, such as surveillance images taken by fixed traffic cameras, or images taken by cameras attached to animals. The second level is a faithful restoration of the shooting content without human intellectual selection despite using a small number of shots and simple shot switches, such as classroom teaching videos, celebrity speech videos, landscape picture videos. The next level involves the use of multiple shots to present the effects of scene changes, frame arrangement, frame selection, frame cutting continuity, and pace changes as well as synchronizing sounds with images, like sport events broadcast. The highest level is a full presentation of the author's creativity and the sequential images' aesthetic feelings. The creation applies skilled multiple scenes shooting techniques and professional continuous shooting, as well as montage effects and artistic audio-to-video synchronization. A typical instance is a feature movie.

---

[551] *See id.*

[552] Yangshi Guoji Wangluo Youxian Gongsi Su Shanghai Juli Chuanmei Jishu Youxian Gongsi Zhuzuo Quan Qinquan Ji Buzhengdang Jingzheng Jiufen (央视国际网络有限公司诉上海聚力传媒技术有限公司著作权侵权及不正当竞争纠纷) [China Cent. Television Int'l Network Co. v. Shanghai SynaCast Media Tech. Co.], 2020 Hu Min Zhong No. 581 (2020 沪 73 民终 581 号) (Shanghai IP Ct. Apr. 20, 2021) [法宝引证码 CLI.C.410250170].

With respect to copyright authorship, China differentiates "cinematographic works and TV play works" from "other audiovisual works." The copyrights of cinematographic works and TV plays are vested in producers by law without any formality requirements.[553] In this situation, the screenwriters, directors, photographers, lyricists, composers, and other authors enjoy the right of signature and the right to receive remuneration according to their agreements with producers.[554] In addition, the authors of script, music, and other works that may be used separately from the motion pictures, can enjoy copyrights on their separable works apart from the whole motion pictures,[555] while the actors cannot claim property rights based on their performance in the motion pictures.[556] In contrast, the copyright ownership of "audiovisual works other than cinematographic works and TV play works" is dominated by agreements. If no agreement or the agreement is unclear, the producers enjoy the copyright, and the author has the right of signature and receives remuneration.[557]

3. Taiwan

Like the United States and China, the originality of work is required for copyright protection in Taiwan. Originality means: (1) the author must create the work originally and independently without copying others' works, and (2) the work must express the author's idea or emotion to a minimal degree of creativity. The degree of creativity does not have

---

[553] *See* Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment) § 17, I.

[554] *See id.* §17, I.

[555] *See id.*

[556] *See* Beijing Shi Gaoji Renmin Fayuan Qinhaiz Huzuoquan Anjian Shenli Zhinan (北京市高级人民法院侵害著作权案件审理指南) [Beijing High People's Court Guidelines for the Trial of Copyright Infringement Cases] (promulgated by the Beijing High People's Ct., Apr. 20, 2018) § 6.2 Film works and performer's right.

[557] *See* Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment) § 17, II.

to be unprecedented as long as sufficient perceivable differences from existing works can show the author's personality.[558] In terms of audiovisual works, they usually contain the director's shooting styles, actors' performances, costume and prop designs, lighting, and music in addition to plots. Therefore, different audiovisual works can be created based on the same plots and subjects.[559]

Similar to China,[560] the person whose name is represented on the work as the work's author is presumed to be the work's author in Taiwan.[561] Once an adversary party shows a counterproof, such as the existence of prior work, the author has the burden to prove his or her authorship, which includes: (1) whether he or she has the required ability, sufficient and reasonable time, and necessary supporting to create the work, and whether he or she can provide any documentations to prove the creative process; (2) the timing of completion; and (3) independent creation without copying.[562]

Unlike China, Taiwan does not distinguish motion pictures from other works in copyright ownership. A motion picture may be a joint work if three elements are met: (1) it is created by two or more people; (2) the parties intend to create a joint work; (3) it is a

---

[558] *See* Zhiguan Keji Gufen Youxian Gongsi Su Heluo Youxi Youxian Gongsi (智冠科技股份有限公司訴河洛遊戲有限公司) [Soft-World Int'l Co. v. Heluo Games Co.], 108 Min Zhu Shang Zi No. 3 (108 年度民著上字第 3 號) (Intell. Prop. Ct. July 30, 2010) (Taiwan).

[559] Yang Wenhui Weifan Zhuzuoquan Fa Anjian (楊文輝違反著作權法案件) [Republic of China v. Yang Wen Hui], 111 Tai Shang Zi No.2231 (111 年度台上字第 2231 號) (Sup. Ct., 2022) (Taiwan).

[560] *See* Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment) § 12, III.

[561] *See* Zhuzuoquan Fa (著作權法) [Copyright Act] (Republic of China, Taiwan) § 13.

[562] Rishang Wumu Shipin Gufen Youxian Gongsi Su Xinglin Shipin Gufen Youxian Gongsi (日商五木食品股份有限公司訴興霖食品股份有限公司) [Japanese Itsuki Foods Co. v. Sing-Lin Food Co.], 103 Tai Shang Zi No.1839 (103 年度台上字第 1839 號) (Sup. Ct. 2014) (Taiwan).

single and inseparable work.[563] The joint work's property rights shall not be exercised without all the coholders' consent, while a joint holder shall not refuse consent without a legitimate reason.[564] It is critical to note that a joint holder must contribute to a work's expression. For example, a commercial client may provide ideas, concepts, or opinions, and has the final say on the script. However, a commercial usually has more than one expression option in image construction, shooting angles, actor's movement and emotion performing, image editing, and music matching. Therefore, if the client does not involve the above expressions, he or she is not considered a joint work holder with the director, cinematographer, actors, and technicians.[565]

Besides joint work, a motion picture can be a work made within the scope of employment[566] or a work made for a commission agreement.[567] Except as otherwise specified in the agreement, the employer owns the property right for a work made within his or her employees' scope of employment, and the commissioned person owns the property right for a work made for a commission agreement, whereas the commissioning party (investor) has the right to exploit the work. Compared to the United States and China, the Taiwanese provision is unfavorable to the investor if he or she omits to acquire copyright in the commission agreement. To ease the severe consequence, the court

---

[563] Yao HsinAn (姚信安), *Cong Dianying Zhizuo Shiwu Lun Shiting Zhuzuo Quanli Guishu Yu Xingshi* (從電影製作實務論視聽著作權利歸屬與行使) [Attribution and Exercise of Copyrights of Audiovisual Works From Film Production Practices], *in* GENERAL REVIEW OF REPUBLIC OF CHINA'S COPYRIGHT LAW UNDER INTERNATIONAL COMPARISON Ⅰ (國際比較下我國著作權法之總檢討上冊) 259, 269 (2014).
[564] *See* Zhuzuoquan Fa (著作權法) [Copyright Act] (Taiwan) § 40-1.
[565] *See* Jian Xuanzong Su Taiwan Aodaili Gufen Youxian Gongsi (簡玄宗訴台灣奧黛莉股份有限公司) [Jian XuanZong v. Sincerity Found. Mfg. Co.], 110 Min Zhu Su Zi No.106 (110 年度民著訴字第 106 號) (Intell. Prop. & Com. Ct. June 14, 2022) (Taiwan).
[566] *See* Zhuzuoquan Fa (Taiwan) § 11.
[567] *See id.* § 12.

sometimes refers to related business customs to infer that the property right belongs to the investor despite no explicit clause.[568]

For example, In *Jian XuanZong v. Sincerity Found. Mfg. Co.*,[569] the defendant commissioned an advertising agency to make a product commercial. The advertising agency subcontracted the project to the plaintiff. To make the commercial, the plaintiff hired a producer and asked him to recruit the director, cinematographer, camera assistant, lighting technician, lighting assistant, design specialist, costume designer, and actors. The court held that the commercial was a work made for a commission agreement. However, the contract between the defendant and the advertising agency, the subcontract between the advertising agency and the plaintiff, and the production contract between the plaintiff and all the talents in the production team did not stipulate who shall enjoy the commercial's property rights. Despite this, based on the above contracts' specific purposes and the producer's general understanding of the business, the court held that the plaintiff enjoyed the property right and had transferred the right to the advertising agency when he delivered the commercial to the advertising agency. For the same reason, the advertising agency had transferred the property right to the defendant when it delivered the commercial to the defendant.

B. THE HUMAN AUTHORSHIP REQUIREMENT

1. United States

Can a machine be an author? If not, who is the author of machine-generated work?

---

[568] *See Jian XuanZong*, 110 Min Zhu Su Zi No.106.
[569] *See id.*

These controversies are long disputed because "neither the Constitution nor the Copyright Act indicates that authors must be human."[570] For these questions, Professor Leaffer has three findings. First, the universal assumption is "the copyright monopoly was intended for human authors."[571] Second, "the raw output of a computer-generated work should be treated as fact, without ownership rights in anybody."[572] Third, "copyright protection should be conferred on the user of the program, not the programmer, but only when the user has added original authorship to the raw output."[573]

Citing the words "the fruits of intellectual labor" and "creative powers of the mind" in *re Trade-Mark Cases*,[574] as well as "original intellectual conceptions of the author" in *Burrow-Giles Lithographic Co* decision,[575] United States Copyright Regulations explicitly requires "human authorship" for copyright registrations.[576] Identical to other types of works, Copyright Regulations prescribes, "motion picture must contain creative human authorship. A motion picture created by a non-human author, created by a purely mechanical process, or generated solely by preexisting software is not copyrightable."[577] The United States Copyright Office (USCO) "will not register works produced by a

---

[570] LEAFFER, *supra* note 80, at 108.
[571] *Id.*
[572] *Id.* at 109.
[573] *Id*; *see also* Design Data Corp. v. Unigate Enter., Inc., 847 F.3d 1169, 1173 (9th Cir. 2017) (stating two opinions among authorities: (a) rejecting copyright could extend to the program's output, and (b) suggesting a computer program copyright "may extend to the program's output if the program 'does the lion's share of the work' in creating the output and the user's role is so 'marginal' that the output reflects the program's contents"); *see also* Huang Chieh, *supra* note 80.
https://www.repository.law.indiana.edu/etd/107
[574] In re Trade-Mark Cases, 100 U.S. 82, 94 (1879).
[575] Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 58 (1884).
[576] COMPENDIUM (THIRD) § 306.
[577] *Id.* § 808.7(C).

machine or mere mechanical process that operates randomly or automatically without any creative input or intervention from a human author."[578]

The following two instances reflect the USCO's consistent position that machine-generated works are not copyrightable:

a. *A Recent Entrance to Paradise*

A two-dimensional artwork titled *A Recent Entrance to Paradise* ("Work") was filed to register a copyright by Stephen Thaler on November 3, 2018.[579] In the application, Thaler noted the work "was autonomously created by a computer algorithm running on a machine" and sought to register the work as "a work made for hire" to him, the machine owner.[580] Because Thaler asserted the work lacked human authorship, the USCO refused his copyright registration (August 12, 2019), his reconsideration request (March 30, 2020), and his second reconsideration request (February 14, 2022).

The USCO disagreed with Thaler's argument that case law did not support the human authorship requirement.[581] In its second reconsideration refusal opinion, the USCO reiterated to follow *Burrow-Giles Lithographic* decision, which described copyright as "the exclusive right of a man to the production of his own genius or intellect,"[582] the "person" who is the cause of the picture which is produced, and "the man" who created or gave effect to the idea. [583] The USCO also cited various lower court cases that rejected the

---

[578] *Id.* § 313.2.
[579] Steven Thaler, Correspondence ID 1-3ZPC6C3; SR # 1-7100387071 (Copyright Rev. Bd. Feb. 14, 2022).
[580] *Id.*
[581] *Id.*
[582] Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 58 (1884).
[583] *Id.* at 60-61.

copyrightability of a book made by non-human spiritual beings,[584] photos taken by a monkey,[585] and the living garden or jellyfish that were created by nature.[586]

Furthermore, the USCO's second reconsideration refusal opinion cited three critical government documents to strengthen its position. First, the National Commission on New Technological Uses of Copyrighted Works Final Report (1978) stated, "the eligibility of any work for protection by copyright depends . . . . upon the presence of at least minimal human creative effort at the time the work is produced."[587] Second, the Register of Copyrights annual report (1965) concluded, "The crucial question appears to be whether the 'work' is basically one of human authorship, with the computer merely being an assisting instrument, or whether the traditional element of authorship in the work (literary, artistic or musical expression or elements of selection, arrangements, etc.) were actually conceived and executed not by man but by a machine."[588] The last document was the USPTO Public Views on Artificial Intelligence and Intellectual Property Policy (2020) stated, "the vast majority of commenters acknowledged that existing law does not permit a non-human to be an author [and] this should remain the law."[589]

Finally, the USCO refuted Thaler's interpretation that a non-human could be an author under the work-made-for-hire doctrine. First, machines could not form employment

---

[584] Urantia Found. v. Kristen Maaherra, 114 F.3d 955, 957–59 (9th Cir. 1997).

[585] Naruto v. Slater, 888 F.3d 418, 426 (9th Cir. 2018).

[586] Kelley v. Chicago Park Dist., 635 F.3d 290, 304 (7th Cir. 2011), Satava v. Lowry, 323 F.3d 805, 813 (9th Cir. 2003).

[587] NATIONAL COMMISSION ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS ("CONTU"), FINAL REPORT at 45 (1978).

[588] U.S. COPYRIGHT OFFICE, SIXTY-EIGHTH ANNUAL REPORT OF THE REGISTER OF COPYRIGHTS FOR THE FISCAL YEAR ENDING JUNE 30, 1965, at 5 (1966).

[589] U.S. PATENT AND TRADEMARK OFFICE, PUBLIC VIEWS ON ARTIFICIAL INTELLIGENCE AND INTELLECTUAL PROPERTY POLICY AT 20-21 (2020), https://www.uspto.gov/sites/default/files/documents/USPTO_AI-Report_2020-10-07.pdf.

agreements or work-for-hire agreements with humans. Second, a work made for hire did not remove the human authorship requirement to be copyrightable. The USCO also indicated that Thaler's policy argument had no grounds under the current copyright law.

The USCO's denial of Thaler's copyright registration was affirmed by the United States District Court for the District of Columbia on August 18, 2023. Regarding "whether a work generated autonomously by a computer system is eligible for copyright," Judge Beryl A. Howell wrote in his memorandum opinion, "In the absence of any human involvement in the creation of the work, the clear and straightforward answer is the one given by the Register: No."[590]

b. *Zarya of the Dawn*

On September 15, 2022, Kristina Kashtanova registered copyright for a visual material titled *Zarya of the Dawn* ("Work") in the USCO.[591] On September 20, she announced on Instagram, "I got Copyright from the Copyright Office of the USA on my Ai-generated graphic novel."[592] The graphic novel was reported widely as the "first known US copyright registration for latent diffusion AI art."[593] Soon, Kashtanova received a USCO notice asking her to "explain in detail exactly how the Work was created."[594]

In her response of November 21, 2022 to the USCO,[595] Kashtanova distinguished

---

[590] Thaler v. Perlmutter, No. CV 22-1564 (BAH), 2023 WL 5333236, at *6 (D.D.C. Aug. 18, 2023).
[591] Zarya of the Dawn, Registration Number VAu001480196, https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?v1=1&ti=1,1&Search%5FArg=VAu001480196&Search%5FCode=REGS&CNT=25&PID=QYVn4FCjDLW8nt-grR26irPWTWWP&SEQ=20230819142409&SID=1.
[592] Kris.Kashtanova, Instagram (Sept. 20, 2022), https://www.instagram.com/p/CivS3iiPigt/ (last visited Aug. 20, 2022).
[593] Benj Edwards, *Artist receives first known US copyright registration for latent diffusion AI art*, ARSTECHNICA (Sept. 22, 2022), https://arstechnica.com/information-technology/2022/09/artist-receives-first-known-us-copyright-registration-for-generative-ai-art/ [https://perma.cc/SBD3-GZAF].
[594] Kristina Kashtanova, Correspondence ID: 1-5GB561K (U.S. Copyright Off. Oct. 28, 2022).
[595] *Id.*

her Work from "A Recent Entrance to Paradise." She asserted Midjourney did not generate images randomly and automatically but was directed by her. Her creative input was either a "prompt" in English or one or multiple pre-existing images. She also designated the image's working part, size, aspect ratio, and generative process. Her prompt might be as short as "dark skin hands holding an old photograph –ar 16:9," or as long as "sci-fi scene future empty New York, Zendaya leaving gates of Central Park and walking towards an empty city, no people, tall trees, New York Skyline forest punk, crepuscular rays, epic scene, hyper realistic, photo realistic, overgrowth, cinematic atmosphere, ethereal lighting."[596] The input for the image "Raya as a Hologram" were two preexisting images, a description of the scene, mood, style-related direction, and output constraints. Kashtanova painstakingly picked Midjourney's output images and iteratively rearranged her input until a satisfactory image was done. She needed to try hundreds of times to make a final image.

USCO recognized Kashtanova's authorship and copyrightability of the Work's text as well as the selection, coordination, and arrangement of the Work's written and visual elements. However, it rejected the human authorship and copyrightability of the individual Midjourney-generated images.[597] Accordingly, USCO canceled Kashtanova's previous registration and reissued a new, more limited registration. USCO denied Kashtanova as the Midjourney-generated images' author based on Midjourney users' inability to predict a particular output image by a particular prompt. USCO analogized the Midjourney user's "prompts function closer to suggestions than orders, similar to the situation of a client who

---

[596] *Id.*

[597] Van Lindberg on behalf of Kristina Kashtanova, Related Correspondence ID: 1-5GB561K (U.S. Copyright Off. Feb. 21, 2023).

hires an artist to create an image with general directions as to its contents."[598] In this situation, the image's author is the artist, not the client. Therefore, no matter how much time, effort, or expense it took Kashtanova to create the work, "sweat of the brow" cannot justify copyright protection for otherwise unprotectable material.

c. USCO AI-generated Work Registration Guidance

To provide guidance for registering works that contain AI-generated material, USCO issued a Copyright Registration Guidance that took effect on March 16, 2023.[599] This guidance reiterates statutory and judicial grounds of the human authorship requirement. A "case-by-case inquiry" is necessary for determining whether a work containing AI-generated material is eligible to claim copyright and "what matters is the extent to which human had creative control over the work's expression and 'actually formed' the traditional elements of authorship."[600]

For the purpose of identifying a work and its existence, ownership, or duration of the copyright,[601] USCO asks the applicant to provide a brief statement in the following fields of the Standard Application:

(1) "Author Created" field:

    (a) For a work incorporating AI generated materials into a larger work, stating which portions of the work is human authored;

---

[598] *Id.*
[599] U.S. COPYRIGHT OFFICE, 37 CFR PART 202 COPYRIGHT REGISTRATION GUIDANCE: WORKS CONTAINING MATERIAL GENERATED BY ARTIFICIAL INTELLIGENCE, 88 Fed. Reg. 16190 (Mar. 16, 2023).
[600] *Id.* at 16193.
[601] 17 U.S.C. § 409(10).

(b) For a work creatively arranges the human and non-human content within a work, stating ''selection, coordination, and arrangement of [describe human-authored content] created by the author and [describe AI content] generated by artificial intelligence;''

(c) AI technology or the company simply using it to create the work should not be listed as an author or co-author.

(2) ''Material Excluded/Other'' field under the ''Limitation of the Claim'' section: stating ''[description of content] generated by artificial intelligence.''

(3) ''Note to CO'' field: Provide additional information.

The above statement requirement applies to applications that are pending or have been registered. The applicant should contact the USCO to disclose the information. USCO may cancel the registration if it becomes aware of the omission.

2. China

Like the United States, human authorship is required to obtain copyright protection in China. Article 2 of Regulation for the Implementation of the Copyright Law of the People's Republic of China provides, "The term 'works' as referred to in the Copyright Law means intellectual creations with originality in the literary, artistic or scientific domain, insofar as they can be reproduced in a tangible form."[602] Accordingly, some courts stress

---

[602] Zhonghua Renmin Gongheguo Zhuzuoquan Fa Shishi Tiaoli (中华人民共和国著作权法实施条例) [Regulation for the Implementation of the Copyright Law of the People's Republic of China] (2013 Revision) § 2.

that human authorship is necessary to obtain copyright.[603]

a. *Tencent v. Yingxun Tech.*

In *Shenzhen Tencent Comput. Sys. Co. v. Shanghai Yingxun Tech. Co.*,[604] the defendant's website posted a financial article that was first published on the plaintiff's website. This article comprised 979 Chinese characters (title included) and nine paragraphs, in which the plaintiff marked, "Tencent Robot Dreamwriter automatically wrote this article." The plaintiff developed and owned the computer program copyright of Dreamwriter that produced this article by collecting and analyzing data from a database. Dreamwriter also proofread and edited this article as well as published it on the relevant platform.

The article's generating process took two minutes and had no human involvement. However, the court found the Dreamwriter development team's intellectual activities, including the arrangement and selection of the data input, trigger conditions setting, template, and corpus style, were directly related to this article's specific expression. Such individual arrangement and selection had originality because there was more than one expressive option. Despite the time interval between human intellectual activities and this article's production, they shall be considered a whole creative process. Therefore, the court

---

[603] Zhou Bo (周波), *Rengong Zhineng Yu Zhuzuoquan Baohu — Zhongguo Fayuan De Sifa Shijian* (人工智能与著作权保护 —中国法院的司法实践) [Artificial Intelligence and Copyright Protection —Chinese Courts' Judicial Practice], https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/conversation_ip_ai/pdf/ms_china_1_zh.pdf [https://perma.cc/42R2-YVRG].

[604] Shenzun Shi Tengxun Jisuanji Xitong Youxian Gongsi Su Shanghai Yingxun Keji Youxian Gongsi Qinhai Zhu Zuoquan Ji Bu Hengdang Jingzheng Jiufen (深圳市腾讯计算机系统有限公司诉上海盈讯科技有限公司侵害著作权及不正当竞争纠纷) [Shenzhen Tencent Comput. Sys. Co. v. Shanghai Yingxun Tech. Co.], 2019 Yue 0305 Min Chu No. 14010 (2019 粤 0305 民初 14010 号) (Shenzhen Nanshan Dist. People's Ct. Dec. 24, 2019) [法宝引证码 CLI.C.97241331].

held this article was an original intellectual creation and was copyrightable.

One commentator comments that the *Tencent* case is related to AI-assisted work instead of AI-generated work.[605] There is always some human involvement in AI-assisted works so that a human author can be found easily. The more difficult question is the copyrightability of an AI-generated work that does not have any human involvement. I think *Tencent's* holding is close to some American courts' "lion's share test" that extends the programmer's copyright to the program's automatic output.[606]

b. *Gao Yang et al. v. Golden Vision*

The issue of *Gao Yang v. Golden Vision (Beijing) Film and Television Culture Co.* regarded whether the pictures extracted from a video that was automatically recorded by a camera carried by a flying hot air balloon were copyrightable.[607] The district court held the pictures extracted from the video were uncopyrightable because the sequential images shot by the balloon's camera were a non-original video recording, which had no human intellectual selection of shooting elements like objects, timing, angle, distance, light, color, and brightness after the balloon released to the sky.

However, Beijing Intellectual Property Court disagreed, stating the pictures were original because there was sufficient human intellectual involvement in making the video, selecting and improving the pictures. The human factors during the video shooting process

---

[605] Zho, *supra* note 603.

[606] *See* Design Data Corp. v. Unigate Enter., Inc., 847 F.3d 1169, 1173 (9th Cir. 2017).

[607] Gaoyang Su Jinse Shizu (Beijing) Yingshi Wenhua Youxian Gongsi Zhuzuoquan Quanshu Qinquan Jiufen (高阳诉金色视族（北京）影视文化有限公司著作权权属、侵权纠纷) [Gao Yang v. Golden Vision (Beijing) Film and Television Culture Co.], 2017 Jing 73 Min Zhong No. 797 (2017 京 73 民终 797 号) (Beijing IP Ct. Apr. 2, 2020) [法宝引证码 CLI.C.97510788]; *see also* Chen Yanru, *Does China Back Copyrights for Automatic Photos from a Hot-Air Balloon?*, CHINA JUST. OBSERVER (Nov. 22, 2020), https://www.chinajusticeobserver.com/a/does-china-back-copyrights-for-automatic-photos-from-a-hot-air-balloon [https://perma.cc/CD8N-NEFU].

mainly included: (1) Shooting purpose: taking pictures of the earth through high-altitude balloons; (2) Subject: the earth; (3) Technique: using balloons to carry a camera in consideration of balloon burst, weather, wind speed, etc.; (4) Equipment: GoPro HERO2 camera; (5) Angle: upside-down shooting; (6) Settings: video recording mode, namely 1080P, 25 frames per second, wide-angle, and sensitivity of 800.[608]

Though the *Gao Yang* case is related to photographic works, the disputed picture is extracted from a video. The court confirms pictures automatically taken by a camera can be original as long as the shooting process presents apparent human involvement, selection, and judgment, even though the author is not on site. For the same reason, a video automatically shot by a camera can be a copyrightable audiovisual work if human factors are present during the shooting process.

3. Taiwan

Like the United States and China, human authorship is required to obtain copyright in Taiwan. Article 3, item 2 of the Republic of China (Taiwan) Copyright Law provides, "Author means a person who creates a work."[609] Also, Article 10 of Taiwan Copyright Law provides that, "the author of a work shall enjoy copyright upon completion of the work."[610] Taiwan Supreme Court has held that a copyrightable work shall be a spiritual creation of human beings and sufficient to reflect the author's personality and uniqueness.[611]

---

[608] Chen Yanru, *Does China Back Copyrights for Automatic Photos from a Hot-Air Balloon?*, CHINA JUST. OBSERVER (Nov. 22, 2020), https://www.chinajusticeobserver.com/a/does-china-back-copyrights-for-automatic-photos-from-a-hot-air-balloon [https://perma.cc/CD8N-NEFU].

[609] *See* Zhuzuoquan Fa (著作權法) [Copyright Act] (Taiwan) § 3(2).

[610] *See id.* § 10.

[611] 97 Tai Shang Zi No. 3121 (97 年台上字第 3121 號) (Sup. Ct. 2008) (Taiwan).

With regard to AI-related works, the Taiwan Intellectual Property Office divides them into two categories:[612] (1) AI-assisted work, which means AI is only used as an assisting tool to complete the work (e.g., a drawing software), is copyrightable and the copyright shall be conferred to the human who contributes the creative inputs; (2) AI independently generated work, which means the work is made by algorithms automatically without human spirit civilization inputs, is uncopyrightable because AI is not human. The following three cases show the boundary between AI-assisted work and AI-independently generated work is arbitrary.

a. *Prudential Life Ins. Co. v. Metro. Life Ins. Co.*

In *Prudential Life Ins. Co. v. Metro. Life Ins. Co.*,[613] the disputed work was a print advertisement that contained various computer analysis charts of several necessary expenses. The plaintiff claimed those computer charts were made by its originally and independently selected and classified parameters. The court held that numerical charts were not copyrightable subject matters pursuant to Article 9 paragraph 1 item 3 of Taiwan Copyright Law.[614] Moreover, the charts were made automatically by computers according to parameter inputs. Since those charts were mathematical calculation results, the court concluded they were not human creations and non-copyrightable subject matters.

---

[612] Taiwan Intellectual Property Office, No. 1111031 Email (Oct. 31, 2022).

[613] Baodexin Guoji Renshou Baoxian Gufen Youxian Gongsi Su Daduhui Guoji Renshou Baoxian Gufen Youxian Gongsi (保德信國際人壽保險股份有限公司訴大都會國際人壽保險股份有限公司) [Prudential Life Ins. Co. v. Metro. Life Ins. Co.], 98 Ming Zu Shang Zi No.16 (98 年度民著上字第 16 號) (Intell. Prop. Ct. Feb. 4, 2010) (Taiwan).

[614] *See* Zhuzuoquan Fa (著作權法) [Copyright Act] (Taiwan) § 9, I, (3): The following items shall not be the subject matter of copyright: . . . 3. Slogans and common symbols, terms, formulas, numerical charts, forms, notebooks, or almanacs.

b. *Fu, Xian Jin case*

In *Fu Xian Jin* case, [615] the disputed works were 59 line-graphs of machine spare parts created by using AutoCAD software. The line graphs were drawn one line by one line. Different operators' works may vary depending on their skills, carefulness, orders given to the software, and extra emphasis and details, even if their subjects were identical machine parts. The court held that the line graphs of machine parts were copyrightable works because they presented the author's minimum originality.

c. *Ding, Fu Ting case*

In *Ding Fu Ting* case,[616] the disputed works were three interior design 3D rendering pictures created by using unnamed software. To draw 3D rendering pictures, the user had to set up the parameters and options such as the materials, season, rays, lighting. The user can only see the 3D effect, the degree of shading, whether the finished product is beautiful or not, etc., after running the program for a long time. Different users' works may vary depending on their design expertise and aesthetic literacy, even if the 3D rendering pictures were derived from the same 2D interior design pictures. The court held that the 3D rendering pictures were copyrightable works because they had presented the author's personality.

Without considering the *Prudential Life* case that is related to uncopyrightable numerical charts, the Taiwanese court's approach is close to the Chinese Court's opinion

---

[615] Fu Xianjin Weifan Zhuzuoquan Fa Anjian (傅先進違反著作權法案件) [Republic of China v. Fu Xian Jin], 107 Xing Zhi Shang Su Zi No.1 (107 年度刑智上訴字第 1 號) (Intel. Prop. Ct. Dec. 27, 2018) (Taiwan).

[616] Ding Fu Ting Weifan Zhuzuoquan Fa Anjian (丁富庭違反著作權法案件) [Republic of China v. Ding Fu Ting], 104 Xing Zhi Shang Yi Zi No. 95 (104 年刑智上易字第 95) (Intel. Prop. Ct. Mar. 10, 2016) (Taiwan).

that admits minimum human intellectual involvement during the creation process would be sufficient to establish human authorship if the finished work is a copyrightable subject matter, whereas the USCO cares more about the author's direct involvement to the final work's expression and whether the author can foresee the results.

C. FIXATION REQUIREMENT

1. United States

Motion pictures must be fixed in video files, videotape, film, or any tangible medium of expression to satisfy the fixation requirement.[617] A live broadcast not fixed on a tangible medium is certainly not copyrightable.[618] In addition, whoever makes the fixation is essential to determine the motion picture's authorship as "fixation must be done 'by or under the authority of the author.'"[619] In *Garcia v. Google, Inc.*,[620] the Ninth Circuit denies Garcia, an actor who performed for five seconds in the film *Innocence of Muslims*, to be the film's author because she is not "the person who translates an idea into a fixed, tangible expression entitled to copyright protection."[621]

2. China

Unlike the United States, there is no fixation requirement in China's copyright law. However, a work must be presented in a certain form to be copyrighted.[622] For a motion

---

[617] COMPENDIUM (THIRD) § 808.6.
[618] *See id.* § 313.
[619] Garcia v. Google, Inc., 786 F.3d 733, 741 (9th Cir. 2015).
[620] *See id.*
[621] *See id.* at 744.
[622] *See* Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment) § 3.

picture, it must be recorded on some material.[623] In *Beijing SINA Internet Info. Serv. Co. v. Beijing Tianying Jiuzhou Network Tech. Co.*,[624] Beijing Intellectual Property Court held that a live sporting event stream over the internet was not entitled to copyright because the work was not fixed. However, the Beijing High People's Court overturned such a narrow opinion, holding that copyright law did not require fixing the signal on a tangible medium. It would be enough to store the signal on a certain material independent from the human mind if it can be reproduced.

3. Taiwan

Unlike the United States, there is no fixation requirement in Taiwan's copyright law. However, the Illustrated Contents of Each Kind of Works in Paragraph One, Article 5 of the Copyright Act (Taiwan) § (7) provides: "Audio-visual works: shall include the images shown in any motion picture, videocassette, videodisk, on computer screen, and the other series of images which could be fixed to any medium with or without sound to be shown by a mechanical device or equipment."[625] Therefore, "fixed to any medium" is necessary to be protected as an audiovisual work.[626]

---

[623] Zhonghua Renmin Gongheguo Zhuzuoquan Fa Shishi Tiaoli (中华人民共和国著作权法实施条例) [Regulation for the Implementation of the Copyright Law of the People's Republic of China] (2013 Revision) § 4(11).

[624] *See* Beijing Xinlang Hulian Xinxi Fuwu Youxian Gongsi Yu Beijing Tianying Jiuzhou Wangluo Jishu Youxian Gongsi Buzhengdang Jingzheng Jiufen (北京新浪互联信息服务有限公司与北京天盈九州网络技术有限公司不正当竞争纠纷) [Beijing SINA Internet Info. Serv. Co. v. Beijing Tianying Jiuzhou Network Tech. Co.], 2020 Jing Min Zai No. 28 (2020 京民再 128 号) (Beijing Mun. High People's Ct. Sept. 23, 2020) [法宝引证码 CLI.C.115093220].

[625] Zhuzuoquan Fa Diwutiao Diyixiang Gekuan Zhuzuo Neirong Liqi (著作權法第五條第一項各款著作內容例示) [The Illustrated Contents of Each Kind of Works in Paragraph One, Article 5 of the Copyright Act] (Taiwan) § 2(7).

[626] CHANG YIYUN (張懿云), RESEARCH ON THE COPYRIGHT PROTECTION OF AUDIOVISUAL WORKS FINAL REPORT (視聽著作權利保護之研究期末報告) 186, (Taiwan Intell. Prop. Off. 2011).

II. MOTION PICTURE'S SCOPE OF COPYRIGHT

A. PROTECTABLE ELEMENTS V. NON-PROTECTABLE ELEMENTS

1. United States

The separation of protectable elements and non-protectable elements of a copyrighted work is important because "[t]he mere fact that a work is copyrighted does not mean that every element of the work may be protected."[627] In *Harper & Row Publishers, Inc. v. Nation Enterprises*, the Supreme Court emphasizes, "the First Amendment protections already embodied in the Copyright Act's distinction between copyrightable expression and uncopyrightable facts and ideas."[628] Therefore, "copyright does not prevent subsequent users from copying from a prior author's work those constituent elements that are not original—for example, quotations borrowed under the rubric of fair use from other copyrighted works, facts, or materials in the public domain."[629] In *Feist Publications, Inc. v. Rural Telephone Service Co.*, the Supreme Court reiterates, "The most fundamental axiom of copyright law is that '[n]o author may copyright his ideas or the facts he narrates.'"[630]

Section 102(b) of the Copyright Act provides that copyright protection does not extend to "any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work."[631] Furthermore, the USCO's Copyright Regulations specify six

---

[627] Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 111 S. Ct. 1282, 1289 (1991).
[628] Harper & Row Publishers, Inc. v. Nation Enterprises, 471 U.S. 539, 560 (1985).
[629] *Id.* at 548.
[630] *Feist Publications, Inc.*, 499 U.S. at 344-345.
[631] 17 U.S.C. § 102(b).

types of uncopyrightable materials:[632] (1) works that have not been fixed, (2) works that lack human authorship, (3) works that do not constitute copyrightable subject matter,[633] (4) works that do not satisfy the originality requirement,[634] (5) specific types of works that may contain uncopyrightable material (e.g. numbers, research, and book designs for literary works), and (6) other types of works that cannot be registered with the USCO (foreign works that are not eligible for copyright protection in the United States; unlawful use of preexisting material in a derivative work, a compilation, or a collective work; government works; government edicts; works in the public domain).

As discussed above, "copyright protections 'extend[ ] only to those components of a work that are original to the author.'"[635] Therefore, in a copyright infringement case, "a court must filter out and disregard the non-protectible elements in making its substantial similarity determination."[636] "[T]his filtration serves 'the purpose of defining the scope of plaintiff's copyright'" [637] and "may ultimately leave behind a 'core of protectable material.'"[638] For example, in *Axelbank v. Rony*,[639] the plaintiff claimed the defendant infringed his copyrighted film, *Tsar to Lenin*, a compilation of the plaintiff's lifelong collection of 1917 Russian Revolution documentary films. By finding the documentary films used in both side's films are in the public domain, the Ninth Circuit found that the

---

[632] COMPENDIUM (THIRD) § 313.

[633] This covers ideas, procedures, processes, systems, methods of operation, concepts, principles, or discoveries; merger of idea and expression; facts; typeface and mere variations of typographic ornamentation; format and layout.

[634] This covers mere copies; de minimis authorship; words and short phrases; works consisting entirely of information that is common property; measuring and computing devices; mere listing of ingredients or contents; blank forms; the name or general idea for a character; scènes à faire; familiar symbols and designs; mere variations of coloring.

[635] Montgomery v. NBC Television, 833 F. App'x 361, 364 (2d Cir. 2020).

[636] Cavalier v. Random House, Inc., 297 F.3d 815, 822 (9th Cir. 2002).

[637] Computer Assocs. Int'l, Inc. v. Altai, Inc., 982 F.2d 693, 707 (2d Cir. 1992).

[638] *Id.*

[639] Axelbank v. Rony, 277 F.2d 314 (9th Cir. 1960).

plaintiff's copyright on *Tsar to Lenin* extended only to "the sequential development, the commentary of Max Eastman, and one map for which [plaintiff] claims credit."[640] As the defendant did not copy the sequential development or the commentary of *Tsar to Lenin*, and the map's copyright protection was limited, the court held the defendant did not infringe on the plaintiff's copyrighted film.

Similarly, in *Idema v. Dreamworks Inc.*,[641] the plaintiffs claimed the defendants' movie *Peacemaker* infringed their copyright of eight works collectively called *Idema's Story*, purported to be actual events of the plaintiff Idema's life. The Ninth Circuit held that many similarities between *Idema's Story* and *Peacemaker* were unprotectable historical facts, stock characters, and *scènes à faire*. As a result, the district court's summary judgment for defendants was affirmed.

2. China

Similar to the United States, ideas, raw materials, information in public domains, creation formats, necessary scenes, or unique or limited expression forms are excluded from copyright protection.[642] In a copyright infringement case, the alleged copyright infringer can assert the following defenses against substantial similarity: (1) Limited expression: the form of expression is very limited;[643] (2) Necessary scene: a scene must be

---

[640] *Id.* at 317.

[641] Idema v. Dreamworks, Inc., 90 F. App'x 496, 498 (9th Cir. 2003), as amended on denial of reh'g (Mar. 9, 2004).

[642] Zhang Xiaoyan Su Lei Xianhe Zhaoqi Sanding Aishuren Yinxiang Tushu Youxian Gongsi (张晓燕诉 雷献和、赵琪、山东爱书人音像图书有限公司) [Zhang Xiaoyan v. Lei Xianhe, Zhao Qi, and Shandong Aishuren DVDs & Books Co., Ltd.], 2013 Min Shen Zi No. 1049 (2013 民申字第 1049 号) (Sup. People's Ct. 2014).

[643] *See* Beijing Shi Gaoji Renmin Fayuan Qinhaiz Huzuoquan Anjian Shenli Zhinan (北京市高级人民法院 侵害著作权案件审理指南) [Beijing High People's Court Guidelines for the Trial of Copyright Infringement Cases] (promulgated by the Beijing High People's Ct., Apr. 20, 2018) § 7.3.

described, or a scene design must be used in expressing a theme;[644] (3) News on current events;[645] (4) Public domain.[646] In addition, the courts do not consider themes, inspirations, or emotions that belong to the category of idea in determining substantial similarity.[647] The mainline and sequence of historical facts based on the same historic theme also fall in the category of idea.[648]

3. Taiwan

Similar to the United States and China, Taiwan also excludes unprotected materials from copyright protection, such as ideas, [649] news on current events, [650] natural phenomena,[651] historical research,[652] limited expressions,[653] *Scènes à Faire*.[654]

With respect to derivative work, the court will examine the incremental parts the latter work adds to the prior work to ascertain the derivative work's scope of copyright.

---

[644] *Id. at* §7.4.

[645] *Id.* at §7.5.

[646] *Id.* at §7.6.

[647] *Id.* at §10.10.

[648] *Id.* at §10.11.

[649] *See* Zhuzuoquan Fa (著作權法) [Copyright Act] (Taiwan) § 10-1.

[650] *See id.* § 9, I(4).

[651] *See* Zhu Peixuan Weifan Zhuzuoquan Fa Anjian (朱珮萱違反著作權法案件) [The Republic of China v. Zhu Pei Xuan], 108 Xing Zhi Shang Su Zi No. 41 (108 年度刑智上訴字第 41 號) (Intell. Prop. Ct. Nov. 14, 2019) (Taiwan).

[652] *See* Huang Pingying Su You Shujun (黃萍瑛訴游淑珺) [Huang Ping Ying v. You Shu Jun], 104 Min Zhu Su No.39 (104 年度民著訴字第 39 號) (Intell. Prop. Ct. Sept. 30, 2016) (Taiwan).

[653] *See* Wang Taichang Su Meishang Maigeluo Xier Guoji Gufen Youxian Gongsi Taiwan Fengongsi (王泰昌訴美商麥格羅．希爾國際股份有限公司台灣分公司) [Wang Tai Chang v. McGraw-Hill Education Co., Taiwan Branch], 104 Tai Shang Zi No.1251 (104 年度台上字第 1251 號) (Sup. Ct. 2015) (Taiwan). This case is related to copyright infringement of an accounting textbook.

[654] *See* Wangyin Guoji Gufen Youxian Gongsi Su Yinhe Xianshang Gufen Youxian Gongsi (網銀國際股份有限公司訴銀河線上股份有限公司) (Wanin Int'l Co., v. Galaxy Co.), 106 Min Zhu Su Zi No. 5 (106 年度民著訴字第 5 號) (Intell. Prop. Ct. Jan. 25, 2018) (Taiwan). This case is related to copyright infringement of an online video game.

For example, in *Endemol Nederland B.V. v. All Full Power Co.*,[655] the court held that the American version *1 vs. 100* TV show originated from a prior Dutch version *1 vs. 100* TV show except for the three types of help: Ask the Mob, Trust the Mob, and Poll the Mob. However, the three helps originated from Endemol Nederland's earlier format and was found in Endemol Nederland's *1 vs. 100* production manuals. An earlier Taiwanese TV show *Who Wants to be a Millionaire* also used a help like Ask the Mob. Therefore, the court held the American version *1 vs. 100* TV show lacked originality for the show's format and could not prevent its competitor from using a similar show format.

B. IDEA V. EXPRESSION

1. United States

The most difficult copyright legal issue is the idea-expression dichotomy, which is often used by the courts in individual cases to adjust the tension between the original authors' exclusive benefits and the prospective authors' creative freedom[656] as well as the public benefits for subsequent improvement and progress of the same subject matter.[657] Examples of ideas include: "a quarrel between a Jewish father and an Irish father, the marriage of their children, the birth of grandchildren, and a reconciliation;"[658] "an idealistic young professional choosing between financial and emotional reward, or of love triangles among young professionals that eventually become strained, or of political forces

---

[655] *See* Helan Shang Endemol Nederland B.V. Zisu Quanneng Zhizuo Gufen Youxian Gongsi (荷蘭商 Endemol Nederland B.V.自訴全能製作股份有限公司) [Endemol Nederland B.V. v. All Full Power Co.], 103 Xing Zhi Shang Yi No. 56 (103 年度刑智上易字第 56 號) (Intell. Prop. Sept. 3, 2015) (Taiwan).
[656] Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231, 240 (2d Cir. 1983).
[657] Reyher v. Children's Television Workshop, 533 F.2d 87 (2d Cir. 1976).
[658] Nichols v. Universal Pictures Corp., 45 F.2d 119, 120 (2d Cir. 1930).

interfering with private action;"[659] "revealing the secrets behind magic illusions and tricks;"[660] or "animated, anthropomorphic car characters."[661]

At the two ends of ideas and expressions, the early courts are prone to ideas (a narrow copyright protection), whereas the later courts are prone to expressions (a broad copyright protection).[662] Several tests have been proposed to distinguish the idea and expression in historical development.

a. Literal Test

This simplistic test had been applied in early cases.[663] In *Stowe v. Thomas*, the Court held that the defendant did not infringe the copyright of the plaintiff's English book by translating it into Germany because he did not copy "a transcript of the language in which the conceptions of the author are clothed."[664] The Court reasoned:

> An author may be said to be the creator or inventor, both of the ideas contained in his book, and the combination of words to represent them . . . . His exclusive property in the creation of his mind, cannot be vested in the author as abstractions, but only in the concrete form which he has given them, and the language in which he has clothed them. When he has sold his book, the only property which he reserves to himself, or which the law gives to him, is the exclusive right to multiply the copies of that particular combination of characters which exhibits to the eyes of another the ideas intended to be conveyed. This is what the law terms copy, or copyright.[665]

Apparently, the *Stowe* Court disavowed that copyright could extend to more than the work's literal expression. However, this simplistic test was rejected by the later

---

[659] Metcalf v. Bochco, 294 F.3d 1069, 1074 (9th Cir. 2002).
[660] Rice v. Fox Broad. Co., 330 F.3d 1170, 1174 (9th Cir. 2003).
[661] Esplanade Prods., Inc. v. Walt Disney Co., No. CV1702185MWFJCX, 2017 WL 5635024, at *8 (C.D. Cal. July 11, 2017).
[662] Edward Samuels, *The Idea-Expression Dichotomy in Copyright Law*, 56 TENN. L. REV. 321, 355 (1989).
[663] *Id.* at 336.
[664] Stowe v. Thomas, 23 F. Cas. 201 (C.C.E.D. Pa. 1853).
[665] *Id.*

131

precedents.[666] In *Nichols v. Universal Pictures Corp.*, the court held, "It is of course essential to any protection of literary property, whether at common-law or under the statute, that the right cannot be limited literally to the text, else a plagiarist would escape by immaterial variations."[667] In *Sheldon v. Metro-Goldwyn Pictures Corp.*, the Court reiterated that "a play may be pirated without using the dialogue," because "[t]he play is the sequence of the confluents of all these [words and gestures and scenery and costume and from the very looks of the actors themselves], bound together in an inseparable unity; it may often be most effectively pirated by leaving out the speech, for which a substitute can be found, which keeps the whole dramatic meaning.[668]

b. Use Test

The formal origin of the idea-expression dichotomy in the United States Supreme Court is *Baker v. Selden*.[669] Baker made a bookkeeping book similar to Selden's but with a different arrangement of columns and headings. The Court held that the copyright of Selden's book should not extend to the ruled lines and account headings necessary to use Selden's bookkeeping system. The Court reasoned:

> The very object of publishing a book on science or the useful arts is to communicate to the world the useful knowledge which it contains . . . . And where the art it teaches cannot be used without employing the methods and diagrams used to illustrate the book, or such as are similar to them, such methods and diagrams are to be considered as necessary incidents to the art, and given therewith to the public; not given for the purpose of publication in other works explanatory of the art, but for the purpose of practical application.[670]

---

[666] Samuels, *supra* note 662, at 344.
[667] Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930).
[668] Sheldon v. Metro-Goldwyn Pictures Corp., 81 F.2d 49, 55, 56 (2d Cir. 1936).
[669] Samuels, *supra* note 662, at 326.
[670] Baker v. Selden, 101 U.S. 99, 103 (1879).

A commentator thought the above reasoning contained "use test" and "merger theory."[671] The test "might be limited to particular kinds of works-architectural works, fashion designs, charts and forms - which by their nature merge idea and expression."[672]

c. Abstraction Test

The abstraction test asserts the copyright protects more than literal expressions. But how far does the copyright extend? In *Nichols*, Judge Hand made a well-known "abstraction test."[673] He described that a play could be generalized, beginning with its text and ending with the most general statement, saying:[674]

> Upon any work, ... a great number of patterns of increasing generality will fit equally well, as more and more of the incident is left out. The last may perhaps be no more than the most general statement of what the [work] is about, and at times consist only of its title; but there is a point in this series of abstractions where they are no longer protected . . . .

In *Sheldon*, Judge Hand reiterated, "the defendants were entitled to use, not only all that had gone before but even the plaintiffs' contribution itself if they drew from it only the more general patterns; that is, if they kept clear of its 'expression."[675] Also, in *Reyher v. Children's Television Workshop*, the court said, "it has been emphasized repeatedly that the essence of infringement lies in taking not a general theme but its particular expression through similarities of treatment, details, scenes, events and characterization." [676] To conclude, the idea is more general, and the expression is more particular. Despite admitting "[n]obody has ever been able to fix that boundary, and nobody ever can," as well as

---

[671] Samuels, *supra* note 662, at 327-329.
[672] *Id.* at 329.
[673] LEAFFER, *supra* note 80, at 420.
[674] Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930).
[675] Sheldon v. Metro-Goldwyn Pictures Corp., 81 F.2d 49, 54 (2d Cir. 1936).
[676] Reyher v. Children's Television Workshop, 533 F.2d 87, 91 (2d Cir. 1976).

"wherever it is drawn, will seem arbitrary,"[677] Judge Hand claimed, "that is no excuse for not drawing it; it is a question such as courts must answer in nearly all cases."[678]

For example, by applying the abstract test, the *Nichols* court found the defendant's film, *The Cohens and The Kellys*, did not copy the plaintiff's play, *Abie's Irish Rose*, more than the ideas. Through comparing the incidents between the works, "the only matter common to the two is a quarrel between a Jewish and an Irish father, the marriage of their children, the birth of grandchildren and a reconciliation," but "the theme was too generalized an abstraction from what she wrote. It was only a part of her 'ideas.'"[679] In contrast, the *Sheldon* Court found the defendant's film, *Letty Lynton*, copied the plaintiff's play, *Dishonored Lady*, because the dramatic significance of the scene between the two works was the same.[680]

Professor Leaffer comments, "The abstractions test is a fact-intensive inquiry that tries to determine at each level of abstraction whether there was sufficient expression to confer copyright protection on the work as a whole."[681] Considering the inherent vagueness of the abstraction test's application, Professor Samuels comments, "Perhaps the major significance of the abstraction test is not in its ability to work as a tool for deciding cases" but "its rejection of the Stowe approach to copyright protection."[682]

---

[677] *Nichols,* 45 F.2d at 121.
[678] *Id.* at 122.
[679] *Id.*
[680] *Sheldon,* 81 F.2d at 56.
[681] LEAFFER, *supra* note 80, at 422.
[682] Samuels, *supra* note 662, at 343-344.

d. Pattern Test

An attempt "to further delineate this boundary" between idea and expression is the pattern test suggested by Professor Chafee, who says, [683]

> [T]he line does lie somewhere between the author's idea and the precise form in which he wrote it down. I like to say that the protection covers the "pattern" of the work . . . . For example, the idea of an Irish-Jewish marriage in a play may be borrowed . . . . but . . . the pattern of the play - the sequence of events and the development of the interplay of the characters - must not be followed scene by scene. Such a correspondence of pattern would be an infringement although every word of the spoken dialogue was changed.[684]

In *Metcalf v. Bochco*, the CBS television series *City of Angels* allegedly infringed the plaintiff's screenplay's copyright. [685] The district court granted summary judgment for the defendants, holding the two works not substantially similar. However, based on "the presence of so many generic similarities and the common patterns in which they arise do help the Metcalfs satisfy the extrinsic test,"[686] the Court of Appeals reversed the summary judgment, saying:

> [P]rotectable expression includes the specific details of an author's rendering of ideas, or "the actual concrete elements that make up the total sequence of events and the relationships between the major characters." . . . The particular sequence in which an author strings a significant number of unprotectable elements can itself be a protectable element. Each note in a scale, for example, is not protectable, but a pattern of notes in a tune may earn copyright protection. A common "pattern [that] is sufficiently concrete . . . warrants a finding of substantial similarity."[687]

---

[683] *Id.* at 340; *see also* Reyher v. Children's Television Workshop, 533 F.2d 87, 91 (2d Cir. 1976); Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231, 240 (2d Cir. 1983).
[684] Zechariah Chafee Jr., *Reflections on the Law of Copyright: I*, 45 COLUM. L. REV. 503, 513-514 (1945).
[685] Metcalf v. Bochco, 294 F.3d 1069 (9th Cir. 2002).
[686] *Id.* at 1074.
[687] *Id.*

e. Purpose or Function Test

In *Whelan Assocs., Inc. v. Jaslow Dental Lab'y, Inc.*, the Third Circuit intended to establish a bright line between the idea and expression for utilitarian or functional works.[688] This copyright infringement case involved a dental-laboratory-operating computer program. The Court concluded the computer program's idea is "the efficient organization of a dental laboratory"[689] and its expression "may extend beyond the programs' literal code to their structure, sequence, and organization."[690] The Court reasoned:

> [T]he line between idea and expression may be drawn with reference to the end sought to be achieved by the work in question. In other words, the purpose or function of a utilitarian work would be the work's idea, and everything that is not necessary to that purpose or function would be part of the expression of the idea . . . . Where there are various means of achieving the desired purpose, then the particular means chosen is not necessary to the purpose; hence, there is expression, not idea.[691]

The *Whelan* test does not apply to works of literature or non-functional visual representations. However, the *Whelan* court's view of expressions has been criticized as overbroad because a computer "program's structure may be viewed as a process, systems, or methods of operation more akin to idea than expression."[692] In *Computer Assocs. Int'l, Inc. v. Altai, Inc*, the *Altai* Court enunciated that the *Whelan* Court erred in assuming that a computer program has only one "idea" so that it ignored a computer program may consist of numerous ideas at each level of abstraction from the lowest-level modules to highest-level modules.[693]

---

[688] Whelan Assocs., Inc. v. Jaslow Dental Lab'y, Inc., 797 F.2d 1222, 1238 (3d Cir. 1986).
[689] *Id.* at 1240.
[690] *Id.* at 1248.
[691] *Id.* at 1236.
[692] LEAFFER, *supra* note 80, at 422.
[693] Computer Assocs. Int'l, Inc. v. Altai, Inc., 982 F.2d 693, 705-707 (2d Cir. 1992).

2. China

Like the United States, the idea is unprotectable under China Copyright Law. In *Zhang Xiaoyan v. Lei Xianhe, Zhao Qi, and Shandong Aishuren DVDs & Books Co., Ltd.*,[694] China Supreme People's Court held:

> What was protected by the Copyright Law in China included original expressions of the author in the work, namely, the expression forms of ideas and emotions, excluding the ideas or emotions reflected in the work. The "ideas" as mentioned above included cognitions of material existence, objective facts, human feelings, and thinking ways and they were objects being described or expressed, which were within the subjective category. In virtue of the material medium, a thinker expressed his conception in forms and transferred an imago to an image, an abstract concept to a concrete object, a subjective thought to an objective form, and an intangible substance to a tangible object.

As described above, the idea is the object to be expressed, an imago, an abstract concept, a subjective thought, and an intangible substance. In contrast, the expression is the outcome of an idea, an image, a concrete object, an objective form, and a tangible object. Additionally, similar to the United States, the expression of literary works is not limited by the literal form in China. For example, in *Gao Yang v. Golden Vision*,[695] Beijing Intellectual Property Court affirmed the district court's opinion that a literary work's expression extends to specific character setup, character mutual relationships, character-specific characteristics, and specific plots centering around the characters.

---

[694] Zhang Xiaoyan Su Lei Xianhe Zhaoqi Sanding Aishuren Yinxiang Tushu Youxian Gongsi (张晓燕诉雷献和、赵琪、山东爱书人音像图书有限公司) [Zhang Xiaoyan v. Lei Xianhe, Zhao Qi, and Shandong Aishuren DVDs & Books Co.], 2013 Min Shen Zi No. 1049 (2013 民申字第 1049 号) (Sup. People's Ct. 2014).

[695] Gaoyang Su Jinse Shizu (Beijing) Yingshi Wenhua Youxian Gongsi Zhuzuoquan Quanshu Qinquan Jiufen (高阳诉金色视族（北京）影视文化有限公司著作权权属、侵权纠纷) [Gao Yang v. Golden Vision (Beijing) Film and Television Culture Co.], 2017 Jing 73 Min Zhong No. 797 (2017 京 73 民终 797 号) (Beijing IP Ct. Apr. 2, 2020) [法宝引证码 CLI.C.97510788].

3. Taiwan

Like the United States and China, Article 10-1 of Taiwan Copyright Law provides, "Protection for copyright that has been obtained in accordance with this Act shall only extend to the expression of the work in question, and shall not extend to the work's underlying ideas, procedures, production processes, systems, methods of operation, concepts, principles, or discoveries." [696] This is the so-called idea and expression dichotomy,[697] which treats ideas and concepts in the public domain. Anyone is free to create a new work based on inspirations of other works' ideas. The purpose is to prevent a monopoly from harming creativity and cultural development. [698] Examples of ideas include: a man and a woman meeting by chance and hitting it off,[699] anthropomorphic water animal cartoon characters,[700] a graphic design symbolizing a butterfly.[701]

Similar to the United States and China, copyright protection in Taiwan is not limited to the copyrighted work's specific language but extends to the interpretation, treatment, arrangement, and sequence that convey the author's ideas or facts. [702] With

---

[696] *See* Zhuzuoquan Fa (著作權法) [Copyright Act] (Taiwan) § 10-1.

[697] *See* Huang Pingying Su You Shujun (黃萍瑛訴游淑珺) [Huang Ping Ying v. You Shu Jun], 104 Min Zhu Su No.39 (104 年度民著訴字第 39 號) (Intell. Prop. Ct. Sept. 30, 2016) (Taiwan).

[698] Helan Shang Endemol Nederland B.V. Zisu Quanneng Zhizuo Gufen Youxian Gongsi (荷蘭商 Endemol Nederland B.V. 自訴全能製作股份有限公司) [Endemol Nederland B.V. v. All Full Power Co.], 103 Xing Zhi Shang Yi No. 56 (103 年度刑智上易字第 56 號) (Intell. Prop. Ct. Sept. 3, 2015) (Taiwan).

[699] Liao Fubin Su Daxin Changpian Gufen Youxian Gongsi (廖福彬訴大信唱片股份有限公司) [Liao Fu-Bin v. Daishin Music Co.], 90 Shang Zi No. 1252 (90 年上字第 1252 號) (Taiwan High Ct. May 25, 2004).

[700] Shouying Chuangyi Gufen Youxian Gongsi Su Yuanjin Guoji Youxian Gongsi (首映創意股份有限公司訴原金國際有限公司) [Sofa Studio Co. v. Engine Studios LLC], 110 Min Zhu Shang Zi No. 4 (111 年度民著上字第 4 號) (Intell. Prop. & Com. Ct. Feb. 9, 2023) (Taiwan).

[701] Jian YongCang Weifan Zhuzuoquan Fa Anjian (簡永倉違反著作權法案件) [Republic of China v. Jian Yong Cang], 104 Xing Zhi Shang Yi Zi No. 75 (104 年刑智上易字第 75 號) (Intell. Prop. Ct. May 13, 2016) (Taiwan).

[702] Guanyu Jixie Gufen Youxian Gongsi Weifan Zhuzuoquan Fa Anjian (冠昱機械股份有限公司違反著作權法案件) [Republic of China v. Guan-Yu Machine Co.], 105 Xing Zhi Shang Su Zi No. 42 (105 年度刑智上訴字第 42 號) (Intell. Prop. Ct. Apr. 16, 2020) (Taiwan).

regard to literary works, the court applies the abstraction test.[703] Simply replacing, deleting, or replacing the copyrighted work's text does not help the infringer to escape infringement liability. With regard to graphical, photographic, artistic, and audiovisual works, the court applies the "total concept and feel" test in examining the quality of substantial similarity.[704]

C. FACT V. EXPRESSION

1. United States

"This dichotomy between facts and their expression"[705] strikes a balance between "the public's interest in stimulating creative activity" and "the public's need for unrestrained access to information."[706] Although a factual compilation is copyrightable if the compiler makes the selection, coordination, and arrangement independently and presents minimal creativity, [707] the constituent historical fact, whether correct or incorrect,[708] is not an original work of authorship and is always uncopyrightable.[709] Even if the "explanatory hypothesis" is involved, the courts hold anyone is free to use it, as long as the follower does not copy the prior author's "original expression of particular facts and theories"[710] or make a "verbatim reproduction" of the prior work.[711] The rationale is that

---

[703] *Id.*

[704] Meishang Meizhou Rentiantang Gufen Youxian Gongsi Su Jinfu Wanju Shiye Youxian Gongsi (美商美洲任天堂股份有限公司訴金富玩具實業有限公司) [Nintendo Co. v. Jin Fu Toy Ltd.], 110 Min Zhu Shang Zi No. 11 (110 年度民著上字第 11 號) (Intell. Prop. & Com. Ct. Nov. 18, 2021) (Taiwan).

[705] Miller v. Universal City Studios, Inc., 650 F.2d 1365, 1368 (5th Cir. 1981).

[706] *Id.* at 1371.

[707] Feist Publications, Inc. v. Rural Telephone Service Co., Inc., 499 U.S. 340, 348 (1991).

[708] Hoehling v. Universal City Studios, Inc., 618 F.2d 972, 979 (2d Cir. 1980).

[709] LEAFFER, *supra* note 80, at 90.

[710] *Hoehling*, 618 F.2d at 974.

[711] *Id.* at 980.

"the cause of knowledge is best served when history is the common property of all, and each generation remains free to draw upon the discoveries and insights of the past."[712]

For example, in *Hoehling v. Universal City Studios, Inc.*, the Second Circuit held Hoehling's hypothesis regarding the motivation and cause of the Hindenburg airship disaster on May 6, 1937, as a "historical interpretation, whether or not it originated with Mr. Hoehling, is not protected by his copyright and can be freely used by subsequent authors."[713] Also, in *Miller v. Universal City Studios, Inc.*,[714] the Fifth Circuit held the plaintiff's hard research on a crime, in which a rich land developer's college daughter was abducted and buried underground alive for five days, is uncopyrightable, as "[t]o hold that research is copyrightable is no more or no less than to hold that the facts discovered due to research are entitled to copyright protection."[715] Besides, in *Abdin v. CBS Broad. Inc.*, the plaintiff claimed the character, a tardigrade named "Ripper," in the defendant's television series *Star Trek: Discovery* infringed the plaintiff's video game copyright. The Second Circuit affirmed the district court's conclusion that all tardigrades had known features of "eight short legs in pairs along a rounded body," "an O-shaped mouth in the center of the face," and the capability "of surviving in space," which are uncopyrightable scientific facts.[716]

Furthermore, "[u]nder the doctrine of copyright estoppel, once a plaintiff's work has been held out to the public as factual, the author-plaintiff cannot then claim that the book is, in actuality, fiction and thus entitled to the higher protection allowed to fictional

---

[712] *Id.* at 974.
[713] *Id.* at 979.
[714] Miller v. Universal City Studios, Inc., 650 F.2d 1365 (5th Cir. 1981).
[715] *Id.* at 1372.
[716] Abdin v. CBS Broad. Inc., 971 F.3d 57, 67 (2d Cir. 2020).

works."[717] For example, in *Marshall v. Yates*, the issue was whether the defendants' film "*France"* infringed the plaintiffs' copyright of the book "*Shadowland*."[718] The Court refuted the book as containing eight "fictional elements," as the book was presented to the public as the true life of the famous actress, Frances Farmer. Again, in *Houts v. Universal City Studios, Inc.*, the plaintiff was an author of the book *Where Death Delights*, whose cover or first page marked "real life detective stories," "N-F," and claimed that "[h]ere is a book that shows that truth can be more brutal than fiction."[719] The court concluded that the book was held out to the public as factual, thus, unable to be protected as a fictional work.

2. China

Like the United States, historical facts are unprotectable in China. For example, in *Zhang Xiaoyan v. Lei Xianhe, Zhao Qi, and Shandong Aishuren DVDs & Books Co., Ltd.*,[720] Zhang alleged her TV series, *A Cavalry Troop on the Plateau*, was infringed by the Lei's TV series, *Last Cavalry*. Lei denied copying and argued his script was based on two novels, *Boundless Sky* and *Traveling to the End of the World by Riding and Carrying a Gun*. Ascertaining that all the TV scripts and novels were based on historical facts about the cavalry downsizing in the middle 1980s, China Supreme People's Court found "the theme mainline and overall sequence of clues . . . were the commonwealth of society and could not be monopolized by any individual."

---

[717] Houts v. Universal City Studios, Inc., 603 F. Supp. 26, 28 (C.D. Cal. 1984).
[718] Marshall v. Yates, No. CV-81-1850-MML, 1983 WL 1148 (C.D. Cal. Oct. 26, 1983).
[719] *Houts*, 603 F. Supp. at 28.
[720] Zhang Xiaoyan Su Lei Xianhe Zhaoqi Sanding Aishuren Yinxiang Tushu Youxian Gongsi (张晓燕诉雷献和、赵琪、山东爱书人音像图书有限公司) [Zhang Xiaoyan v. Lei Xianhe, Zhao Qi, and Shandong Aishuren DVDs & Books Co., Ltd.], 2013 Min Shen Zi No. 1049 (2013 民申字第 1049 号) (Sup. People's Ct. 2014).

Similarly, in *Gao Yang v. Golden Vision*,[721] Beijing Intellectual Property Court affirms the district court's opinion that a factual literary work has both characteristics of documentary and literature. With regard to such works, the court must filter out the facts and exclude them from copyright protection. In addition, the court must analyze which elements are protectable expressions and which elements are unprotectable ideas from a literature perspective. In this case, the plaintiffs published a short article entitled *The Naughty Kids Who Chase Balloons* that was based on their activity to shoot the earth by using a camera attached to a hot air balloon. Three months later, the defendant released a commercial of the same name on the Youku website. Though both works mentioned that the balloon flew over 300 kilometers and was retrieved on a paddy field. The court held these were simple facts rather than expressions.

3. Taiwan

Like the United States and China, the fact is uncopyrightable in Taiwan. Article 9 paragraph 1 item 4 of Taiwan Copyright Law provides, "oral and literary works for news reports that are intended strictly to communicate facts" are not the subject matter of copyright.[722] On the one hand, the dissemination of facts is related to the public interest and should not be monopolized by any person. On the other hand, facts are not original creations and thus not copyrightable.[723] However, if the reporter has added personal

---

[721] Gaoyang Su Jinse Shizu (Beijing) Yingshi Wenhua Youxian Gongsi Zhuzuoquan Quanshu Qinquan Jiufen (高阳诉金色视族（北京）影视文化有限公司著作权权属、侵权纠纷) [Gao Yang v. Golden Vision (Beijing) Film and Television Culture Co.], 2017 Jing 73 Min Zhong No. 797 (2017 京 73 民终 797 号) (Beijing IP Ct. Apr. 2, 2020) [法宝引证码 CLI.C.97510788].

[722] *See* Zhuzuoquan Fa (著作權法) [Copyright Act] (Taiwan)] § 9, I, (4).

[723] Zeng Fanqi Su Zhonghua Dianshi Gufen Youxian Gongsi (曾繁祺訴中華電視股份有限公司) [Zeng Fan Qi v. Chinese Television Sys. Co.], 109 Min Zhu Su Zi No. 76 (109 年度民著訴字第 76 號) (Intell. Pro. Ct. Nov. 9, 2020) (Taiwan).

thoughts into the news title, material selection, writing, structure, arrangement, and news interpretation so as to express the reporter's personality and writing style to minimum creativity, the news report shall be a copyrightable work.[724]

Moreover, copyright protection does not extend to natural phenomena. For example, cats are natural animals that have the usual appearance of circle faces, sharp ears, white fur with yellow stripes, stripes distributed from the top of the head to the sides of the cheeks or have a prone position with the front palms together.[725] The first user cannot monopolize these natural features and prevent the second user from using these features.

Furthermore, copyright protection does not extend to historical research. For example, in a case regarding two studies about Taiwanese folk beliefs to the souls of unmarried dead women, the court held that descriptions with respect to the customs or temples worshiping unmarried female souls were Taiwanese customs, not the plaintiff's original creations.[726] Even though the plaintiff did extensive field research, she cannot monopolize those facts.

D. MERGER DOCTRINE V. *SCÈNES À FAIRE*

1. United States

Unlike the idea and expression dichotomy focuses on "whether the work constitutes

---

[724] *Id.*

[725] Zhu Peixuan Weifan Zhuzuoquan Fa Anjian (朱珮萱違反著作權法案件) [The Republic of China v. Zhu Pei Xuan], 108 Xing Zhi Shang Su Zi No. 41 (108 年度刑智上訴字第 41 號) (Intell. Prop. Ct. Nov. 14, 2019) (Taiwan).

[726] *See* Huang Pingying Su You Shujun (黃萍瑛訴游淑珺) [Huang Ping Ying v. You Shu Jun], 104 Min Zhu Su No.39 (104 年度民著訴字第 39 號) (Intell. Prop. Ct. Sept. 30, 2016) (Taiwan).

idea or expression," [727] the merger doctrine focuses on whether there are capable "alternative expressions" to use the same idea.[728] "When there is essentially only one way to express an idea, the idea and its expression are inseparable, and copyright is no bar to copying that expression."[729] In other words, since the idea and expression have "merged" in this situation, to avoid the monopoly of the idea, everyone is free to copy the only available expression for using the idea. The courts have extended the merger doctrine from cases with only one way to express an idea to cases with only a few ways to express an idea.[730] Typically, the merger doctrine is applied to utilitarian works.

A related doctrine applied to fictional or narrative works is *Scènes à Faire*,[731] which means "sequences of events which necessarily follow from a common theme"[732] or "incidents, characters or settings which are as a practical matter indispensable, or at least standard, in the treatment of a given topic."[733] Judge Yankwich borrowed the French term to describe details necessary to the environment or setting of an action.[734] *Scènes à Faire* is not copyrightable "[b]ecause it is virtually impossible to write about a particular historical era or fictional theme without employing certain 'stock' or standard literary devices."[735] For example, in *Walker v. Time Life Films, Inc.*,[736] the plaintiff claimed the defendant's policeman film *Fort Apache: The Bronx* shared many similarities with his

---

[727] Samuels, *supra* note 662, at 383.

[728] *Id.*

[729] Concrete Machinery Co. v. Classic Lawn Ornaments, Inc., 843 F.2d 600, 606 (1st Cir.1988).

[730] "The merger doctrine is applied as a prophylactic device to ensure that courts do not unwittingly grant protection to an idea by granting exclusive rights to the only, or one of only a few, means of expressing that idea." Gates Rubber Co. v. Bando Chem. Indus., Ltd., 9 F.3d 823, 838 (10th Cir. 1993).

[731] *See* LEAFFER, *supra* note 80, at 87-88.

[732] Reyher v. Children's Television Workshop, 533 F.2d 87, 91 (2d Cir. 1976).

[733] Hoehling v. Universal City Studios, Inc., 618 F.2d 972, 979 (2d Cir. 1980).

[734] Cain v. Universal Pictures Co., 47 F. Supp. 1013, 1017 (S.D. Cal. 1942).

[735] *Hoehling*, 618 F.2d at 979.

[736] Walker v. Time Life Films, Inc., 784 F.2d 44 (2d Cir. 1986).

book *Fort Apache*. However, the Court held "drunks, prostitutes, vermin and derelict cars" are *scenes a faire* "that necessarily result from the choice of a setting or situation" about the policemen's work in South Bronx.[737]

Judge Posner described "s*cènes à faire*—standard expressions, like language itself, without which the would-be author of an expressive work would be speechless."[738] In *Bucklew*, he enunciated the doctrine:[739]

> The doctrine of *scènes à faire* (another confusing label, literally "scenes for action," which the Oxford English Dictionary tells us is a theatrical term meaning "the most important scene in a play or opera, made inevitable by the action which leads up to it"—which is not the legal doctrine at all) teaches, sensibly enough, that a copyright owner can't prove infringement by pointing to features of his work that are found in the defendant's work as well but that are so rudimentary, commonplace, standard, or unavoidable that they do not serve to distinguish one work within a class of works from another.

Actually, the litigated works in the *Bucklew* are computer forms. The *Bucklew* Court found three forms "either were *scènes à faire* or weren't expressive elements at all." Also, the *Gates* Court used *scènes à faire* to exclude some elements of a computer program that were dictated by external factors.[740] Professor Leaffer comments it is "unfortunate" that the courts use the *Scènes à Faire* instead of the merger doctrine in utilitarian works, which results in the loss of the precise meaning of *Scènes à Faire*.[741]

Scènes à Faire has been frequently applied to negate substantial similarity. For example, "cowboys, bank robbers, and shootouts in stories of the American West,"

---

[737] *Id.*
[738] Bucklew v. Hawkins, Ash, Baptie & Co., LLP., 329 F.3d 923, 930 (7th Cir. 2003).
[739] *Id.*
[740] Gates Rubber Co. v. Bando Chem. Indus., Ltd., 9 F.3d 823, 838 (10th Cir. 1993).
[741] *See* LEAFFER, *supra* note 80, at 87-88.

"electrified fences, automated tours, dinosaur nurseries, and uniformed workers . . . of a dinosaur zoo," "a space ship, space travel, and alien encounters . . . in the science fiction genre,"[742] and "the sequencing of first performing the trick and then revealing the secrets behind the trick."[743] Professor Samuels comments that some *scènes à faire* cases regarding stock treatments do not "involve works which are capable of only one or only a few expressions."[744] Such cases may be better held as "not original" or the second work only copies "generalized treatment that is in the public domain."[745]

2. China

Like the United States, the *Scènes à Faire* doctrine is adopted by Chinese courts. Certain incidents, characters, settings, and scenes that are generally used or must be used to express a specific idea are uncopyrightable.[746] When the idea and expression have merged or there are only one or limited-expression forms to a specific idea, such expression is deemed as an idea and thus uncopyrightable.

For example, in *Zhang Xiaoyan v. Lei Xianhe*,[747] the plaintiff's TV series, *A Cavalry Troop on the Plateau*, and the defendant's TV series, *Last Cavalry*, shared the same theme of "the end of a hero and the cavalry's peak of poetic perfection" and depicted

---

[742] Abdin v. CBS Broad. Inc., 971 F.3d 57, 71-72 (2d Cir. 2020).
[743] Rice v. Fox Broad. Co., 330 F.3d 1170, 1177 (9th Cir. 2003).
[744] Samuels, *supra* note 662, at 386.
[745] *Id.*
[746] Gaoyang Su Jinse Shizu (Beijing) Yingshi Wenhua Youxian Gongsi Zhuzuoquan Quanshu Qinquan Jiufen (高阳诉金色视族（北京）影视文化有限公司著作权权属、侵权纠纷) [Gao Yang v. Golden Vision (Beijing) Film and Television Culture Co.], 2017 Jing 73 Min Zhong No. 797 (2017 京 73 民终 797 号) (Beijing IP Ct. Apr. 2, 2020) [法宝引证码 CLI.C.97510788]; *see also* Zhang Xiaoyan Su Lei Xianhe Zhaoqi Sanding Aishuren Yinxiang Tushu Youxian Gongsi (张晓燕诉雷献和、赵琪、山东爱书人音像图书有限公司) [Zhang Xiaoyan v. Lei Xianhe, Zhao Qi, and Shandong Aishuren DVDs & Books Co], 2013 Min Shen Zi No. 1049 (2013 民申字第 1049 号) (Sup. People's Ct. 2014).
[747] *Zhang Xiaoyan,* 2013 Min Shen Zi No. 1049.

the cavalry life under a specific historical context. China Supreme People's Court held that such characters' settings and relationships as a love triangle, superior and subordinate, and cavalry and the people were inevitable and necessary scenes in this type of cavalry drama. Since such expression forms were limited, they were not protected by the Copyright Law.

3. Taiwan

Like the United States and China, the Taiwan Supreme Court holds that when there is only one way or there are extremely limited ways to express an idea or a concept, the expression is not copyrightable.[748] Otherwise, the copyright holder will monopolize the idea or concept and thus inhibit human cultural and artistic development. Such copyright protection also invades people's freedom of speech, teaching, writing and publication.

For example, in *Prudential Life*,[749] the defendant's print advertisement was alleged to infringe the copyright of the plaintiff's insurance product advertisements, based on both works having the same coverage for living expenses, parents' filial piety pension, education expenses, mortgage, rent expenses, final expenses, inheritance tax preparation, retirement expenses, other expenses, and medical expenses. However, those items were often used by insurers to design their products because those items were what most customers cared about. The court held the plaintiff's marketing materials lacked originality because there were no better substitutes or extremely limited choices based on the market

---

[748] Wang Taichang Su Meishang Maigeluo Xier Guoji Gufen Youxian Gongsi Taiwan Fengongsi (王泰昌訴美商麥格羅．希爾國際股份有限公司台灣分公司) [Wang Tai Chang v. McGraw-Hill Education Co., Taiwan Branch], 104 Tai Shang Zi No.1251 (104 年度台上字第 1251 號) (Sup. Ct. 2015) (Taiwan).
[749] Baodexin Guoji Renshou Baoxian Gufen Youxian Gongsi Su Daduhui Guoji Renshou Baoxian Gufen Youxian Gongsi (保德信國際人壽保險股份有限公司訴大都會國際人壽保險股份有限公司) [Prudential Life Ins. Co. v. Metro. Life Ins. Co.], 98 Ming Zu Shang Zi No.16 (98 年度民著上字第 16 號) (Intell. Prop. Ct. Feb. 4, 2010) (Taiwan).

demand. It will also hinder fair market competition if allowing the plaintiff to prevent other people's use of the same selection by enforcing copyright.

Regarding *Scènes à Faire*, the Taiwan Intellectual Property Court holds that it is a supplemental rule of the merge doctrine.[750] *Scènes à Faire* means the inevitable or standardized events, characters, settings, and layouts for creating a work of a specific subject. Such events, characters, settings, and layouts do not constitute copyright infringement even though they are substantially similar to others' work.

E. CHARACTERS

1. United States

Characters are essential components of motion pictures, but stock or stereotyped characters are not entitled to copyright protection.[751] Examples include:

> [A] drunken old bum . . . a drunken suburban housewife, a gesticulating Frenchman, a fire-breathing dragon, a talking cat, a Prussian officer who wears a monocle and clicks his heels, a masked magician . . . [752]

Following Judge Hand's statement in *Nicholes* "that the less developed the characters, the less they can be copyrighted; that is the penalty an author must bear for marking them too indistinctly,"[753] the Second Circuit holds that copyright does not protect "basic or 'stock' character types . . . that are as a practical matter indispensable or standard

---

[750] Wangyin Guoji Gufen Youxian Gongsi Su Yinhe Xianshang Gufen Youxian Gongsi (網銀國際股份有限公司訴銀河線上股份有限公司) (Wanin Int'l Co., v. Galaxy Co.), 106 Min Zhu Su Zi No. 5 (106 年度民著訴字第 5 號) (Intell. Prop. Ct. Jan. 25, 2018) (Taiwan).
[751] DC Comics v. Towle, 802 F.3d 1012, 1019 (9th Cir. 2015).
[752] Gaiman v. McFarlane, 360 F.3d 644, 660 (7th Cir. 2004).
[753] Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930).

in the treatment of a given topic"[754] or "generic and generalized character traits such as race, gender, and hair color."[755]

Likewise, the Ninth Circuit holds only "characters that are 'especially distinctive' or the 'story being told'" can enjoy copyright protection.[756] To receive copyright protection, the Ninth Circuit's three-part test (Towle test) reviews if "(1) the character has 'physical as well as conceptual qualities,' (2) the character is 'sufficiently delineated to be recognizable as the same character whenever it appears' and 'display[s] consistent, identifiable character traits and attributes,' and (3) the character is 'especially distinctive' and 'contain[s] some unique elements of expression.'"[757] Another test for character's copyrightability is "the story being told test" (Warner Brothers test).[758] To be copyrightable, the character must really constitute "the story being told," not being "only the chessman in the game of telling the story."[759]

2. China

Like the United States, simple and abstract character identities, relationships, and characteristics are unoriginal or uncopyrightable ideas if not depicted sufficiently concretely by story plots.[760] For example, middle school students, classmate relationships, and gender are uncopyrightable.

---

[754] CK Co. v. Burger King Corp., 122 F.3d 1055 (2d Cir. 1995).
[755] Abdin v. CBS Broad. Inc., 971 F.3d 57, 67 (2d Cir. 2020).
[756] Rice v. Fox Broad. Co., 330 F.3d 1170, 1175 (9th Cir. 2003).
[757] Daniels v. Walt Disney Co., 958 F.3d 767, 771 (9th Cir. 2020); *see also* DC Comics v. Towle, 802 F.3d 1012, 1021 (9th Cir. 2015); Esplanade Prods., Inc. v. Walt Disney Co., No. CV1702185MWFJCX, 2017 WL 5635024, at *8 (C.D. Cal. July 11, 2017).
[758] *Daniels*, 958 F.3d at 774.
[759] Warner Bros. Pictures v. Columbia Broad. Sys., 216 F.2d 945, 950 (9th Cir. 1954).
[760] Gaoyang Su Jinse Shizu (Beijing) Yingshi Wenhua Youxian Gongsi Zhuzuoquan Quanshu Qinquan Jiufen (高阳诉金色视族（北京）影视文化有限公司著作权权属、侵权纠纷) [Gao Yang v. Golden

3. Taiwan

Like the United States and China, if a character has been highly developed, sufficiently delineated, and its story is fully told, the character will be more prone to expression than the idea.[761] Therefore, others have less room to use such characters. For example, Linghu Chong, Yang Guo, Little Dragon Girl, Wei Xiaobao from Jin Yong's novels, Zoro, Sanji from One Piece, and Pikachu from Pokémon. If someone uses these characters that have well-known personalities and stories to create a derivative work, he or she will infringe the author's adoption rights.

III. MOTION PICTURE'S COPYRIGHT INFRINGEMENT

If a motion picture's copyright is infringed, the copyright owner must prove three elements to prevail an action for infringement: "(1) ownership of a valid copyright in the work, (2) copying by the defendant, and (3) that the defendant's copying constitutes an improper appropriation."[762] "Improper appropriation" means "the defendant copied a sufficient amount of the protectable elements of the plaintiff's copyrighted work as to render the two works substantially similar."[763] However, if only "literal copying of a small and usually insignificant portion of the plaintiff's work" is shown, such de minimis copying does not support copyright infringement.[764] The following section will focus on the "copying" and "substantial similarity" elements.

Vision (Beijing) Film and Television Culture Co.], 2017 Jing 73 Min Zhong No. 797 (2017 京 73 民终 797 号) (Beijing IP Ct. Apr. 2, 2020) [法宝引证码 CLI.C.97510788].

[761] Taiwan Intellectual Property Office, No. 1070110 Email (Jan. 10, 2018); Taiwan Intellectual Property Office, No. 1081219 Email (Dec. 19, 2019).

[762] LEAFFER, *supra* note 80, at 413.

[763] *Id.* at 418; *see also* Abdin v. CBS Broad. Inc., 971 F.3d 57, 66 (2d Cir. 2020).

[764] Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231, 242 (2d Cir. 1983).

A. COPYING

1. United States

When the defendant denies copying, there are two ways to establish copying: direct and indirect proof of copying. The defendant does not have the burden to produce any evidence showing independent creation or a common source before the plaintiff establishes copying.[765] It is insufficient to prove copying by demonstrating the defendant "has duplicated a few ordinary phrases and paraphrased largely factual statements" if the defendant was "creating an entirely different kind of story."[766] On the other hand, it may be enough to prove copying by demonstrating the defendant has made close and frequent paraphrases of protected expression over necessity.[767] If there is no direct evidence of copying, the plaintiff can establish circumstantial copying by showing that the defendant had "access" to the plaintiff's work and both works are "substantially similar."[768]

To prove the "access" requirement, the plaintiff must demonstrate that the defendant had a "reasonable opportunity or reasonable possibility of viewing" the allegedly infringed work.[769] If the evidence of access is absent, the plaintiff can still establish copying when the similarities between the two works are "so striking as to preclude the possibility that plaintiff and defendant independently arrived at the same result."[770] On the contrary, "[n]o amount of proof of access will suffice to show copying if there are no

---

[765] *See* LEAFFER, *supra* note 80, at 417-418.

[766] Narell v. Freeman, 872 F.2d 907, 912 (9th Cir. 1989) (rejecting the plaintiff's copyright violation claim against word-for-word copying and paraphrasing over 300 words).

[767] Salinger v. Random House, Inc., 811 F.2d 90, 98 (2d Cir.), opinion supplemented on denial of reh'g, 818 F.2d 252 (2d Cir. 1987).

[768] Funky Films, Inc. v. Time Warner Ent. Co., L.P., 462 F.3d 1072, 1076 (9th Cir. 2006).

[769] Astor-White v. Strong, 817 F. App'x 502, 503 (9th Cir. 2020).

[770] Arnstein v. Porter, 154 F.2d 464, 468 (2d Cir. 1946).

similarities."[771] Neither can a high degree of access relax the required substantial similarity, at least in the Second, Fifth, Seventh, Ninth, and Eleventh Circuits. These circuits have expressly denied "a lower standard of proof of substantial similarity when a high degree of access is shown."[772] In *Skidmore v. Led Zeppelin*, the Ninth Circuit Court of Appeal states:[773]

> [T]he concept of "access" is increasingly diluted in our digitally interconnected world . . . . Given the ubiquity of ways to access media online, from YouTube to subscription services like Netflix and Spotify, access may be established by a trivial showing that the work is available on demand . . . . To the extent "access" still has meaning, the inverse ratio rule unfairly advantages those whose work is most accessible by lowering the standard of proof for similarity . . . . By rejecting the inverse ratio rule, we are not suggesting that access cannot serve as circumstantial evidence of actual copying in all cases; access, however, in no way can prove substantial similarity.

2. China

Like the United States, Chinese courts may hold the defendant copied the plaintiff's work if: (1) the alleged infringing work is identical or substantially similar to the plaintiff's work, and (2) the defendant had or was likely to have access to the defendant's prior work when he/she created the alleged infringing work.[774]

With regard to determining "access," the court generally considers: (1) whether the prior work has been published; (2) if the prior work is not published, whether the plaintiff has ever submitted contributions to the defendant or has cooperated with the defendant;

---

[771] Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp., 562 F.2d 1157, 1172 (9th Cir. 1977).
[772] Skidmore as Tr. for Randy Craig Wolfe Tr. v. Led Zeppelin, 952 F.3d 1051, 1066 (9th Cir. 2020).
[773] *Id.* at 1068-1069.
[774] *See* Beijing Shi Gaoji Renmin Fayuan Qinhaiz Huzuoquan Anjian Shenli Zhinan (北京市高级人民法院侵害著作权案件审理指南) [Beijing High People's Court Guidelines for the Trial of Copyright Infringement Cases] (promulgated by the Beijing High People's Ct., Apr. 20, 2018) § 10.7.

and (3) whether the expression of two works are identical or strikingly similar as to exclude the possibility of independent creation. If so, whether the defendant has a reasonable explanation.[775]

3. Taiwan

Unlike the United States, Taiwan does not divide direct and indirect proof of copying. Like China, Taiwanese courts always use "access" and "substantial similarity" to analyze copyright infringement cases (see **Table 4-1**).[776]

Table 4-1 The Comparison of Copyright Infringement Elements

Countries	Element 1	Element 2
U. S.	Direct proof of copying or Indirect Proof of Copying (Access + Substantial Similarity)	Substantial Similarity (Improper Appropriation)
China	Access	Substantial Similarity
Taiwan	Access (Direct or Indirect)	Substantial Similarity

The "access" element requires the alleged infringer actually or has reasonable opportunity or likeness to read or hear the protected work. It is further divided into direct access and indirect access: (1) direct access means the alleged infringer has read or obtained the protected work from the copyright holder, e.g., the alleged infringer has participated in the creation of the plaintiff's work;[777] (2) indirect access means the alleged infringer has a

---

[775] *Id.* at §10.8.
[776] Caituan Faren Laihe Wenjiao Jijinhui Su Lianhe Baike Dianzi Chuban Youxian Gongsi (財團法人賴和文教基金會訴聯合百科電子出版有限公司) [Laiho Cultural and Educ. Found. v. United Digital Publ'n Co.], 108 Tai Shang Zi No.1315 (108 年度台上字第 1315 號) (Sup. Ct. Oct 9, 2019) (Taiwan).
[777] *See* Zhiguan Keji Gufen Youxian Gongsi Su Heluo Youxi Youxian Gongsi (智冠科技股份有限公司訴河洛遊戲有限公司) [Soft-World Int'l Co. v. Heluo Games Co.], 108 Min Zhu Shang Zi No.3 (108 年度民著上字第 3 號) (Intell. Prop. Ct. July 30, 2010) (Taiwan).

reasonable opportunity to access the protected work under reasonable circumstances, e.g., the protected work has been published and anyone can easily obtain the protected work on the market,[778] or two works are strikingly similar to exclude the possibility of independent creation, common errors, incorrect footnotes, or redundant pleonasm.[779]

B. SUBSTANTIAL SIMILARITY

1. United States

Substantial similarity can be used to prove circumstantial copying or improper appropriation. Though the requirements for each purpose are not necessarily the same, many cases use the same proofs of similarity for both purposes.[780] Some scholars and courts prefer to use the term "probative similarity" for showing circumstantial copying, in order to distinguish the term "substantial similarity" used for showing improper appropriation.[781]

a. Second Circuit v. Ninth Circuit

(1) Two-Part Test

For assessing substantial similarity, the Second Circuit applies a two-part test, non-

---

[778] Guanyu Jixie Gufen Youxian Gongsi Weifan Zhuzuoquan Fa Anjian (冠昱機械股份有限公司違反著作權法案件) [Republic of China v. Guan-Yu Machine Co.], 105 Xing Zhi Shang Su Zi No. 42 (105 年度刑智上訴字第 42 號) (Intell. Prop. Ct. Apr. 16, 2020) (Taiwan).

[779] Chuanqi Wangle Youxi Gufen Youxian Gongsi Su Xingyu Hudong Yule Keji Gufen Youxian Gongsi (傳奇網路遊戲股份有限公司訴星宇互動娛樂科技股份有限公司) X-Legend Ent. Co. v. Loftstar Interactive Ent. Inc.), 107 Min Zhu Shang Yi Zi No.9 (107 年度民著上易字第 9 號) (Intell. Prop. Ct. Jan. 24, 2019) (Taiwan).

[780] *See* LEAFFER, *supra* note 80, at 417.

[781] *See id.*; *see also* Laureyssens v. Idea Grp., Inc., 964 F.2d 131, 140 (2d Cir. 1992) (stating "Copying may be established either by direct evidence of copying or by indirect evidence, including access to the copyrighted work, similarities that are probative of copying between the works, and expert testimony. If actual copying is established, a plaintiff must then show that the copying amounts to an improper appropriation by demonstrating that substantial similarity to protected material exists between the two works.)

infringing "copying" and infringing "illicit copying."[782] The former "may be inferred from substantial similarities between the two works," whereas the latter "demands that such similarities relate to protectible material."[783] Expert testimony is admissible in proving "copying" but not "illicit copying."[784]

The courts evaluate substantial similarity regarding "the expression of ideas, not the ideas themselves."[785] "After a detailed examination of the works themselves,"[786] the standard or general test is whether a "reasonable observer could find them substantially similar beyond the level of generalized or otherwise nonprotectible ideas"[787] Some courts apply a "more discerning" observer test when the allegedly infringed work "incorporates unprotectible elements from the public domain,"[788] as the test requires to "extract the unprotectible elements from . . . consideration and ask whether the protectible elements, standing alone, are substantially similar."[789]

Alternatively, the Ninth Circuit applies "the extrinsic test and the intrinsic test."[790] Only when both tests are satisfied can a jury find substantial similarity. On the one hand, the extrinsic test is an objective test, which examines similarities in two works' expressive elements such as "the plot, themes, dialogue, mood, setting, pace, characters, and sequence

[782] Walker v. Time Life Films, Inc., 784 F.2d 44, 51 (2d Cir. 1986).
[783] Id.
[784] Id. at 52.
[785] Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231, 239 (2d Cir. 1983).
[786] Walker, 784 F.2d at 48-49.
[787] Id. at 49; see also Warner Bros. Inc. v. Am. Broad. Companies, Inc., 654 F.2d 204, 208 (2d Cir. 1981) (stating "whether an average lay observer would recognize the alleged copy as having been appropriated from the copyrighted work."); Abdin v. CBS Broad. Inc., 971 F.3d 57, 66 (2d Cir. 2020) (stating "whether an ordinary observer, unless he set out to detect the disparities, would be disposed to overlook them, and regard [the] aesthetic appeal as the same.")
[788] Abdin, 971 F.3d at 66.
[789] Id. at 73-74.
[790] Kouf v. Walt Disney Pictures & Television, 16 F.3d 1042, 1045 (9th Cir. 1994).

of events" and "often requires analytical dissection of a work and expert testimony."[791] On the other hand, the intrinsic test is a subjective test, which "focuses on the total concept and feel of the two works"[792] and "examines an ordinary person's subjective impressions of the similarities between two works."[793]

In applying the extrinsic test, the courts need to "filter out and disregard non-protectable elements."[794] Thus, proffering a list that "emphasizes random similarities scattered throughout the works" may not support substantial similarity if "any similarities in sequencing are commonplace or, given the plot premise, unremarkable."[795] For example, in *Benay v. Warner Bros. Entm't, Inc.*,[796] both the plaintiff's screenplay and the defendant's film center around an American war veteran who travels to Japan in the 1870s. The Ninth Circuit held the two works not substantially similar after stripping unprotectable elements regarding "historical facts, familiar stock scenes, and characteristics that flow naturally from the works' shared basic plot premise."[797] However, a case may survive the extrinsic test "when considered as a whole the overall selection and sequence of generic elements was substantially similar" even though "the similarities between the works were not individually protectable."[798]

(2) Question of Fact or Question of Law

Though substantial similarity is a question of fact reserved for the jury, the district

---

[791] Rice v. Fox Broad. Co., 330 F.3d 1170, 1179 (9th Cir. 2003).
[792] *Kouf*, 16 F.3d at 1045.
[793] Funky Films, Inc. v. Time Warner Ent. Co., L.P., 462 F.3d 1072, 1077 (9th Cir. 2006).
[794] *Id.* at 1077.
[795] Carlini v. Paramount Pictures Corp., No. 21-55213, 2022 WL 614044, at *2 (9th Cir. Mar. 2, 2022).
[796] Benay v. Warner Bros. Entm't, Inc., 607 F.3d 620, 625 (9th Cir. 2010).
[797] *Id.*
[798] Esplanade Prods., Inc. v. Walt Disney Co., No. CV1702185MWFJCX, 2017 WL 5635024, at *10 (C.D. Cal. July 11, 2017).

courts in the Second Circuit may grant a summary judgment for the defendant, "either because the similarity between two works concerns only 'non-copyrightable elements of the plaintiff's work,' or because no reasonable jury, properly instructed, could find that the two works are substantially similar."[799] Such summary judgment will be affirmed by the Court of Appeals if "the lack of substantial similarity between the protectible aspects of the works was 'so clear as to fall outside the range of disputed fact questions' requiring resolution at trial."[800] Both the "total concept and overall feel" and "commonalities in the works' theme, characters, plot, sequence, pace, and setting" are considered by the courts for making the summary judgment.[801] Some courts even grant the defendant's motion to dismiss for failure to state a claim when two works are not substantially similar.[802]

Similarly, the Ninth Circuit has "frequently affirmed summary judgment in favor of copyright defendants on the issue of substantial similarity"[803] when "no reasonable juror could find substantial similarity of ideas and expression,"[804] or "grant of summary judgment for plaintiff . . . . where works are so overwhelmingly identical that the possibility of independent creation is precluded."[805] The courts only apply the extrinsic test on making the summary judgment,[806] as the intrinsic test "is exclusively the province of the jury."[807] The defendant will lose on summary judgment if a triable factual issue is shown in the

---

[799] Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231, 240 (2d Cir. 1983).
[800] Walker v. Time Life Films, Inc., 784 F.2d 44, 48 (2d Cir. 1986).
[801] Montgomery v. NBC Television, 833 F. App'x 361, 363-364 (2d Cir. 2020).
[802] Abdin v. CBS Broad. Inc., 971 F.3d 57 (2d Cir. 2020); *see also* Esplanade Prods., Inc. v. Walt Disney Co., No. CV1702185MWFJCX, 2017 WL 5635024, at *9 (C.D. Cal. July 11, 2017) (stating "the Complaint identifies only two instances of similar dialogue, which is insufficient to show any extended dialogic similarity between Zootopia, which has a 110-minute running time, and Looney, which proposed running time is unknown.)
[803] Funky Films, Inc. v. Time Warner Ent. Co., L.P., 462 F.3d 1072, 1077 (9th Cir. 2006).
[804] Kouf v. Walt Disney Pictures & Television, 16 F.3d 1042, 1045 (9th Cir. 1994).
[805] Twentieth Century-Fox Film Corp. v. MCA, Inc., 715 F.2d 1327, 1330 (9th Cir. 1983).
[806] *Kouf*, 16 F.3d at 1045.
[807] *Funky Films, Inc*., 462 F.3d at 1077.

extrinsic test, whereas a plaintiff will lose on summary judgment if the extrinsic test is not satisfied. The district court may also dismiss the plaintiff's complaint as a matter of law based on a lack of substantial similarity under Federal Rule of Civil Procedure Rule 12(b)(6).[808]

b. Fictional Works v. Factual Works

The term "substantial similarity" is used by the courts to "strike a delicate balance between the protection to which authors are entitled under an act of Congress and the freedom that exists for all others to create their works outside the area protected against infringement."[809] In doing so, "the degree of substantial similarity required to show infringement varies according to the type of work and the ideas expressed in it."[810]

Fictional works and factual works are usually distinguished in determining substantial similarity. A fictional work has infinite variable expressions for an idea, like a "boy meets girl" idea or an "aliens-attack" idea.[811] Accordingly, "[a] resemblance in details of setting, incident, or characterization that falls short of close paraphrase may be enough to establish substantial similarity and infringement," [812] even though numerous dissimilarities may exist between the two works. In contrast, "a factual work often can choose from only a narrow range of expression."[813] Thus, "similarity of expression may have to amount to verbatim reproduction or very close paraphrasing before a factual work

---

[808] *See* White v. Twentieth Century Fox Corp., 572 F. App'x 475, 476 (9th Cir. 2014); Carlini v. Paramount Pictures Corp., No. 21-55213, 2022 WL 614044 (9th Cir. Mar. 2, 2022).
[809] Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231, 245 (2d Cir. 1983).
[810] Landsberg v. Scrabble Crossword Game Players, Inc., 736 F.2d 485, 488 (9th Cir. 1984).
[811] *See* Williams v. Gaye, 895 F.3d 1106, 1120 (9th Cir. 2018).
[812] *Landsberg*, 736 F.2d at 488.
[813] *Id.*

will be deemed infringed."[814] In sum, a work capable of a wide range of expression is granted "broad copyright protection," while a work capable of a narrow range of expression "merits only thin copyright protection."[815]

For example, in *Universal Pictures Co. v. Harold Lloyd Corp.*, the Court found the defendant Universal Pictures' movie, *So's Your Uncle*, "copied and misappropriated" the "characters, characterization, motivation, treatment, action and sequence of action" of a sequence in the plaintiff Harold Lloyd's movie, *Movie Crazy*.[816] The sequence is a fictional work depicting a man misidentified as another person attending a dinner. He mistakenly put on a magician's coat, and various funny stuff fell out from the coat when he danced with ladies. He was then identified and thrown out of the place. The defendant argued the two movies had numerous dissimilarities in the locale, actors, characters, dialogues, costumes, and purposes. Nonetheless, the Court held that "[t]he whole picture need not be copied to constitute infringement. The mere copying of a major sequence is sufficient . . . . Slight difference and variations will not serve as a defense."[817]

c. Story v. Character

The courts apply different methods to determine the substantial similarity issue of story and character. In determining the infringement of a copyrighted story, the courts compare the "plot, themes, dialogue, mood, setting, pace, characters, and sequence of events" between two works when applying the extrinsic test.[818]

---

[814] *Id.*
[815] See *Williams*, 895 F.3d at 1120.
[816] Universal Pictures Co. v. Harold Lloyd Corp., 162 F.2d 354, 360 (9th Cir. 1947).
[817] *Id.* at 361.
[818] Kouf v. Walt Disney Pictures & Television, 16 F.3d 1042, 1045 (9th Cir. 1994).

For example, in *Funky Films, Inc. v. Time Warner Ent. Co.*,[819] the Ninth Circuit affirmed the district court's summary judgment for the defendant, holding that Time Warner's miniseries *Six Feet Under* did not infringe Funky Films' screenplay of *The Funk Parlor*. Both stories centered around two brothers running a small funeral home inherited from their father, whose romantic life, and competition with a large funeral business. The two works share several similarities. But *The Funk Parlor* was driven by a series of murders, whereas *Six Feet Under* explored relationships and death's meaning. By holding the similarities shared by the two works were mere general plot ideas, *scenes à faire*, or coincidental, the court concluded they were not substantially similar because of "greater, more significant differences and few real similarities at the levels of plot, characters, themes, mood, pace, dialogue, or sequence of events."[820]

By contrast, in determining the infringement of a copyrighted character, the Second Circuit generally considers the visual resemblance, the totality of the character's attributes and traits, the total concept and feel, and the similarities and differences between the two characters at issue.[821] The second comer's potential to evoke the copyrighted character does not necessarily amount to substantial similarity.[822] For example, in *Warner Bros. Inc. v. Am. Broad. Companies, Inc.*,[823] Warner Bros contended Ralph Hinkley, the protagonist in ABC's television series "*The Greatest American Hero*" (*Hero*), infringed Warner Bros' copyrighted character Superman. The Second Circuit affirmed the district court judge's summary judgment for the defendants because the "overall perception of the way Hinkley

---

[819] Funky Films, Inc. v. Time Warner Ent. Co., L.P., 462 F.3d 1072 (9th Cir. 2006).
[820] *Id.* at 1078.
[821] *See* Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231, 241 (2d Cir. 1983); Walker v. Time Life Films, Inc., 784 F.2d 44, 50 (2d Cir. 1986).
[822] *See Warner Bros. Inc.*, 720 F.2d at 242.
[823] *Id.*

looks and acts marks him as a different, non-infringing character who simply has some of the superhuman traits popularized by the Superman character and now widely shared within the superhero genre."[824]

Besides, many characters have a long history of popularity and have been presented in different mediums or series of movies. Certain early materials might have entered the public domain because of the copyright term expiration or lack of required copyright notice if they were first published before January 1, 1978.[825] This raises a unique issue of whether anyone can freely exploit these characters in the public domain materials. The answer is the entitled exploitation cannot do more than what the public-domain materials have presented or in new ways that "comes into conflict with valid copyright" of the later works.[826]

For example, in *Warner Bros. Ent. v. X One X Prods.*,[827] the defendants extracted several movie characters from the public-domain publicity materials like movie posters, lobby cards, still photographs, and press books. These characters include Dorothy, Tin Man, Cowardly Lion; Scarecrow in *The Wizard of Oz*; Scarlett O'Hara and Rhett Butler in *Gone with the Wind*; and Tom and Jerry in the cartoons *Tom & Jerry*. The court held that the defendants could reproduce faithfully in two dimensions the entire image or any one portion of an image from any one item of publicity material. But the defendants were restrained from juxtaposing a public-domain image with a phrase, combining two or more

---

[824] *Id.* at 243.
[825] Warner Bros. Ent. v. X One X Prods., 644 F.3d 584, 593 (8th Cir. 2011) (addressing "the general rule under the 1909 Copyright Act is that a work published in the United States without the statutorily required copyright notice fell into the public domain.")
[826] *Id.* at 596.
[827] *Id.*

separate public-domain images, or transforming any two-dimensional public-domain image into three dimensions. This is because these extra uses will infringe the copyright of the corresponding movie characters' incremental original elements that have not entered the public domain.[828]

2. China

Like the United States, the plaintiff has the burden to prove substantial similarity. The plaintiff must submit a comparison list or a table, demonstrating the two works' similar points.[829] Then the defendant shall explain or proffer counterevidence to each point that he or she objects. Upon motion or necessity, the court may invite an expert to assist the trial or appoint a specialist institution to compare the expression of two works.

With respect to substantial similarity, the court examines whether the expressions (including author's choices, selection, arrangement, and design) between two works constitute substantial similarity disregarding ideas such as theme, inspiration, and emotion.[830] Also, the mainline and sequence of historical facts based on the same historical event, or inevitable incidents, characters, setting, and scenes shared by the same theme, must be excluded.[831] The determination of substantial similarity applies a totality of circumstances analysis,[832] covering: (1) dialogue and aside; (2) setting and relationship of characters; (3) logic arrangement of storylines; (4) same errors in grammatical expression,

---

[828] *Id.* at 603.
[829] *See* Beijing Shi Gaoji Renmin Fayuan Qinhaiz Huzuoquan Anjian Shenli Zhinan (北京市高级人民法院侵害著作权案件审理指南) [Beijing High People's Court Guidelines for the Trial of Copyright Infringement Cases] (promulgated by the Beijing High People's Ct., Apr. 20, 2018) § 10.9.
[830] *See id.* §10.10.
[831] *See id.* §10.11.
[832] *See id.* §10.10.

logical relationship, historical facts or the like; (5) same special design of details; (6) whether the similar parts fall in the prior work's core content; and (7) other factors.

For example, in *Zhang Xiaoyan v. Lei Xianhe*,[833] the two cavalry television series had six identical or similar plots. However, after filtering out the unprotectable historical theme, scènes à faire, and usual terms under the specific context, China Supreme People's Court concluded that the two script were not substantial similar on the whole because:

> [They] had different plots, description emphases, and characters of heroes, and endings, the identical and similar plots accounted for a very low proportion in the two works, such plots were minor in the whole stories, and they were not the main body of Lei's script and they would not make readers and audiences have the identical or similar appreciation experience.

Recently, a new type of copyright infringement has emerged in China as the internet literature becomes popularized and AI writing tools are accessible. Some "writers" allegedly use AI writing tools to produce literature by discomposing and reorganizing numerous prior internet literature.[834] Zhou, Jing's (周静) novel-turned-soap opera, *JinXiu WeiYang* (锦绣未央), is suspected of such a case. According to Beijing Chaoyang District People's Court's twelve judgments (see **Table 4-2**) rendered in 2019, Zhou was found to copy 594 places, 22 plots, nearly a total of 114,000 words from 16 internet novels, including Chen, Wen Wen's (沈文文) *Shen Li Liu Di Chong Bu Shuai* (身历六帝宠不衰),

---

[833] Zhang Xiaoyan Su Lei Xianhe Zhaoqi Sanding Aishuren Yinxiang Tushu Youxian Gongsi (张晓燕诉雷献和、赵琪、山东爱书人音像图书有限公司) [Zhang Xiaoyan v. Lei Xianhe, Zhao Qi, and Shandong Aishuren DVDs & Books Co], 2013 Min Shen Zi No. 1049 (2013 民申字第 1049 号) (Sup. People's Ct. 2014).

[834] Yuan Bo (袁博), *Cong Jinxiu Weiyang Shexian Chaoxi Kan Zhuzuoquan Qinquan Xin Leixing* (从《锦绣未央》涉嫌抄袭看著作权侵权新类型) [From the Suspected Plagiarism of "Jinxiu Weiyang" See New Type of Copyright Infringement], CHINA PRESS PUBL'N RADIO FILM &TELEVISION J (中国新闻出版广电报) (Feb. 23, 2017, 1:45 PM), http://www.ezhicai.com/news/66347.htm [https://perma.cc/D4ZE-NZSW].

Pei, YunFu's (裴云) *Zhong Sheng Zhi Yao Xiang* (重生之药香), Shi Jin's (傅世瑾) *Yi Hu Zhu* (一斛珠), etc. As Zhou's internet novel has three versions of 1.54 million words, 2.7 million words, and 2.87 million words,[835] the 114,000-word illegal copying scattered across Zhou's long novel is unnoticeable.

In *Wen Rui An v. Zhou Jing*,[836] the court stated that copyright infringement can be found at the sentence level. On the one hand, no literary work can stand without borrowing from predecessors. Those idioms, allusions, rhetorical devices, sentence types, and daily use languages repeated in different literary works are in the public domain that no one can monopolize. On the other hand, sentences are the foundation of literature. Those sentences having such unique rhetoric, detailed narrative, or specific description of characters or plots that reflect the author's personality are protected expressions. In determining whether there is copyright infringement on sentences, the court considered sentences' similarity, quantity, and context cohesion on the whole and applied a totality circumstance test rather than observing sentences individually or dissecting them for comparison.

Though Zhou trickily copied certain parts of Wen's sentences and replaced some words, the court held Zhou had illegally copied Wen's sentences at thirteen places. The court held Zhou infringing Wen's copyright protection for sentences in three ways: (1) using the same unique metaphors or adjectives on the same subject matter; (2) using the same or similarly detailed narratives on characters or events; (3) using a large amount of similar language combination and transition in narrating the same historical facts.

---

[835] Wen Ruian Yu Zhoujing Deng Qinhai Zhuzuoquan Jiufen (温瑞安与周静等侵害著作权纠纷) [Wen Rui An v. Zhou Jing], 2017 Jing 0105 Min Chu No.932 (2017 京 0105 民初 62752 号) (Beijing Chaoyang Dist. People's Ct. May 30, 2019) [法宝引证码 CLI.C.83245364].
[836] *Id.*

In addition, the court states that copyright infringement can be found in individual plots. The court holds substantial similarity between plots can be established if the character setup, character relationship, sequence of events, and specific detail designs are basically the same, after filtering out unprotectable ideas, *scènes à faire*, and commonplace elements in the public domain. This is also true even if the alleged infringer uses substitute sentences in narrating the same plot. In this case, the court held Zhou infringed Wen's copyright prediction for three plots.

It is unclear whether *Jin Xiu Wei Yang* (锦绣未央) is really produced by an AI writing tool. Assuming that is true. Then this case highlights the AI writing tool's strength to combine numerous elements extracted from diverse source works into a new work. Still, it also exposes AI's writing tool's weakness that copying clues may leave in specific sentences or plots. Because the new work looks significantly different from any of its source works, it will be easier for the author of a source work to prove substantial similarity in the level of sentences or plots rather than the whole work.

Table 4-2 *Jin Xiu Wei Yang* (锦绣未央) Copyright Infringement Cases

Case Number	Plaintiff	Infringed Work	Fact	Judgment
2019.5.8 （2017）Jing 0105 Minchu No.932	Chen, Wen Wen （沈文文）	Shen Li Liu Di Chong Bu Shuai （身历六帝宠不衰)	Copying 116 places, 2 plots, nearly 30,000 words	Lost profit RMB￥120,000 and reasonable litigation expense RMB￥16,500
2019.5.30 （2017）Jing 0105 Minchu No.937	Pei, Yun （裴云）	Zhong Sheng Zhi Yao Xiang （重生之药香)	Copying 19 places, 1 plot, nearly 7,000 words	Lost profit RMB￥30,000 and reasonable litigation expense RMB￥10,400
2019.5.30	Fu, Shi Jin (傅	Yi Hu Zhu	Copying 27	Lost profit RMB￥

（2017）Jing 0105 Minchu No.953	世瑾)	（一斛珠)	places, 1 plot, nearly 8,000 words	40,000 and reasonable litigation expense RMB￥10,400
2019.5.30 （2017）Jing 0105 Minchu No.959	Zuo, Juan (左娟)	Liu Shui Tiao Tiao (流水迢迢)	Copying 41 places, 3 plots, nearly 12,000 words	Lost profit RMB￥50,000 and reasonable litigation expense RMB￥10,400
2019.5.30 （2017）Jing 0105 Minchu No.970	Wang, Yu Hong （王玉红)	Dou Jin Tang (斗锦堂)	Copying 7 places, 4 plots, nearly 5,000 words	Lost profit RMB￥20,000 and reasonable litigation expense RMB￥10,400
2019.5.30 （2017）Jing 0105 Minchu No.974	Wang, Guo Hua (王国华)	Yan Zhi Lei Zhuang (胭脂泪妆)	Copying 117 places, 1 plot, nearly 7,000 words	Lost profit RMB￥30,000 and reasonable litigation expense RMB￥10,400
2019.5.30 （2017）Jing 0105 Minchu No.982	Cheng, Yun Feng （程云峰)	Shi Hun Zhi Shen Gui Yuan Nu （世婚之深闺怨女)	Copying 9 places, 1 plot, nearly 7,000 words	Lost profit RMB￥30,000 and reasonable litigation expense RMB￥10,400
2019.5.30 （2017）Jing 0105 Minchu No.987	Zhang, Zhi Fan (张之帆)	Yi Shi Wei Chen (一世为臣)	Copying 173 places, 2 plots, nearly 12,000 words	Lost profit RMB￥50,000 and reasonable litigation expense RMB￥10,400
2019.5.30 （2017）Jing 0105 Minchu No.989	Zhu, Xiao Bai (朱笑白)	Shu Nu Sheng Cun Shou Ce （庶女生存手册)	Copying 51 places, 1 plot, nearly 4,000 words	Lost profit RMB￥20,000 and reasonable litigation expense RMB￥10,400
2019.5.30 （2017）Jing 0105 Minchu No.993	Guo, Hui (郭慧)	Xiao Jian Zhi Jiang Shan （啸剑指江山)	Copying 9 places, 2 plots, nearly 7,000 words	Lost profit RMB￥40,000 and reasonable litigation expense RMB￥10,400
2019.5.30 （2017）Jing 0105 Minchu No.1128	Huang, Lin Da (黄琳达)	Gong Zi Wu Chi (公子无耻)	Copying 12 places, 1 plot, nearly 8,000 words	Lost profit RMB￥40,000 and reasonable litigation expense RMB￥10,400

2019.5.30 （2017）Jing 0105 Minchu No.62752	Wen, Rui An (温瑞安)	Wen Rou Yi Dao (温柔一刀) Ji Mo Gao Shou (寂寞高手) Ni Shui Han (逆水寒) Jian Qi Zhang Jiang (剑气长江) Jiang Shan Ru Hua (江山如画)	Copying 13 places, 3 plots, nearly 7,000 words	Lost profit RMB￥ 50,000 and reasonable litigation expense RMB￥ 100,000

3. Taiwan

To establish copyright infringement, the Taiwanese courts follow two steps:[837] (1) determining whether an idea or expression is infringed. Only expression is protected. The expression includes the language, interpretation, treatment, arrangement, and sequence that convey the ideas and facts; (2) determining whether the alleged infringer had access to the copyrighted work and whether the allegedly infringed work constitutes substantial similarity to the protected work in the aspect of expressions.

Like the United States and China, Taiwanese courts make a case-by-case decision. The court uses quantitative and qualitative analysis to examine substantial similarity,[838] and analyzes both the text and non-text elements.[839] The quantitative analysis focuses on what proportion of work is copied. Neither verbatim similarity nor substantial similarity

---

[837] Guanyu Jixie Gufen Youxian Gongsi Weifan Zhuzuoquan Fa Anjian (冠昱機械股份有限公司違反著作權法案件) [Republic of China v. Guan-Yu Machine Co.], 105 Xing Zhi Shang Su Zi No. 42 (105 年度刑智上訴字第 42 號) (Intell. Prop. Ct. Apr. 16, 2020) (Taiwan).; Career Consulting Co., Ltd. v. Yuan Tai Technology, LLC., 108 Min Zhu Shang Zi No. 12, Intellectual Property Court (June 18, 2020).

[838] Longxing Keji Gufen Youxian Gongsi su Xuwei Wenhua Shiye Youxian Gongsi (龍行科技股份有限公司訴徐薇文化事業有限公司) [Long Xing technology Co. v. Xu Wei Culture Co.], 106 Tai Shang Zi No.2673 (106 年度台上字第 2673 號) (Sup. Ct. Nov. 7, 2018) (Taiwan).

[839] Taiwan Nuohua Gufen Youxian Gongsi Su Shengda Huaxue Zhiyao Gufen Youxian Gongsi (台灣諾華股份有限公司訴生達化學製藥股份有限公司) [Novartis (Taiwan) Co. v. Standard Chem. & Pharm. Co.], 105 Min Zhu Shang Zi No.4 (105 年度民著上字第 4 號) (Intell. Prop. Ct. Oct. 20, 2016) (Taiwan).

throughout the entire work is required.[840] The required amount depends on the nature of the work. With regard to factual works, they have features of limited room for fabrication, confined creative and expressive freedom, as well as overlapping information sources. Thus, the court asks for higher substantial similarity to factual works.[841] On the contrary, the court asks for lower substantial similarity to fictional, Sci-Fi, poems and other highly creative works. The qualitative analysis focuses on whether the copied portion is significant. It is sufficient if substantial similarity exists in the copyrighted work's original content. Even if only a tiny portion of the copyrighted work is copied, substantial similarity can still be found if the copied portion is a significant part of the copyrighted work.[842]

Regarding literary works or computer programs, the court applies the abstraction test.[843] First, the court must abstract the work into generality. The work will become increasingly general, along with increased abstraction. The general and highly abstract patterns are considered in the public domain. Second, the court must filter out the general and abstract patterns from the disputed works. Third, the court compares the common characteristics and patterns of the remaining elements. The compared elements must be

---

[840] Longxing Keji Gufen Youxian Gongsi su Xuwei Wenhua Shiye Youxian Gongsi (龍行科技股份有限公司訴徐薇文化事業有限公司) [Long Xing technology Co. v. Xu Wei Culture Co.], 106 Tai Shang Zi No.2673 (106 年度台上字第 2673 號) (Sup. Ct. Nov. 7, 2018) (Taiwan).

[841] *See* Zhiguan Keji Gufen Youxian Gongsi Su Heluo Youxi Youxian Gongsi (智冠科技股份有限公司訴河洛遊戲有限公司) [Soft-World Int'l Co. v. Heluo Games Co.], 108 Min Zhu Shang Zi No.3 (108 年度民著上字第 3 號) (Intell. Prop. Ct. 2d instance CIV. judgment July 30, 2010) (Taiwan).

[842] *Novartis (Taiwan) C*o., 105 Min Zhu Shang Zi No.4.

[843] Tianyi Caijin Keji Fuwu Gufen Youxian Gongsi Su Zhongcheng Zixun Guanli Gufen Youxian Gongsi (天逸財金科技服務股份有限公司訴仲城資訊管理股份有限公司) [Vteam Financial Technology Service Co. v. Fortress System Co.], 102 Min Zhu Shang Zi No. 7 (102 年度民著上字第 7 號) (Intell. Prop. Ct. Dec. 22, 2016) (Taiwan); Guanyu Jixie Gufen Youxian Gongsi Weifan Zhuzuoquan Fa Anjian (冠昱機械股份有限公司違反著作權法案件) [Republic of China v. Guan-Yu Machine Co.], 105 Xing Zhi Shang Su Zi No. 42 (105 年度刑智上訴字第 42 號) (Intell. Prop. Ct. Apr. 16, 2020) (Taiwan); Jiuye Qingbao Zixun Gufen Youxian Gongsi Su Yuantai Shuwei Keji Youxian Gongsi (就業情報資訊股份有限公司訴元太數位科技有限公司) [Career Consulting Co. v. Yuan Tai Tech., LLC.], 108 Min Zhu Shang Zi No. 12 (108 年度民著上字第 12 號) (Intell. Prop. Ct. June 18, 2020) (Taiwan).

sufficiently specific and detailed. Only when quantitative and qualitative substantial similarity exists, and the similarities are not commonplace, does the defendant constitute copyright infringement.

For example, in *Liao, Fu-Bin v. Daishin Music Co.*,[844] the defendant's MTV video *Late Marriage* was alleged to infringe the copyright of the plaintiff's illustrated book *Turn left, Turn Right*. The court abstracted the illustrated book into general layouts and structures as follows: (1) a man and a woman lived closely but in different lifestyles; (2) they felt lonely and did not know each other; (3) they hit it off when they first met in a park; (4) they exchanged phone numbers but separated in a hurry; (5) they could not find each other for some reasons; (6) they missed each other very much; (7) although their lives intersected occasionally, they did not know it; (8) when they couldn't bear the sense of loss and decided to move away, they met again.

The court held the above layout and structure were commonplace among romances and were uncopyrightable ideas. The court further compared the two works' character interactions and the sequence of events. By finding 3 similar places but 7 dissimilar places, the court concluded the character interactions and sequence of events between the two works were not substantially similar. With regard to the three similar places, the court stated they were protected expressions. However, the court held in favor of the defendant because the evidence was insufficient to prove the defendant accessed the plaintiff's work, and there were similar expressions in other sources. It is critical that the court asked for a

---

[844] Liao Fubin Su Daxin Changpian Gufen Youxian Gongsi (廖福彬訴大信唱片股份有限公司) [Liao Fu-Bin v. Daishin Music Co.], 90 Shang Zi No. 1252 (90 年上字第 1252 號) (Taiwan High Ct. May 25, 2004).

high proportion of similarity here because this story was a common romance. A lower proportion of substantial similarity would be sufficient if it were rootless sci-fi.

However, applying analytical and dissectible methods for comparing the substantial similarity of graphical, photographic, artistic, and audiovisual works may be tricky or inappropriate. Therefore, the Taiwan Supreme Court holds special attention shall be paid to the "total concept and feel" of graphical, photographic, artistic, and audiovisual works while applying qualitative analysis.[845] When using the "total concept and feel" test, the court does not dissect the work to compare each detail individually. The judgment criteria are the response and impression of the reasonable general audience instead of the expert testimony of someone with special knowledge or expertise.[846]

For example, in *Sofa Studio Co. v. Engine Studios LLC*,[847] the defendant's animation character *Manui* was alleged to infringe the plaintiff's copyright to an animation character *MuMu*. *Manui* was created by a defendant's employee who also helped to create *MuMu* while he was working for the plaintiff. Though the two characters shared the same idea of anthropomorphic water animals, *MuMu's* prototype was a combination of whale and dolphin, while *Manui's* prototype was a newt. The court found the two characters' skin color, facial features, limbs, size, shape, and proportion were so different that they

---

[845] Jiage Shipin Gufen Youxian Gongsi Su Tongfang Shengwu Keji Gufen Youxian Gongsi (佳格食品股份有限公司訴統芳生物科技股份有限公司) [Standard Foods Co. v. Top Greats Biotech Co.], 103 Tai Shang Zi No.1554 (103 年度台上字第 1544 號) (Sup. Ct. July 31, 2014) (Taiwan).

[846] Meishang Meizhou Rentiantang Gufen Youxian Gongsi Su Jinfu Wanju Shiye Youxian Gongsi (美商美洲任天堂股份有限公司訴金富玩具實業有限公司) [Nintendo Co. v. Jin Fu Toy Ltd.], 110 Min Zhu Shang Zi No. 11 (110 年度民著上字第 11 號) (Intell. Prop. & Com. Ct. Nov. 18, 2021) (Taiwan).

[847] Shouying Chuangyi Gufen Youxian Gongsi Su Yuanjin Guoji Youxian Gongsi (首映創意股份有限公司訴原金國際有限公司) [Sofa Studio Co. v. Engine Studios LLC], 110 Min Zhu Shang Zi No.4 (111 年度民著上字第 4 號) (Intell. Prop. & Com. Ct. Feb. 9, 2023) (Taiwan).

presented an obvious difference in total concept and feel. Therefore, the defendant did not infringe the plaintiff's character copyright.

IV. CONCLUSION

The copyright statutes and judicial precedents share a high commonality among the United States, China, and Taiwan. As discussed above, copyright law protects human original expressions but only requires minimum creativity. The scope of copyright does not extend to elements that do not originate from the author. In copyright infringement cases, the courts must evaluate whether the defendant copied the plaintiff's original expression and whether the two works have substantial similarities. Low originality works (i.e., factual works) require a higher substantial similarity whereas high originality works (i.e., fictional works) require a lower substantial similarity.

In the next chapter, I will apply these principles to examine the new copyright violation risks that a motion picture producer cannot ignore during the data-driven creation process and the new copyright violation risks of the final work. Moreover, the new copyrightability issues and the scope of copyright issues regarding data-driven motion pictures will be addressed.

CHAPTER FIVE: OLD COPYRIGHT ISSUES ARE NEW AGAIN

Confronting with the data-driven trend, there are three copyright issues regarding motion pictures that I will address in this chapter. The first is the new copyright violation risks in the data-driven era, including the risks in the data-driven creation process (data preparation, model training, fine-tuning, prompt) and the risk of final works. The second is the copyrightability issues of data-driven works, including the originality and human authorship requirements. The third is the scope of copyright issues regarding a data-driven work, including the necessity and hardship of ascertaining its scope of copyright.

Only by learning from the past can we know what to do next. I will start by studying how these old copyright topics profoundly influenced today's motion picture industry and then individually analyze the new challenges that a reasonable motion picture maker must be cautious about in the data-driven era. In many cases, I will use two ChatGPT-generated scripts to demonstrate data-driven work's bright future and dark trap. These findings will be the premise for my proposals in the next chapter.

I. OLD COPYRIGHT ISSUES FOR MOTION PICTURES

A. Back to the Beginning – from Photographs to Dramatic Performance

Today, people get used to motion pictures that are long enough to narrate a dramatic or fictional story. But in the late nineteenth and early twentieth centuries, many films were

simply short documentaries of travel or newsworthy events [848] or recordings of performances and acts of vaudeville performers, bodybuilders, dancers, and boxing players. [849] At that time, these "short, attention-grabbing subjects designed to evoke immediate, visceral response from audiences" were considered uncopyrightable because they were "immoral or mere spectacle."[850] Besides, copyright law was blank as to whether the new medium – film – was a proper subject of copyright. Therefore, duping of films, which meant creating a negative from a rival company's positive film and then using the new negative to print more positive films for sale, was a common practice among film companies.[851] Against this historical background, a novel question faced by the courts was whether a motion picture could be a proper subject of copyright as a photograph.[852]

1. *Edison v. Lubin*

Before motion pictures were added to the U.S. copyright law in 1912, a landmark decision, *Edison v. Lubin*, had admitted a motion picture to be protected as a photograph.[853] In this case, Edison sought an injunction restraining Lubin from duping *Christening and Launching Kaiser Wilhelm's Yacht "Meteor,"* a film of an invite-only event that Edison bought the right to make an exclusive recording.[854] The film recorded the ceremony scene from one point where a pivoted camera was placed. Obviously, the film's subject was an uncopyrightable historical fact.

---

[848] *See* DECHERNEY, *supra* note 125, at 63.
[849] *See id.* at 38.
[850] *Id.* at 42.
[851] *See id.* at 19.
[852] *See id.* at 18.
[853] Edison v. Lubin, 122 F. 240, 242 (3d Cir. 1903).
[854] DECHERNEY, *supra* note 125, at 23.

Edison registered a copyright for the film's positive celluloid sheet and placed copyright notices on the front of the copies thereof and at one end of the sheet. Without Edison's permission, Lubin duped 1,500 pictures from the film's 4,500 pictures. Noting that copyright law did not explicitly protect motion pictures, the district court described the film as "an aggregation of photographs" that was unavailable for copyright protection because each picture of the film was not separately registered, and the copyright notice was not placed on each picture.[855] The Third Circuit Court of Appeals reversed such a strict interpretation. The Court of Appeals recognized the film as an advanced type of photograph covered by the statutory language of "any photograph" and admitted the film as a unit being "a single photograph of the whole" in its substance and value.[856] Furthermore, the court held the film to be a work of fine art that "embod[ied] artistic conception and expression" because it required "a study of lights, shadows, general surroundings, and a vantage point adapted to securing the entire effect."[857]

The above precedent clearly recognized that a motion picture possessed an inherent copyrightable character independent from its subject. In other words, a motion picture can be duly copyrighted even though its subject is an uncopyrightable fact. But after the courts outlawed duping of films and reality-based moving photographs gradually developed into a new form of drama in 1903,[858] things started to change. A new type of piracy, remaking a competitor's film shot by shot to create a new version of its own, became more common

---

[855] Edison v. Lubin, 119 F. 993 (C.C.E.D. Pa. 1903).
[856] *Edison*, 122 F. at 242.
[857] *Id.*
[858] *See* DECHERNEY, *supra* note 125, at 34.

among industrial participants.[859] Consequently, the following question submitted to the courts was whether remaking a film without permission constituted copyright infringement.

2. *Biograph v. Edison*

In *American Mutoscope & Biograph Co. v. Edison Mfg. Co.*,[860] Edison's film, *How a French Nobleman Got a Wife Through the New York Herald Personal Columns*, was alleged to infringe Biograph's film, *Personal*. The two films depicted the same plot about a crowd of women who read a wealthy young man's personal ad in a newspaper running after the man across the countryside. Both films were duly registered for copyright. Edison's photographer denied copying *Personal* as to the characters' costumes, expressions, figures, bearings, poses, gestures, postures, and actions.

Without argument or discussion about whether the film subject was an original story, the Court found Biograph's film was a copyrighted photograph by reaffirming a photograph was a proper subject of copyright if it were "not only a light-written picture of some object, but also an expression of an idea, or thought, or conception of the one who takes it."[861] The Court also confirmed a photograph "may be copyrighted to the same extent that any literary composition expressive of an author's ideas or conceptions may be copyrighted" no matter whether it was taken by a camera placed at one point or at numerous points.[862] However, it refused to grant a preliminary injunction for Biograph. The Court held that Biograph had not proved unlawful copying because the two films' titles, shooting

---

[859] *See Id.* at 63.
[860] Am. Mutoscope & Biograph Co. v. Edison Mfg. Co., 137 F. 262 (C.C.D.N.J. 1905).
[861] *Id.* at 265.
[862] *Id.* at 266.

times, actors, and costumes were different, although the two films were so similar in scenes, plots, and shooting places.[863]

The generous attitude of *Biograph v. Edison* toward film remaking provided a free environment for the young film industry to exchange creativity.[864] Accordingly, a great masterpiece can come from the inspiration of numerous remade versions. However, as films became longer and capable of telling a more complex story after 1907, many filmmakers began to adapt contemporary literary works and stage plays "in terms of both narrative detail and psychological depth."[865] This transformation quickly induced the concerns of publishers and theater producers. Ultimately, publishers and theater producers brought the issue before the Supreme Court to determine whether a motion picture adapted from a novel would violate the novel's copyright.

3. *Kalem Co. v. Harper Bros.*

In *Kalem Co. v. Harper Bros.*,[866] the Supreme Court held that a motion picture could infringe on the dramatizing right of a literary work for the first time. This case involved a copyright infringement action filed by Harper & Bros., a prominent publisher owning the copyright of Lew Wallace's authored novel called "Ben Hur," and Klaw & Erlanger, a major Broadway producer owning the novel's theatrical right, as well as Henry Wallace, the novel author's heir.[867] The defendant was Kalem Company, a motion picture producer and a member of Edison Trust. Kalem hired a writer to make a script based on

---

[863] *Id.* at 267

[864] *See* DECHERNEY, *supra* note 125, at 66-67.

[865] *Id.* at 43.

[866] Kalem Co. v. Harper Bros., 222 U.S. 55 (1911).

[867] *See* DECHERNEY, *supra* note 125, at 45.

"Ben Hur" and produced a 15-minute film exhibiting the novel's certain scenes. Kalem gave the film to exhibitors for public exhibitions and advertised it under the novel's title.

Justice Holmes confirmed the lower court's decision that the exhibition of a motion picture might infringe on the author's right to dramatize his novel. He pointed out that "[a]ction can tell a story, display all the most vivid relations between men, and depict every kind of human emotion, without the aid of a word."[868] Likening a mirror directly reflecting real men's actions on the stage, the moving pictures just exhibited a less vivid visual impression in a more complex way.[869] Even if the motion pictures were duly copyrighted photographs, they might infringe on the author's rights, just as an innocent mirror might be used for unlawful purposes.[870]

Although Kalem argued that it did not exhibit the films by itself, Justice Holmes held it "was liable to not only expected but invoked by advertisement the use of its films for dramatic reproduction of the story. That was the most conspicuous purpose for which they could be used and the one for which especially they were made."[871] Justice Holmes denied such interpretation would extend copyright to ideas because Congress definitely conferred the authors the exclusive rights to dramatic reproduction of their works.[872]

B. CONGLOMERATION OF STORIES INTO MOTION PICTURES

After the *Kalem* decision was handed down, the film companies quickly realized two things. First, they must protect themselves from copyright infringement claims that

---

[868] *Kalem Co.*, 222 U.S. at 61.
[869] *Id.* at 61–62.
[870] *Id.* at 62.
[871] *Id.* at 62–63.
[872] *Id.* at 63.

may come from industrial competitors and countless writers. Second, they needed to collect

more dramatizing rights to obtain a monopoly in the film marketplace.

1. Controlling Copyright Violation Risks

a. Industrial Measures

Several measures were quickly taken by film companies to protect themselves from

copyright claims.[873] For example, Edison Co. had a full-time team to check if the proposed

scripts infringed the copyrights of stories in magazines, books, and papers. It made

agreements with major publishers and well-known writers to fix the sums that it would pay

if infringement unfortunately happened. It also requested screenwriters to sign a standard

warranty clause when assigned their stories like below:[874]

> Said author warrants that he has the full right to convey the interest herein
> assigned; that he has not executed and will not execute any agreement in
> conflict herewith; that the said dramatic composition is original with him
> and that no incident therein described is, to the best of the author's
> knowledge and belief, the same as or any colorable imitation of any incident
> in a copyrighted book or play.

However, not all proposed scripts can be luckily picked by the studios to make

motion pictures. For example, only four percent of scripts sent to Edison Co. were selected

in 1912.[875] In the 1990s, more and more disappointed writers sued studios for infringing

copyrights without compensation.[876] The frequency and number of copyright lawsuits from

the angry writers taught the studios not to read unsolicited scripts unless they were

---

[873] *Townsend Copyright Amendment: Complete File of Arguments Before the Committee on Patents, House of Representatives, on H.R. 15263 and H.R. 20596, Commencing January 24, 1912*, 62nd Cong. 9 (1912) (statement of Frank L. Dwyer, President of The Edison Electric Co.).
[874] *Id.*
[875] *Id.*
[876] 1 Entertainment Industry Contracts P 2.04 (2023) (LEXIS).

submitted through an agent and accompanied by a submission agreement that releases the studios from copyright lawsuit risks.[877]

Securing a policy of errors and omissions (E&O) insurance is another crucial measure for minimizing copyright violation risks.[878] This insurance may cover high legal expenses caused by defending and/or indemnifying potential copyright claims. A typical E&O insurance covers one million for one claim or three million for aggregated claims with a deductible of $25,000 for a term of three years.[879] A motion picture without E&O coverage would be hard to sell to distributors and exhibitors.[880]

To evaluate the copyright litigation risk, the motion picture producer is asked to complete an application form disclosing all the information to the insurer.[881] Providing a title report and a copyright report is inevitable. The former indicates clearance of the motion picture's title. The latter indicates the clearance of the motion picture's source materials.[882] The motion picture producer has to follow all the clearance procedures with a complex checklist.[883] Experienced clearance attorneys and copyright research companies are hired to complete the documentation and procedures. All clearance must be done before the insured motion picture's first exhibition.[884]

b. Legislative Measure

Admitting industrial measures could not prevent all possible copyright violations

---

[877] *Id.*; MICHAEL C DONALDSON, CLEARANCE & COPYRIGHT: EVERYTHING THE INDEPENDENT FILMMAKER NEEDS TO KNOW 118-119 (2d ed. 2003).

[878] *See* EVE LIGHT HONTHANER, THE COMPLETE FILM PRODUCTION HANDBOOK 127 (4th ed. 2010).

[879] 2 Entertainment Industry Contracts Form 15-16 (2023) (LEXIS).

[880] *Id.*

[881] *Id.*

[882] *Id.*

[883] 1 Entertainment Industry Contracts P 2.02 (2023) (LEXIS).

[884] FORM 15-16, *supra* note 879.

because of the prevalence of copyrighted materials, Edison Co. pleaded with House Representative Edward Waterman Townsend to introduce bills to limit the statutory damage that an innocent filmmaking infringer would have to pay.[885] After several hearings, the proposals were integrated into a bill and became law on August 24, 1912.[886] The so-called Townsend Amendment added motion pictures to the category of protected works on the one hand and limited statutory damages that could be claimed against innocent motion picture studios on the other hand.

The motion picture industry was particularly concerned about statutory damages because such damages were calculated based on the times of infringements. Under the prior Copyright Act of 1909, in the case of infringing dramatic or dramatic-musical works, it was $100 for the first and $50 for the subsequent infringing performance; in the case of infringing undramatized or nondramatic works, it was $1 dollar for every copy.[887] As copies of motion pictures were released to nationwide exhibitors simultaneously for exhibiting several times a day, an infringed copyright owner might be able to claim very high statutory damages in a short period without needing to prove any actual damages and the infringers' real profits.[888]

To address such concerns, the 1912 amendment of the Copyright Act set new limitations for statutory damages against motion picture makers who were not aware of infringing and could not have reasonably foreseen the infringement.[889] The limitation

---

[885] H.R. 15263, 62nd Cong. (1911); H.R. 20596, 62nd Cong. (1912).
[886] Motion Picture Copyright Amendments, Pub. L. 62-303, 37 Stat. 488.
[887] Copyright Act of 1909, Pub. L.60-349, 35 Stat. 1075.
[888] H.R. REP. No. 62-756, at 3-4 (1912).
[889] Motion Picture Copyright Amendments, *supra* note 886.

amounted from $50 to $100 for undramatized or nondramatic works and $250 to $5,000

for dramatic or dramatic-musical works. The House report of the bill clearly explained:[890]

> The new limitations do not operate as "compulsory license." They operate only in cases of innocent infringement. The protection terminates with notice, and thereafter the full arbitrary sums are recoverable, This fact and the great expense of making the films, the right of the copyright proprietor to recover actual damages and profits, and the right to impound and destroy all infringing articles and all devices used in their production, are believed to be sufficient to induce the makers of films to continue to use all diligence to avoid any and all infringements.

However, the new limitations seemed not to help a lot for motion picture makers.

Forty-six years after the Amendment was enacted, there have been no reported cases

because authors were inclined to sue motion picture studios for actual damages or

profits.[891] Nevertheless, the Copyright Act 1976 maintains the limitation on statutory

damages against innocent infringers with new amounts and extends this limitation to all

innocent infringers regardless of the type of infringing works.[892]

2. Expanding Monopoly to Original Stories

a. Assignment and License

One quick way to monopolize the dramatizing rights of literary works or dramatic

works was forming an exclusive alliance with publishers and theaters.[893] Cross-media

monopolies contributed to the establishment of large studios that were keen on making

franchises, series, and authorized adaptations.[894] New and independent film companies,

---

[890] H.R. REP. No. 62-756, at 4 (1912).
[891] Ralph S. Brown, *Jr, The Operation of the Damage Provisions of the Copyright Law: An Exploratory Study*, 23 COPYRIGHT LAW REVISION STUDY 81 (1958).
[892] 17 U.S.C § 504(c).
[893] *See* DECHERNEY, *supra* note 125, at 55.
[894] *Id.*

such as Paramount and Universal, did this way much better than Edison and Biograph and ultimately displaced those ever film moguls.

To save money, film companies also purchased many not yet known stories from a new class of writers called "scenario writers" springing up in the country,[895] and hence stimulated the birth of screenplay schools.[896] In 1912, the Edison Co. received about 100 proposals a week but was only able to use 4 percent of them. [897] To be assured of monopolies on stories, film companies preferred original stories rather than stories in the public domain. Consequently, films became longer in order to illustrate these new stories to audiences who had never known about these stories.[898]

b. Works Made for Hire

In addition to shopping into the option of preexisting works' dramatizing rights, the film companies also hired writers to create their own original stories. Under the work-made-for-hire doctrine, film companies were not only legal authors but also copyright owners of those original stories created by their hired writers. Work-made-for-hire is even a more favorable copyright arrangement than assignments or licenses for film companies because the actual creators do not have the renewal rights under the Copyright Act of 1909 §24 and the termination right under the Copyright Act of 1976 §203, 304.[899]

Confronting with the industrial moguls' exploitation, several writer clubs arose in

---

[895] *Townsend Copyright Amendment: Complete File of Arguments Before the Committee on Patents, House of Representatives, on H.R. 15263 and H.R. 20596, Commencing January 24, 1912*, 62nd Cong. 9 (1912) (statement of Frank L. Dwyer, President of The Edison Electric Co.).
[896] *See* DECHERNEY, *supra* note 125, at 56.
[897] Statement of Frank L. Dwyer, *supra* note 895.
[898] *See* DECHERNEY, *supra* note 125, at 56.
[899] *See* NIKOLAUS REBER, FILM COPYRIGHT, CONTRACTS AND PROFIT PARTICIPATION 68, 76 (2000).

1913 to struggle for writers' interests.[900] A later formed writer group called the Screen Writers' Guild (SGA) transformed to be a labor union in 1933 and then became today's Writers Guild of America, East (WGAE) and the Writers Guild of America West (WGAW) after 1954.[901] WGAE and WGAW jointly represent writers of motion pictures, television, radio, and internet programming to negotiate collective bargaining agreement with the Alliance of Motion Picture and Television Producers, Inc.[902]

II. NEW COPYRIGHT VIOLATION ISSUES FOR DATA-DRIVEN WORKS

As discussed above, after the *Kalem* case, the two most important copyright lessons for motion picture studios are averting infringement risks and ensuring a monopoly on original stories. Thus, a reasonable motion picture producer will not recklessly jump into the data-driven current without addressing the two issues.

Before further discussion, I would like to show AI's state of the art writing a film script. I used the popular free version ChatGPT to generate a film script on June 27, 2023 (see **Appendix C**).[903] Suppose data science told me that the audience would be interested in a romance between law students during the COVID pandemic. I instructed ChatGPT to

---

[900] Hilary Swett, *The Screen Writers' Guild: An Early History of the Writers Guild of America*, WRITERS GUILD FOUNDATION (2020), https://www.wgfoundation.org/screenwritersguild-history [https://perma.cc/S3GZ-7VP5].

[901] A HISTORY OF WGA CONTRACT NEGOTIATIONS AND GAINS, https://www.wga.org/the-guild/about-us/history/a-history-of-wga-contract-negotiations-and-gains [https://perma.cc/RN27-D2ES].

[902] *See* Press Release, Joint Statement from WGA & DGA Presidents Regarding Max Credit Changes (May 24, 2023), https://www.wga.org/news-events/news/press/joint-statement-from-wga-and-dga-presidents-regarding-max-credit-changes [https://perma.cc/7L4E-N3WQ].

[903] ChatGPT was an application of the Large Language Model (LLM) developed by OpenAI and released at the end of 2022. It can perform natural language tasks according to a user's prompt. *See* INTRODUCING CHATGPT, *supra* note 422.

"*Write a film script based on the words: law school students, online bar exam, covid pandemic, love.*" Except for the prompt, I did not limit the number of words or give any other instructions. Appendix C is a 625-word script that ChatGPT generated in a few seconds after I clicked the enter button.[904] This script looks like being written by a human. I would not have known it was written by a machine if I were not the user. But I dislike the title, *Virtual Verdicts*. Therefore, I instructed ChatGPT to revise the title three times until I was satisfied with the last one, *Trials of Passion*.[905]
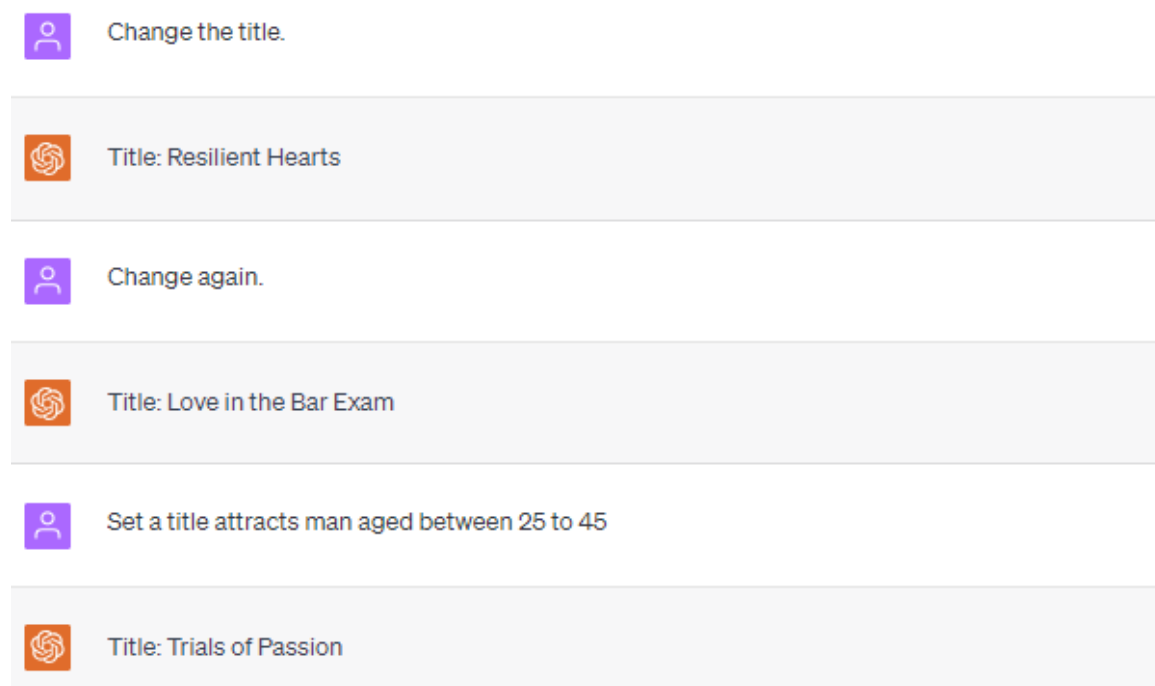
Although AI can write like a human writer, a reasonable motion picture producer will not randomly throw out a prompt to ChatGPT to generate a script and produce a motion picture accordingly. He will need an effective model(s) that helps him: (1) predict the

---

[904] Virtual Verdicts (June 27, 2023), https://chat.openai.com/share/2c110e76-1294-42f1-9d00-e19def1c0c4a [https://perma.cc/T8NW-W6YT]

[905] Virtual Verdicts (changing title) (June 27, 2023), https://chat.openai.com/share/553e5cde-a0b7-497a-853f-50c683385405 [https://perma.cc/56PE-XTKN].

patterns of hit motion pictures, and (2) make a script based on the predicted patterns. To achieve these goals, a large amount of data must be used in diverse ways for different purposes at various stages. First, I will focus on the potential copyright violation issues on data uses, including data preparation, model building, model fine-tuning, and prompting.

A. COPYRIGHT VIOLATION RISKS IN DATA-DRIVEN CREATION PROCESS

1. Data Preparation

Data tells what has ended and what is ongoing. From learning the past, we can predict what may happen. Whether humans or existing models will make predictions or if we want to create a new model for making predictions, the first steps are all the same: collecting, managing, and storing data. Suppose we have decided to build a model to predict motion pictures' success rate. We need two data sets:[906] (1) a **training data set** to teach the model what hit motion pictures are. (2) a **test set** to test the model's actual performance to predict hit motion pictures. An overall dataset is usually split into two sets: one-third for training and two-thirds for testing. If there is more than one model to select, (3) a **validation set** (or development set) is needed to validate each candidate model's performance and choose the model that performs best. [907]

What types of data sets are needed depends on the problem to solve and the type of models are selected.[908] For predicting movies' chance of success, titles, data set about budgets, box offices, reviews, writers, directors, actors, audience demographic, viewing

---

[906] KOTU, *supra* note 345, at 30.
[907] STUART J. RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 666 (4th ed. 2021).
[908] *Id.* at 704-705.

histories, countries, and release time may be important. For generating scripts, a text data set is proper. For generating images, an image data set is proper. Not every data set is copyrighted. And not all data preparation activity requires copyright licenses unless the activity involves the reproduction of copyrighted materials. Whether a data-preparing activity constitutes reproduction is a technical and legal question. It must be determined case by case:

a. Hypothetical Case 1: Spreadsheet (Low risk)

To discover the pattern of movie features, John Doe prepares a spreadsheet listing hundreds of attributes (features) that he has already known and sends out to investigators to complete the questionnaire (see below).[909] New features can be added. The investigators complete the spreadsheet by watching movies in cinemas, on televisions or streaming platforms; buying or renting authorized DVDs; borrowing authorized DVDs from libraries; and buying or renting movies on Amazon. There is no reproduction involved.

Movie	Attribute 1	Attribute 2	Attribute 3	Attribute 4	…….	Attribute n
Movie 1	x		x	x	…….	…….
Movie 2	x	x			…….	…….
Movie 3	x		x	x	…….	…….
Movie 4		x		x	…….	…….
…….	……	……	…….	…….	…….	…….
Movie n	…….	…….	…….	…….	…….	…….

---

[909] A dataset with hundreds to thousands of attributes is common in data science. Some algorithms only accept a data set in numeric and normalized in rows and columns. *See* KOTU, *supra* note 345, at 23-28.

b. Hypothetical Case 2: Watching Robot (Low risk)

Jane Doe collects movie data by using a robot that can mimic humans' watching behavior to "watch" movies played on televisions or streaming platforms and can summarize the movies' ideas automatically and instantaneously. But this Robot does not "memorize" the movies except for the summary.

In *Cartoon Network LP, LLLP v. CSC Holdings, Inc.*,[910] the Second Circuit Court of Appeals held that a copy must satisfy both "embodiment" and "duration" requirements to infringe reproduction rights. In cases 1 and 2, the embodiment requirement is not satisfied because the motion pictures are not embodied on a medium.

c. Hypothetical Case 3: Screenshot Robot (Low to Medium risk)

Jane Doe collects data using a robot that can summarize the movie's ideas. Slowing down the video playing speed may be necessary to match the Robot's computational capabilities. The Robot continuously takes screenshots of the movie on play and selects the keyframes to analyze and summarize the image of each shot.[911] Those shots stay on the Robot's RAM for 2 seconds and are overwritten automatically and permanently as soon as it is processed.

In *Cartoon Network LP, LLLP v. CSC Holdings, Inc.*,[912] the Second Circuit Court of Appeals concluded that data residing in buffers less than 1.2 seconds was not a copy because the transitory duration is too short.[913] In case 3, although the shots stay on the

---

[910] Cartoon Network LP, LLLP v. CSC Holdings, Inc., 536 F.3d 121, 127 (2d Cir. 2008).
[911] This description is not a full fiction. A new area called video mining has emerged to develop models and techniques to browse, index, skim, summarize, and understand video content. *See* VIDEO MINING (Azriel Rosenfeld et al. eds., 2003).
[912] *Cartoon Network LP*, 536 F.3d at 127.
[913] *Id.* at 130.

RAM for more than 1.2 seconds, the courts probably will accept an argument that the difference between 1.2 seconds and 2 seconds is not significant. Therefore, the copyright violation risk is low to medium.

d. Hypothetical Case 4: Deformed Script (High Risk)

John Doe collects hundreds of copyrighted scripts online. He converts the scripts from an unstructured text format to a semi-structured data set so that they can be used for training models. [914] The semi-structured data set is not meaningful to humans, but algorithms can recognize the patterns, frequencies, and relationships between words and therefore perform language tasks.

In *White-Smith Music Pub. Co. v. Apollo Co.*,[915] the Supreme Court held that perforated rolls were not copies of musical composition because they were parts of a machine, not intended to and could not enable human eyes to read as a written musical composition, even though the rolls could reproduce the same sound as playing the musical composition. However, this precedent has been overruled by section 1(e) of the Copyright Act of 1909[916] and the later section 102(a) of the Copyright Act 1976.[917] Therefore, the deformed scripts are copies if they can be communicated to humans with the aid of a machine or device,

e. Hypothetical Case 5: Eidetic Memory Robot (High risk)

Jane Doe collects data by using a machine that not only can mimic humans'

---

[914] *See* KOTU, *supra* note 345, at. 283.
[915] White-Smith Music Pub. Co. v. Apollo Co., 209 U.S. 1, 18 (1908).
[916] Copyright Act of 1909, Pub. L.60-349, 35 Stat. 1075.
[917] "Copyright protection subsists . . . . in original works of authorship fixed in any tangible medium of expression, . . . can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device." 17 U.S.C. § 102(a).

watching behavior to "watch" movies on televisions or the internet but also has a highly retentive memory. Thereafter, the machine can answer any questions about the movies' content upon Jane's request except for playing the movies.

No matter what technology is used for the function, Jane commits a reproduction in the technical sense and the legal sense. Jane would be liable for unauthorized reproduction unless fair use is met.

f. Hypothetical Case 6: Database (High risk)

John Doe reproduces and stores copyrighted movies into a cloud-based database for data mining use without permission. He would infringe copyrights unless complying with fair use.

It is critical to note that once a data-preparing activity has satisfied the requirements of "embodiment" and "duration," the infringer cannot contend that his reproduction is only transitory steps en route to a non-infringing work. In *Walt Disney Prods. v. Filmation Assocs.*, the court held that Defendant *Filmation's* script, storyboard, story reel, etc., had infringed Walt Disney's copyright even its film was not final "[b]ecause the right of reproduction affords a copyright owner protection against an infringer even if he does not also infringe the § 106(3) right of distribution."[918] Therefore, in the cases of example 5 and 6, data preparation is not a valid defense against unauthorized reproduction.

2. Model Training

Model training is a required procedure for machine learning. Learning means the

---

[918] Walt Disney Prods. v. Filmation Assocs., 628 F. Supp. 871, 876 (C.D. Cal. 1986).

model can optimize its performance by being trained by data, i.e., past examples.[919] Unlike ordinary computer programs, learning algorithms can perform various tasks by adjusting the values of model internal parameters to cope with the best performance benchmark. After repetitive training on many examples, the model's parameters are gradually calibrated to the best mode.

It is unclear how the training data is processed and whether a copy is made during the training process. Does the algorithm only read, browse, and visit a database (not copying), or download, scan, and print the data to itself (copying) when it is trained on the data sets?[920] Does it only extract the underlying abstract patterns (idea) or retain the whole contents (expression)? Can it provide or recompose the whole content upon request later? No matter how the training is processed, if the training datasets stay on the model for more than a transitory duration or remain on the model and are available for communicating to users upon requests, this activity will meet the embodiment and duration requirements and thus constitutes reproduction under the meaning of copyright law.

For example, GPT-3 is trained on five massive text datasets total of 300 billion tokens, including Common Crawl (almost a trillion words collected from web pages), WebText2 (extracted from Reddit submissions), Book 1 and Book 2 (two internet-based books corpora), and English version Wikipedia.[921] The GPT series models are "trained to predict the next word in a document."[922] Though OpenAI says their training data are

[919] ETHEM ALPAYDIN, MACHINE LEARNING 27-28 (rev. & updated ed. 2021).
[920] Hillel Aron, *Microsoft and GitHub ask court to scrap lawsuit over AI-powered CoPilot*, COURTHOUSE NEWS (May 4, 2023), https://www.courthousenews.com/microsoft-and-github-ask-court-to-scrap-lawsuit-over-ai-powered-copilot/[https://perma.cc/E5ML-V98Y].
[921] Brown, *supra* note 410 at 8-9.
[922] GPT-4, OPENAI (March 14, 2023), https://openai.com/research/gpt-4 [https://perma.cc/H4HD-GA96].

"publicly available data (such as internet data) as well as data they have licensed," it does not immune from copyright violation lawsuits.

On June 28, 2023, two writers, Paul Tremblay and Mona Awad, brought a class action on behalf of book authors against OpenAI in California Northern District Court.[923] OpenAI was sued for copyright violation by reproducing Plaintiffs' books' expressive information and distributing the expressive information by ChatGPT's outputs. Plaintiffs alleged that GPT-3 used training datasets in violation of copyright. Particularly, the training datasets allegedly included a Book Corpus consisting of 7,000 free, but most were copyrighted novels available online. Besides, GPT-3's training datasets book21 and book2 allegedly contained 63,000 titles and 294,000 titles that may have come from unauthorized sources because no such large-size dataset was known. Plaintiff alleged OpenAI's copy was inferred by ChatGPT's capability to correctly summarize the Plaintiffs' books and mimic individual writers' styles to respond to ChatGPT user's prompts.

If motion picture makers use ChatGPT to generate scripts and such scripts contain copyrighted materials, they may infringe on reproduction and adaptation rights.[924] Any part of the scripts that "unlawfully" uses preexisting material cannot be copyrighted.[925]

3. Model Fine-Tuning

LLMs built for general language tasks are not good at special tasks that require creativity and authenticity. For example, writing poetry and song lyrics.[926] Like how an

[923] Tremblay v. OPENAI, INC., 3:23-cv-03223 (N.D. Cal. 2023). *See* Lauren Berg, *ChatGPT Is 'Trained' Using Copyrighted Books, Authors Say*, Law360 (June 29, 2023, 9:56 PM EDT), https://www.law360.com/articles/1694688/chatgpt-is-trained-using-copyrighted-books-authors-say [https://perma.cc/QG8A-EH9Q].
[924] 17 U.S.C. § 106.
[925] *Id.* § 103(a).
[926] Jooyoung Lee, et al., *Do Language Models Plagiarize?* WWW '23: PROCEEDINGS OF THE ACM WEB CONF. 3637, 3638 (2023).

ordinary person needs special education to be an expert in small domains, a standard approach is post-train LLMs for specific tasks. This post-training process is called Fine-Tuning.[927]

Fine-tuning requires specific datasets that may not be available to the public. Typically, LLMs are fine-tuned by a new tasks-oriented supervised dataset that includes thousands to hundreds of thousands of labeled examples in a specific domain.[928] The dataset used in the fine-tuning process has the same copyright violation concerns if copyrighted materials are used, and the person or entity performing the Fine-Tuning is liable for copyright infringement. For example, if a motion picture studio fine-tunes GPT-4 for writing film scripts with hundreds of thousands of copyrighted scripts without permission, they may infringe on adaptation rights.[929] Any part of the scripts that "unlawfully" use preexisting material cannot be copyrighted.[930]

4. Prompt

A user's instruction to a machine is called a prompt.[931] For example, ChatGPT accepts a user's prompt in text format to generate text output.[932] GPT-4 can accept prompts in both image and text formats to generate text output.[933] Midjourney can accept prompts in text format to produce an image output.[934] Different models have different limits to the

---

[927] *Id.*
[928] Brown, *supra* note 410 at 6.
[929] 17 U.S.C. § 106.
[930] *Id.* § 103(a).
[931] Floridi, *supra* note 409 at 684.
[932] *See* INTRODUCING CHATGPT, *supra* note 422.
[933] *See* OpenAI, *supra* note 424.
[934] MIDJOURNEY, PROMPTS, https://docs.midjourney.com/docs/prompts[https://perma.cc/4J29-YUJZ].

prompt's length.[935]

Machine is designed to respond to the user's prompts as best as possible. The machine's output may rewrite or contain the prompt's original content. If a machine uses a copyright-infringing prompt to generate an output, the user will be liable for copyright violation. For example, entering the text of a novel as a prompt to instruct ChatGPT to generate a film script. The user is infringing the novel author's adaptation right. Any part of the scripts that "unlawfully" use preexisting material cannot be copyrighted.[936]

B. COPYRIGHT VIOLATION RISKS FOR FINAL DATA-DRIVEN WORKS

Following the previous section which focuses on the copyright violation risks involved in creation process activities, this section will discuss the copyright violation risks on the final work. "Copying" and "substantial similarity" elements are studied below.

1. Copying Concerns

a. Copy from Training Data into the Output

Many people wonder if machines just copy and paste something that has existed.[937] For this question, we have two reasons to assume that AI does more than copy and paste. First, there is no reasonable purpose to invest a large amount of money and effort to create a machine that can only copy old things. There are many cheaper and easier substitute means, e.g., photocopy machines, scan machines, traditional database systems, and the

---

[935] 100 tokens ~= 75 words. Raf, *What are tokens and how to count them?*, OPENAI, https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them[https://perma.cc/3X6Z-TEWT]; Michael Mrbullwinkle & Eric Urban, *Learn how to work with the ChatGPT and GPT-4 models*, AZURE (May 05, 2023), https://learn.microsoft.com/en-us/azure/cognitive-services/openai/how-to/chatgpt?pivots=programming-language-chat-completions [https://perma.cc/U4G5-8W25].
[936] 17 U.S.C. § 103(a).
[937] *See* Boden, *supra* note 58 at 224, 224.

Google search engine. Second, all the AI-generated artworks that have been publicly reported, displayed, or for sale should have been already suited for unlawful reproduction. If they have not been sued, that may be because they only reproduced something in the public domain, or those artworks are worthless. For example, ShortlyAI, a downstream AI writing tool connected to GPT-3, asserts that ShortlyAI's output is completely original and passes all common plagiarism checkers.[938]

However, nothing is absolute. A study shows GPT-3's predecessor, GPT-2, doing more than copying and pasting training data, but also finds GPT-2 plagiarizing training data in three ways: verbatim copying, paraphrasing, and mimicking ideas.[939] It is worth noting that verbatim copying and paraphrasing are no different under the sense of copyright if a protected expression has been "used" whether it has been quoted.[940]

This study indicates LLM's size and its training data's similarity degree are positively correlated with plagiarism rates.[941] LLM's memorization effect has become a significant concern.[942] A website has warned that ChatGPT's output may occasionally contain training data.[943] Users are suggested to verify the content and use plagiarism checkers.

In November 2022, thousands of anonymous computer programmers filed a lawsuit

---

[938] FREQUENTLY ASKED QUESTIONS, SHORTLYAI, https://www.shortlyai.com/pricing#faqs[https://perma.cc/93R2-8SG4].

[939] Lee, *supra* note 926, at 3638.

[940] Salinger v. Random House, Inc., 811 F.2d 90, 98 (2d Cir.), opinion supplemented on denial of reh'g, 818 F.2d 252 (2d Cir. 1987).

[941] Lee, *supra* note 926.

[942] Stella Biderman et al., *Emergent and Predictable Memorization in Large Language Models*, ARXIV: 2304.11158v2 (2023).

[943] Govind Dheda, *Is Chat GPT Plagiarism? Unveiling the Truth Behind Generated Content*, OPENAIMASTER (May 24, 2023), https://openaimaster.com/is-chat-gpt-plagiarism/ [https://perma.cc/F63J-HAL3].

in California Northern District Court against OpenAI and Github.[944] GiHub is the largest internet software projects hosting service provider. Plaintiffs published their copyrighted codes on GitHub's public repositories and licensed anyone to use them with proper attribution to copyright owners. In 2021, OpenAI collaborated with Github to release Copilot (AI coding tools) and Codex (AI coding tools with natural language). Copilot and Codex were trained on billions of lines of code, including codes stored in public GitHub repositories.

Plaintiffs alleged that Copilot generated codes from the training data about 1%, but GitHub and OpenAI violated their copyright licenses for not including any attribution to Plaintiffs. The May 11, 2023, order of Judge Tigar recognizes the Plaintiffs' standing to seek injunction relief on the claim of injury to property rights because the Plaintiffs plausibly alleged a realistic danger that Codex and Copilot would reproduce Plaintiff's licensed codes as output.[945] This case's subsequent development deserves our attention.

b. Copy from Previous Output into New Output

If GPT-3 would copy training data, it may copy its previous output as well. This may cause trouble between two users that happen to give similar prompts. OpenAI says GPT-4 does not learn from its experience.[946] However, based on the same data that CPT-4 is trained on, we cannot exclude the possibility that it will give similar outputs to similar

---

[944] Dave Simpson, *GitHub Users Say Microsoft Ripped off Their Code for AI Tool*, Law360 (November 3, 2022, 9:15 PM EDT).
[945] Doe v. Github, Inc., No. 22-cv-06823-JST, 2023 U.S. Dist. LEXIS 86983, at *14 (N.D. Cal. May 11, 2023).
[946] OpenAI, *GPT-4 Technical Report*, arXiv:2303.08774 (2023).

prompts. To test if ChatGPT will reuse the same content for the same prompt, I reused the same prompt on July 1, 2023, to test ChatGPT's output (see **Appendix D**).[947]

Because the logline is the first expansion of my prompt, within the limitations of words and creative freedom, several words and phrases are inevitably overlapping. As the story develops into incidents, there is more creative freedom. We can find few words or phrases overlapping except for the keywords in the prompts. Therefore, we can find that ChatGPT does not reuse the same literal output in this case. The comparison of two stories is as follows (see **Table 5-1**):

Table 5-1 Comparison between Two Scripts Generated by ChatGPT

	06/27/2023 [Appendix C]	07/01/2023 [Appendix D]
Prompt	Write a film script based on the words: law school students, online bar exam, covid pandemic, love.	Write a film script based on the words: law school students, online bar exam, covid pandemic, love.
Title	Virtual Verdicts	Love Amidst Bar Exam
Logline	In the midst of a global pandemic, a group of law school students prepares for an unprecedented online bar exam. As they navigate the challenges of studying remotely and the uncertainties of the COVID-19 era, two students find love and support in each other, giving them the strength to face their fears and fight for their dreams.	Amidst a global pandemic, a group of law school students navigates the challenges of preparing for an online bar exam. As they strive to achieve their professional dreams, they unexpectedly find love and support in each other, proving that even in the toughest times, love can still triumph.
Characters	EMMA (23, a diligent and intelligent law student) ADAM (25, a charismatic and charming law student)	JULIA (late 20s, determined but anxious law student) ALEX (early 30s, charismatic and supportive law student) LUCAS (ambitious and focused law student) EMMA (strong-willed and compassionate law student)

---

[947] Love Amidst Bar Exam (July 1, 2023), https://chat.openai.com/share/397b5542-bc4e-4868-868c-35c61b21cc14. [https://perma.cc/RYV2-9D3M]

Incident 1	LAW SCHOOL LIBRARY - DAY	LAW SCHOOL LIBRARY - DAY
	The bustling law school library is filled with determined students pouring over textbooks, their faces hidden behind masks. We focus on EMMA, a diligent and intelligent law student, and ADAM, a charismatic and charming law student. EMMA, 23, sits alone at a table, scrolling through her laptop, researching online bar exam preparations. ADAM, 25, notices her from across the room. ADAM (whispering) Hey, is this seat taken? EMMA looks up, startled, but her eyes soften as she notices ADAM's warm smile. EMMA (smiling) No, go ahead. ADAM sits down, and they strike up a conversation about the challenges of studying during the COVID-19 pandemic and their fears about the online bar exam.	A group of LAW SCHOOL STUDENTS sits at a long table, studying books and legal documents. Laptops and online study materials are scattered everywhere. The atmosphere is tense. JULIA (late 20s, determined but anxious) looks up from her notes and notices ALEX (early 30s, charismatic and supportive) sitting across from her. JULIA (stressed) Alex, do you think we'll be ready for this online bar exam? The pressure is overwhelming. ALEX (sincerely) I understand, Julia, but we've come this far. We're going to conquer this exam together. We have to believe in ourselves. They share a brief moment of connection.
Incident 2	EMMA'S APARTMENT - NIGHT EMMA sits at her desk, surrounded by books and papers, feeling overwhelmed. Her laptop screen illuminates her determined face. EMMA (V.O.) This online bar exam is turning my life upside down. I don't know if I can do it. Just then, her phone buzzes with a notification. It's a text from ADAM. ADAM (TEXT) You're stronger than you think. Believe in yourself. I believe in you. EMMA smiles and takes a deep breath, finding comfort and inspiration in ADAM's message.	N/A
Incident 3	VIRTUAL CLASSROOM - DAY	VIRTUAL CLASSROOM - NIGHT

	EMMA and ADAM, along with their fellow law school students, log into the virtual classroom for a bar exam preparation lecture. The professor, PROFESSOR WILSON, appears on the screen. PROFESSOR WILSON Good morning, everyone. Today, we'll be discussing contract law... EMMA spots ADAM's video feed and catches his eye. They exchange encouraging smiles, their connection growing stronger.	The students gather for an online class, their faces displayed in individual video windows. The professor explains the challenges of conducting the bar exam remotely due to the COVID PANDEMIC. PROFESSOR (through video) Class, due to the ongoing pandemic, the bar exam will be administered online. It's going to be tough, but I believe in each and every one of you. The students exchange worried glances but remain determined.
Incident 4	N/A	VIRTUAL STUDY GROUP - DAY Julia initiates a virtual study group with her classmates, including Alex, LUCAS (ambitious and focused), and EMMA (strong-willed and compassionate). They discuss legal concepts and share study tips. EMMA (excitedly) Guys, I found a helpful online resource! It breaks down the complex concepts we're struggling with. They all eagerly take notes, bonding over their shared determination to succeed.
Incident 5	COFFEE SHOP - DAY EMMA and ADAM sit across from each other at a cozy coffee shop, their textbooks and notes scattered on the table. ADAM You know, studying law during a pandemic feels like an uphill battle, but having you by my side makes it easier. EMMA blushes, grateful for his support. EMMA I feel the same way. We can do this together. They hold hands, united in their determination to overcome the challenges ahead.	N/A

Incident 6	N/A	PARK - DAY Julia and Alex meet for a study break. They stroll through the park, enjoying a moment of respite from their intense preparation. JULIA (softly) Alex, I appreciate your constant <mark>support</mark>. I couldn't have made it this far without you. ALEX (gently) Julia, we've been through so much together. I want you to know that I've fallen in <mark>love</mark> with you, too. I believe in us, not just as study partners but as something more. They share a heartfelt embrace, realizing their connection goes beyond the bar exam.
Incident 7	ONLINE EXAMINATION HALL - DAY EMMA and ADAM sit in their respective homes, facing their laptops. The virtual <mark>bar exam</mark> begins, and tension fills the air. EMMA (V.O.) This is it. Months of hard work and sacrifices come down to this moment. As they navigate through the exam questions, they exchange reassuring glances over video chat, their <mark>love</mark> and support fueling their determination.	VIRTUAL EXAM ROOM - DAY The day of the <mark>online bar exam</mark> arrives. The students, including Julia and Alex, sit at their respective desks, laptops open. Anxiety fills the air as they prepare for the momentous test.
Incident 8	<mark>GRADUATION CEREMONY</mark> - DAY EMMA and ADAM, dressed in their graduation gowns, stand among their law school peers. Their eyes meet, filled with pride and <mark>love</mark>. DEAN WILLIAMS, a distinguished figure, addresses the graduates. DEAN WILLIAMS Today, we celebrate your resilience, your adaptability, and your unwavering commitment to the pursuit of justice. You have overcome extraordinary challenges to be here. EMMA and ADAM hold hands tightly, a	LAW SCHOOL AUDITORIUM - DAY (MONTHS LATER) A <mark>graduation ceremony</mark> takes place, celebrating the students' successful completion of law school and passing the <mark>online bar exam</mark>. As Julia, Alex, Lucas, and Emma walk across the stage, they receive their diplomas and celebrate their hard-earned achievements.

	silent acknowledgment of the obstacles they've faced and the love that carried them through.	
Incident 9	GRADUATION PARTY - NIGHT EMMA and ADAM ==dance together, surrounded by their friends== and ==loved== ones, celebrating their accomplishments. EMMA (whispering) I couldn't have done it without you. ADAM (==kiss==ing her forehead) And I couldn't have done it without you. We make a great team. They share a loving embrace, their journey as ==law school students== in the midst of a ==pandemic== now a chapter in their lives, forever marked by resilience, ==love==, and triumph. FADE OUT.	RECEPTION HALL - NIGHT A celebration is in full swing. Music plays, and laughter fills the room as the law school graduates enjoy their success. Julia and Alex ==dance together==, their eyes full of ==love== and pride for each other. JULIA (happily) Who would've thought that love would bloom amidst the chaos of law school and a ==pandemic==? ALEX (grinning) It's true what they say: ==love== conquers all. We faced incredible challenges together, and now we get to build a future side by side. They share a ==kiss==, ==surrounded by their friends== who have become their chosen family. FADE OUT.

c. Copy from Web Browsing

OpenAI has fun-tuned GPT-3 to give it the ability to search for up-to-date information on the internet.[948] This function was provided to ChatGPT users in March 2023.[949] WebGPT might be good news for users because GPT-3's knowledge is no longer frozen on training data's cut-off date. But from the perspective of copyright law, it will add more copyright violation risks. Nonetheless, this risk seems controllable because WebGPT is designed to cite the source.

---

[948] WEBGPT: IMPROVING THE FACTUAL ACCURACY OF LANGUAGE MODELS THROUGH WEB BROWSING, OPENAI (Dec.16, 2021), https://openai.com/research/webgpt [https://perma.cc/3X2E-9BH3].
[949] OPENAI, CHATGPT PLUGINS, https://openai.com/blog/chatgpt-plugins [https://perma.cc/XEM3-89B9].

d. Influence on Proof of Access

Chapter Four mentions that when the defendant denies copying, it is the plaintiff's burden to prove "access," i.e., the defendant had a "reasonable opportunity or a reasonable possibility of viewing" the allegedly infringed work.[950]

If the plaintiff can prove the defendant was using an AI to make the allegedly infringing work, and the infringed work was included in the AI's training data, it may be sufficient to prove that the plaintiff has access to the infringed work. Some may argue that ChatGP's access does not equal the user's access. But it is not deniable that the infringed work's data is ready for ChatGPT to use, and ChatGPT is ready for the user to use. In other words, ChatGPT is the bridge (access) between the user and the infringed work. Therefore, a movie studio using GPT-3 to develop film scripts would expose itself to the risk of accessing all the works in GPT-3's training data. And there are a lot of them.

2. Substantial Similarity Concerns

a. Idea

Chapter Four has introduced that the courts examine substantial similarity regarding the expression, not the idea.[951] A prompt, like "Write a film script based on the words: law school students, online bar exam, covid pandemic, love," is short, general, and abstract. It would be considered only an idea and not included in considering substantial similarity.

---

[950] Astor-White v. Strong, 817 F. App'x 502, 503 (9th Cir. 2020).
[951] Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231, 239 (2d Cir. 1983).

b. Expression

As shown in **Table 5-1**, ChatGPT does a decent job in avoiding the same text output for the same prompts. Suppose the first script (Appendix C) is copyrightable (the first script's copyrightability will be discussed in the next section), does ChatGPT also do a good job in avoiding substantial similarity? Here I apply the Ninth Circuit's two-part test to see if the two scripts (Appendix C and Appendix D) share substantial similarities.[952] The extrinsic test examines "the plot, themes, dialogue, mood, setting, pace, characters, and sequence of events."[953] The intrinsic test focuses on the "total concept and feel."[954]

For the characterization of the two scripts, the characters' personalities do not have obvious differences. The male student is mentally stronger than the female student and supports the female student along the way. The plots of incident 1,3,7,8,9, and the sequences of events are similar. Both dialogues focus on bar exams and mutual encouragement. The setting of the library, classroom, bar exam, commencement, and graduation party are the same. Two scripts have the same mistake because the bar exam shall happen after the commencement. Both the extrinsic test and the intrinsic test are met. Therefore, I conclude the two scripts are substantially similar. (Some may argue that many similarities are Scènes à Faire and should not be included in the comparison.) This test tells us LLM could generate substantial similar text output for the same prompts.

---

[952] Kouf v. Walt Disney Pictures & Television, 16 F.3d 1042, 1045 (9th Cir. 1994)
[953] Rice v. Fox Broad. Co., 330 F.3d 1170, 1179 (9th Cir. 2003).
[954] *Kouf*, 16 F.3d at 1045.

III. NEW COPYRIGHTABILITY ISSUES FOR DATA-DRIVEN WORKS

Data-driven creation has touched writers' nerves. No one will just sit and watch AI taking away their bread and butter, although AI does not need to eat at all. Without hiding despise for AI-generated output, a Writers Guild of America West's tweet directly challenged AI-generated works' copyrightability, including the lack of originality and human authorship, which stated:[955]

> It is important to note that AI software does not create anything. It generates a regurgitation of what it's fed. If it's been fed both copyright-protected and public domain content, it cannot distinguish between the two. Its output is not eligible for copyright protection, nor can an AI software program sign a certificate of authorship. To the contrary, plagiarism is a feature of the AI process.

A. ORIGINALITY CONCERNS

Originality requires a copyrightable work to be independently created by the author and possess minimal creativity.[956] Therefore, I will discuss the independent creation element and the minimal creativity element separately.

1. Independent Creation Element

Independent creation disallows the author from copying other works,[957] but it does not demand the author to have full creative autonomy and forbids the author from using any source materials.[958] I will not say a film script is not an independent creation because

---

[955] Writers Guild of America West (@WGAWest), Tweeter (Mar 22, 2023, 8:47 PM), https://twitter.com/WGAWest/status/1638643544977195008?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweet embed%7Ctwterm%5E1638643544977195008%7Ctwgr%5Eac8451a069a50959ba745ad0e69e4c0a084520 1f%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fvariety.com%2F2023%2Fbiz%2Fnews%2Fwriters-guild-artificial-intelligence-proposal-1235560927%2F [https://perma.cc/A5WU-NUTQ].
[956] Feist Publications, Inc. v. Rural Telephone Service Co., Inc., 499 U.S. 340, 345 (1991).
[957] *Id.*
[958] Berkic v. Crichton, 761 F.2d 1289, 1294 (9th Cir. 1985).

its writer followed the movie producer's simple instructions and borrowed ideas from source materials. For the same reason, I would not jump to conclude that Appendix C script is not an independent creation just because it followed my instructions and was produced out of data.

Therefore, the material question is whether the script of Appendix C is copied from training data. If the whole script copies another work, the entire script is not independently created. If part of the script copies another work, this part is not independently created. It is critical to note that simple paraphrasing is usually insufficient to prove independent creation if the script substantially resembles the paraphrased works.[959] Without knowing the specific training data used for generating this script, it is unclear whether Appendix C is an independent creation.

2. Minimal Creativity Element

Chapter One mentions three forms of creativity: combination, exploration, and transformation. It is possible for an AI to simulate the former two forms of creativity.[960] AI containing genetic algorithms can also perform the third form of creativity.[961] Such AI can imitate biological evolution to make random changes to generate novel and surprising content based on the program's task-oriented rules. The rules can include changing the rule itself. Even so, whether AI could be creative is still an unanswered philosophical question.[962] Some philosophers have asserted three features are essential for creativity:

---

[959] *See* LEAFFER, *supra* note 80, at 420; *see also* ". . . though the 'ordinary' phrase may be quoted without fear of infringement, a copier may not quote or paraphrase the sequence of creative expression that includes such a phrase." Salinger v. Random House, Inc., 811 F.2d 90, 98 (2d Cir.), opinion supplemented on denial of reh'g, 818 F.2d 252 (2d Cir. 1987).
[960] *See* Boden, *supra* note 58 at 229.
[961] *See id.* at 230.
[962] *See id.* at 224.

"autonomy, intentionality, and consciousness (including valuation and emotion)."[963] But there is still disagreement on this assertion, and there are different opinions about what these features mean. Most AI developers are not interested in discussing whether machines can really think.[964] For example, the computer scientist Edsger Dijkstra says this question is similar to asking whether submarines can swim.[965]

ChatGPT can respond to a human user's instruction with language output without realizing the meaning of human language because it was repeatedly trained until its performance was rewarded with human trainers' positive feedback.[966] This situation can be explained by the philosopher John Searle's famous Chinese Room argument.[967] His argument describes a hypothetical English speaker who knows nothing about the Chinese language in a room. Following the direction of an English-written rulebook, he is busy accepting pieces of paper containing Chinese symbols and sending out the completed document in Chinese through the room door. From the views of people outside the room, his Chinese is fluent. But he does not understand Chinese at all. Without considering ChatGPT is not a human, will copyright law care about what the author thinks and perceives his works? The answer seems to be negative.

In *Star Athletica, L. L. C. v. Varsity Brands, Inc.*, when examining if Varsity's uniform design is eligible for copyright, the Supreme Court rejected to "consider evidence of the creator's design methods, purposes, and reasons" because there is no ground in the

---

[963] *See id.* at 236.
[964] RUSSELL & NORVIG, *supra* note 907, at 984.
[965] *Id.*
[966] *See* INTRODUCING CHATGPT, *supra* note 422.
[967] RUSSELL & NORVIG, *supra* note 907, at 985.

copyright law.[968] Justice Thomas wrote:[969]

> The statute's text makes clear, however, that our inquiry is limited to how the article and feature are perceived, not how or why they were designed. See *Brandir Int'l, Inc. v. Cascade Pacific Lumber Co.*, 834 F.2d 1142, 1152 (C.A.2 1987) (Winter, J., concurring in part and dissenting in part) (The statute "expressly states that the legal test is how the final article is perceived, not how it was developed through various stages").

Siding with the Supreme Court, the U.S Copyright Office makes it clear that it "will focus solely on the appearance or sound of the work that has been submitted for registration to determine whether it is original and creative" without considering "the author's inspiration for the work, creative intent, or intended meaning,"[970] "the author's skill, experience, or artistic judgment,"[971] and "the amount of time, effort, or expense required to create the work."[972]

Therefore, to determine the minimal creativity element, I only examine the script of Appendix C on its appearance without looking into the creation process. This script contained all the film script elements, including title, genre, plot, themes, dialogue, mood, setting, pace, characters, and sequence of events.[973] It is short but looks nice and understandable. This script has minimal creativity when the human authorship requirement is not considered.

---

[968] Star Athletica, L.L.C. v. Varsity Brands, Inc., 580 U.S. 405, 422 (2017).
[969] *Id.* at 422-423.
[970] COMPENDIUM (THIRD) § 310.5.
[971] *Id.* § 310.6.
[972] *Id.* § 310.7.
[973] Kouf v. Walt Disney Pictures & Television, 16 F.3d 1042, 1045 (9th Cir. 1994).

B. HUMAN AUTHORSHIP CONCERNS

As discussed in Chapter Four, human authorship is required in the United States, China, and Taiwan. It seems China and Taiwan care more if a human author is involved in the creative process, while the United States cares more about the author's involvement in the expression of final work. However, we have not seen a case directly related to Large Language Model generated works like ChatGPT.

In Chapter One, I divided the scriptwriting process into eight steps: (1) theme setting up; (2) raw materials selection; (3) logline; (4) outline; (5) treatment; (6) first draft; (7) rewrite, and (8) polish. In addition, I listed three types of human and AI collaboration. I will answer as follows:

1. Type 1: Machine First, Human Last (Machine → Human)

[Example]

A film script is completed by a machine contributing to step (1) through step (3) (creative contribution 22%), and a human writer contributing to step (4) through step (8) (creative contribution 78%).

[Comments]

This situation is like the human writer making the script's expression based on the ideas provided by AI. The human writer is the author and copyright owner of the full script's expression but does not have a monopoly on the ideas.

2. Type 2: Human First, Machine Last (Human → Machine)

[Example]

A film script is completed by a human contributing to step (1) through step (3) (creative contribution 22%), and a machine contributing to step (4) through step (8) (creative contribution 78%).

[Comments]

This situation is like AI making the script's expression based on the human writer's ideas. The AI is not eligible to be the author and copyright owner for the script's expression. The writer does not have a monopoly on his ideas. He is also not the author and copyright owner of the script's expression.

However, it may be argued that AI is just the human writer's tool, and the human writer should be the author and copyright owner for the full script's expression.

3. Type 3: Human And Machine Back And Forth (Human ⇄ Machine)

[Example]

A film script is completed by a human and a machine working together on steps (1) through step (8) of the writing process. The machine generates content based on human-created content, and the human creates content based on machine-generated content. Either the human or the machine revises each other's content back and forth. Humans and machines contribute more or less to each step until all steps are finished. The human creativity efforts and the machine's outputs are commingled. Overall, humans contribute 30%, and machines contribute 70%.

[Comments]

It is necessary to examine the final script to decide which part is completed by AI and which part is completed by the human writer. It is critical to note the examination is not limited to the literal text but extends to treatment, details, scenes, events, and characterization.[974] The human writer is the author and copyright owner of the part that he completed. But he does not have a monopoly on the part that AI made. However, it may be argued that AI is just the human writer's tool, and the human writer should be the author and copyright owner for the full script's expression.

Returning to the script of Appendix C, my only contribution is the prompt "Write a film script based on the words: law school students, online bar exam, covid pandemic, love." This prompt is an idea at most. If John Doe copies my prompt to generate another script, I cannot sue him for unlawfully copying my prompt. As to the script's expression, I had no creative control over it and did not predict the content generated by ChatGPT.[975] Therefore, I am not the author of the script. The script's expression is completely completed by ChatGPT, not a human, so it is not copyrightable.

IV. NEW SCOPE OF COPYRIGHT ISSUES FOR DATA-DRIVEN WORKS

A. CORE OF PROTECTABLE MATERIAL

As discussed above, only minimum creativity is required for a motion picture to be copyrighted. However, it is critical to note that copyright protection does not extend to non-

---

[974] Reyher v. Children's Television Workshop, 533 F.2d 87, 91 (2d Cir. 1976).
[975] U.S. COPYRIGHT OFFICE, 37 CFR PART 202 COPYRIGHT REGISTRATION GUIDANCE: WORKS CONTAINING MATERIAL GENERATED BY ARTIFICIAL INTELLIGENCE, 88 Fed. Reg. 16190 (Mar. 16, 2023).

original elements, e.g., ideas, facts, *Scènes à Faire*, and other elements that do not originate from the author. These non-original elements must be filtered out to ascertain the work's scope of copyright. In a copyright infringement case, only after defining the "core of protectable material,"[976] can the courts evaluate whether the alleged infringer copied the plaintiff's original expression and whether the similarities between the protected work and the allegedly infringing work constitute substantial similarity. Therefore, in a data-driven work, we must filter out the non-human expression generated by machines to ascertain the scope of copyright.

B. DISAPPEARING HUMAN EXPRESSION

Recent data science and AI developments have allowed machines to provide the idea and expression of the motion picture (see **Table 5-2**). More than that, most human expressions may be dominated by data science because they can catch the patterns of what the audience wants in detail. For a data-driven work with minimum human creativity, it will be difficult to separate the very thin human expression from the non-human expression based on the work's appearance. To detect human expression in a data-driven work would be like searching for a tiny pearl in a large ocean (Type 1: Machine → Human), extracting a handful of salt dissolved in tons of seawater (Type 2: Human → Machine), or identifying dolphin genes and whale genes from a Wholphin (Type 3: Human ⇄ Machine).[977]

---

[976] Computer Assocs. Int'l, Inc. v. Altai, Inc., 982 F.2d 693, 707 (2d Cir. 1992).

[977] Wholphin is "a mixture of two different types of cetaceans, her mom is a bottlenose dolphin, and her dad is a false killer whale!" born in 1985 in Sea Life Park, Hawaii. *See* DOLPHINS & WHALES, SEA LIFE PARK, https://www.sealifeparkhawaii.com/conservation-n-education/animal-profiles/dolphins---whales [https://perma.cc/FX5A-CTAP].

In a copyright infringement case where an innocent plaintiff files against a notorious pirate, a sympathetic court may have no option but to adopt two approaches to protect the plaintiff: (1) move the boundary of the abstract test in favor of the plaintiff to find expressions that were traditionally considered ideas,[978] or (2) accept trivia clues to prove substantial similarity. For example, a few insignificant similarities between words and phrases. Neither approach is welcomed to the copyright system because copyright shall not extend to ideas,[979] ordinary phrases, and small groups of words.[980]

Table 5-2 Interaction between Human and Machine in Filmmaking

Who's idea	Who's expression	Example	Description
Human	Human	The oldest film Fred Ott's Sneeze (1894) preserved in the U.S. Copyright Office	Filmmakers' ideas and expressions.
	Human	Use of box office, sneak previews, fan mails, exhibitor's opinions, theatrical tryouts, etc., to guess audience preference	Since the 1910s, Hollywood had begun to collect audience preference, but the collected information is limited.
	Human	Use of interviews, questionnaires, IBM cards, and machines to collect and analyze moviegoers' opinions with statistical methods.	Since the 1940s, Hollywood has hired the Audience Research Institute (ARI) and the Motion Pictures Research Bureau (MPRB) to conduct audience research systematically. The collected information is more precise but still limited.
Human (Master) & Machine	Human	The World Between Us (2019),[981] Who Killed the Good Man (2021)[982]	The writer gives keywords to the data science team to mine ideas from data. Then the writer combines her own idea

---

[978] Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930).
[979] 17 U.S.C. § 102(b).
[980] Narell v. Freeman, 872 F.2d 907, 912 (9th Cir. 1989) (rejecting the plaintiff's copyright violation claim against word-for-word copying and paraphrasing over 300 words).
[981] *See* Lin, *supra* note 7.
[982] *See* Yang, *supra* note 9.

(Slave)			and the ideas suggested by data to write the script.
Human (Client) & Machine (Service Provider)	Machine	Some students used Shortly to develop a 3.5-minute screenplay named Solicitors.[983]	In 2020, Shortly completed everything except for the open scene ("Barb's reading a book. A knock on the door. She stands and opens it. Rudy, goofy-looking, stands on the other side") and the first two dialogue lines given by the students.
		Sunspring (2016) is the first one written entirely by Artificial Intelligence	Ross Goodwin and Oscar Sharp trained Benjamin to write screenplays by feeding him dozens of 1980s and 1990s sci-fi screenplays.
		The scripts of Appendix C and Appendix D	The plot, dialogue, mood, setting, pace, characters, and sequence of events are all generated by ChatGPT except for the keywords.

V. CONCLUSION

As discussed above, the copyright violation risks of data-driven works, particularly during the creation processes, are exceedingly high and uncertain. Data-driven works are also disadvantaged in asserting independent creation defense if the plaintiff can prove the allegedly infringed work existed in the training data. A cautious motion picture producer would not want to take risks to create a movie script model or use available LLMs to create scripts unless it assures all the copyrighted data's proprietary right or license. In other words, whoever controls the data will win the battle in the data-driven era. This reminds us of the history that happened in 1912: who controlled the stories would monopolize the motion picture industry.

---

[983] Ramirez, supra note 416.

With regard to the copyrightability issue, a data-driven motion picture will be eligible to copyright if there is minimal human creativity. It would not be a big problem for the motion picture makers because they just need to hire some human writers to add a minimum of creativity to the data-driven work. Still, machines cannot substitute a director, cinematographer, camera assistant, lighting technician, lighting assistant, design specialist, costume designer, and actor soon. Thus, a motion picture can always be copyrightable if minimum human intellectual input can be found in its expression.

The most difficult question is how to define an appropriate scope of copyright for data-driven works made with the collaboration of humans and machines. Unless the "core of protectable materials" of these data-driven works is inspected before the court, they will deter others from using the whole contents as if they are made of pure gold. As their human authorship may range from 0.000 . . . . 1% to 99.99999 . . . .%, the expansive and time-consuming litigation may prevent the competitors and the public from using the materials. This is the most core problem that the next chapter wants to address.

CHAPTER SIX: PROPOSALS AND CONCLUSION

The closing chapter of this dissertation is divided into three parts. The first part recaps this research's salient findings and demonstrates new copyright challenges for motion pictures in the data-driven era. The second part enumerates policy considerations for supporting a narrower scope of copyright for motion pictures to confront with data-driven trends. The third part contains my proposals and conclusion.

I. MOTION PICTURE TECHNOLOGY AND COPYRIGHT

A. MOTION PICTURE, DATA TECHNOLOGY, AND AI

Motion pictures were born of technology,[984] and their format, content, and distribution continuously evolved with the advancement of technology.[985] As an industrial-technological product, film production is notoriously complex and expensive.[986] It requires a production team made of various talents and technicians. Nearly eighty percent of American features cost more than US$1 million. A blockbuster could even cost US$300 million. No matter whether motion pictures are technological products or artworks, there would be no motion pictures if there were no business. Therefore, all the time, businesspeople like studio executives and producers are the decision-makers in Hollywood.[987] They believe "giving the audience what it wants" is the key to success and place audience opinions above the film artists' desire to express.

---

[984] *See supra* Chapter Two. I. A.
[985] *See supra* Chapter Two. I. B.
[986] *See supra* Chapter Two. II. A.
[987] *See supra* Chapter Two. II. B.

Since the 1910s, Hollywood has measured audience preferences with industry-created methods such as box office figures, sneak previews, fan mails, exhibitors' opinions, and theatrical tryouts.[988] In the 1940s, scientific audience research like George Gallup's Audience Research Institute and Leo A. Handel's Motion Pictures Research Bureau started to conduct scientific audience surveys with traditional statistical and psychological methods such as interviews, questionnaires, IBM cards, and audience reaction profile charts.[989] However, scientific audience research relied on the sample audience's honest feedback and could only indicate a general preference. Also, it was not good at creating new content. Moreover, scientific research requires considerable amounts of time and labor. The motion picture producers' urgent requests and fast film production pace often compromised its accuracy.

Things dramatically changed in the internet age. The prevalence of digital data, social media, and streaming services increases the convenience, instantaneousness, volume, width, and depth of audience data and content data collection.[990] Data engineering eases the burden of managing and storing vast data.[991] Data science makes it possible to detect the hidden patterns and interrelationships between audience and content.[992] With vast data and data technology, several companies, e.g., Cinelytic, ScriptBook, LargoAi, provide the motion picture industry with film analytics services, including script diagnosis, revenue predictions, talent suggestions, and marketing strategies.[993]

---

[988] *See supra* Chapter Two. III. A.
[989] *See supra* Chapter Two. III. B & C.
[990] *See supra* Chapter One. I. B, Three. I. A.
[991] *See supra* Chapter One. I. B, Three. I. B.1.
[992] *See supra* Chapter One. I. B, Three. I. B.2.
[993] *See supra* Chapter Three. II.

More than this, massive data is widely used to train AI models to generate new content. For example, Disney research's scriptwriter-assisting models, [994] GPT-3's downstream application SortlyAI, [995] Deepmind's Dramatron, [996] Ross Goodwin's Benjamin,[997] and IBM's Watson.[998] These attempted uses of AI in generating film content have worried many writers and have become one of the difficult issues on the Minimum Basic Agreement negotiation table between the Writers Guild of America and the Alliance of Motion Picture and Television Producers in 2023.[999] In addition to generating scripts, data-driven trends have expanded throughout motion picture pre-production, production, and post-production phrases and have even extended to screen education.[1000]

B. MOTION PICTURE AND COPYRIGHT

As mentioned in Chapter Four, the original authorship requirement and idea-expression dichotomy are fundamental copyright principles. These two principles and their downstream fact-expression distinction, merger doctrine, and *Scènes à Faire* determine a motion picture's copyrightability and its scope of copyright. A motion picture's scope of copyright protection is the most debated and complicated issue in an alleged copyright infringement litigation because it is the premise to examine whether the copying element and substantial similarity element are met to prove copyright infringement. One way to

---

[994] *See supra* Chapter Three. III. A.
[995] *See supra* Chapter Three. III. B.
[996] *See supra* Chapter Three. III. C.
[997] *See supra* Chapter Three. IV. C.
[998] *See supra* Chapter Three. IV. D.
[999] WGA NEGOTIATIONS—STATUS AS OF MAY 1, 2023, WGA ON STRIKE, https://www.wgacontract2023. org/the-campaign/wga-negotiations-status-as-of-5-1-2023[https://perma.cc/LLV7-8JHT].
[1000] *See supra* Chapter Three. IV.

approach these questions is to map out the four possible combinations, as illustrated by the box below (**Table 6-1**).

Table 6-1 Scope of Copyright Protection

	Original Authorship	**Non-original Authorship**
Idea	Unprotectable	Unprotectable
Expression	Protectable	Unprotectable

It is critical to note that only original human expression is protected by copyright law. The copyright statutes and judicial precedents of the United States, China, and Taiwan share high commonality in many aspects. Nonetheless, from my observation of the few cases, the Taiwanese and Chinese courts tend to admit that minimum human intellectual involvement during the creation process would be sufficient to establish original human authorship.[1001] By comparison, the USCO requires human authors' foreseeability and direct control over the final work's expression.[1002]

Today's motion picture industry is deeply shaped by copyright. Before the motion picture was added to the U.S. copyright law in 1912, it was protected as a photograph.[1003] Early court's tolerance toward remaking rival film companies' materials gave the young film industry an open environment to exchange creativity.[1004] Things changed after the

---

[1001] *See supra* Chapter Four. I. B. 2. 2.
[1002] *See supra* Chapter Four. I. B. 1.
[1003] *See supra* Chapter Five. I. A. 1.
[1004] *See supra* Chapter Five. I. A. 2.

publishers and theater producers took the film companies to court.[1005] Since then, the

motion picture industry has learned the necessity to prevent copyright violation claims.

Copyright warranty, refusal to accept unsolicited materials, submission release agreement,

E&O insurance, and clearance of titles and source materials have become common

practices.[1006] Due to the motion picture producers' lobby, a new limitation to statutory

damage against innocent filmmaking infringers was introduced to the 1912 Townsend

Amendment.[1007] Most importantly, motion picture producers realized an effective way to

monopolize the marketplace was to control original stories.[1008] Exclusive cross-media

alliances created entertainment empires. Numerous writers became worker bees producing

original stories in the dream factory.[1009] Subsequently, the Writers Guild of America was

formed to safeguard the writers' interests against motion picture producers'

exploitation.[1010]

Once again, data-driven trends powered by data technology and AI bring new

copyright challenges to the motion picture industry. Among these challenges, I do not

worry much about the copyrightability issue.[1011] This is because I do not foresee data

technologies and AI will soon replace all types of film artists' jobs and can create a motion

picture good enough to exhibit in cinemas, televisions, or streaming platforms on their

own.[1012] No matter how powerful the GPT-4 is, it is a language model that can only perform

---

[1005] *See supra* Chapter Five. I. A. 3.
[1006] *See supra* Chapter Five. I. B. 1. a.
[1007] *See supra* Chapter Five. I. B. 1. b.
[1008] *See supra* Chapter Five. I. B. 2.
[1009] *See supra* Chapter Two. II. B. 1.
[1010] *See supra* Chapter Five. I. B. 2.
[1011] *See supra* Chapter Five. III..
[1012] *See supra* Chapter Three. IV.

language tasks.[1013] It cannot replace a director, cinematographer, camera assistant, lighting technician, lighting assistant, design specialist, costume designer, and actor. A motion picture can always be copyrighted if minimum human authorship is involved in its creation.

Nonetheless, I am concerned with the copyright violation risks in the data-driven process and the copyright violation risks of data-driven works. Many activities in the data-driven creation process, like data preparation, model training, model fine-tuning, and prompting, hide uncertain copyright violation risks.[1014] Moreover, data-driven works derived from common or similar datasets may share similarities. But in copyright infringement litigations, data-driven works are unfavorable to prove no access to Plaintiffs' work if Plaintiffs' works are buried in tons of training data, fine-tuning data, and prompts that are parsed from the internet.[1015] Given the high and uncertain copyright violation risks, small and cautious motion picture producers will be kept out of data and AI technologies. It is foreseeable that whoever has control over copyrighted film materials will become the overlord of the motion picture industry in the data-driven era. This is the first problem that my proposal wants to address.

Furthermore, I am concerned with the scope of copyright for motion pictures made with human and machine collaboration.[1016] Advanced data technology and AI have broken the boundary between original and non-original creation. They can either produce ideas or make human-like expressions. Although many detection tools have been developed to

---

[1013] *See supra* Chapter Three. III. C.
[1014] *See supra* Chapter Five. II. A.
[1015] *See supra* Chapter Five. II. B.
[1016] *See supra* Chapter Five. IV.

identify AI-generated content, it seems like an endless game of catching up.[1017] Even OpenAI is disappointed with its AI classifier's low accuracy rate and stopped its use on July 20, 2023.[1018] Some have attempted to "apply traceable credentials to digital work upon creation."[1019] But a credential attached to a digital script will not stay on a motion picture made with the script. From the appearance of data-driven works, particularly those that humans have processed, it is difficult to identify which part is the human-made expression and which part is AI made with data. Expansive and time-consuming copyright litigation may prevent others from accessing many low human authorship input materials. This is the second problem that my proposal wants to address.

II. POLICY CONSIDERATIONS

Before proceeding with my proposals, I would like to address several essential policy considerations below.

A. SUSTAIN THE VALUE OF HUMAN CREATIVITY AND CONTRIBUTION

Regulate use of artificial intelligence on MBA-covered projects: AI can't write or rewrite literary material; can't be used as source material; and MBA-covered material can't be used to train AI.[1020]

—Writers Guild of America Proposal

[1017] Tiffany Hsu & Steven Lee Myers, *Another Side of the A.I. Boom: Detecting What A.I. Makes*, N.Y. TIMES (May 18, 2023), https://www.nytimes.com/2023/05/18/technology/ai-chat-gpt-detection-tools.html?searchResultPosition=1 [https://perma.cc/45BX-7GW5].

[1018] Jan Hendrik Kirchner et.al., *New AI classifier for indicating AI-written text*, OPENAI (Jan. 31, 2023), https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text [https://perma.cc/JR82-BBRZ].

[1019] Hsu & Myers, *supra* note 1017.

[1020] WGA Negotiations, *supra* note 999.

At the end of this research, a historical Hollywood writer's strike is ongoing.[1021] Motion picture makers will not abandon human writers completely if the human authorship requirement remains unmoved in copyright law. Besides, a powerful Large Language Model, like GPT-3 or its next-generation GPT-4, may substitute a craftsman screenwriter who only loyally and mechanically follows the producer's instruction to execute writing assignments. It cannot replace a writer who can make a creative prompt for the machine to perform the writing task.

Nonetheless, because data-driven scripts can be copyrightable with minimal human expression, motion picture producers may want to hire as few writers as possible. Probably, these human writers' salaries are even cheaper than the costs of maintaining AI and acquiring data. Therefore, an incentive to encourage retaining human writers is necessary.

B. MOTION PICTURES CREATIVE FREEDOM IS LIMITED

As a popular work accessible to the public, a commentator says, "the evaluation of the level of creativity of a [motion picture] project is more in the hands of the public than the professional field."[1022] Moreover, as high-cost industrial products, motion pictures must attract as many audiences as possible. Its content needs to address the audience's opinions. Because data can be used to predict the audience's future behaviors based on their viewing history, adding that AI models can produce a huge amount of content in seconds and work twenty-four hours, it is not an alarmist talk that content generated by

---

[1021] Jason P. Frank, *The 2023 WGA Strike for Dummies*, VULTURE (updated June 30, 2023), https://www. vulture.com/article/wga-strike-2023.html [https://perma.cc/BJ75-SLM6].
[1022] BLOORE, *supra* note 88, at 134.

221

data and AI will exhaust all the available expressions for motion pictures. Therefore, it is necessary to preserve more creative freedom for human creators.

C. MORE MOTION PICTURE VERSION CHOICES FOR THE AUDIENCE

Congress could have imposed a compulsory license for film adaptations, as it did with song performances and recordings. Many companies could then have adapted the same novel as long as they paid the author the price determined by the statute.[1023]

—Peter Decherney, *Hollywood's Copyright Wars*

As mentioned in Chapter Four, due to "the difference between literary and graphic expression," no matter how detailed the literal description might be, "[t]he description of a character in prose leaves much to the imagination."[1024] One story may have one hundred versions of imagination in one hundred people's minds. Therefore, the filmmaker enjoys quite a lot of creative choices for making a motion picture apart from the based literary work. However, since the motion picture industry uses copyright to monopolize stories, we can only see one version of Superman on the large screen. Considering data-driven works containing thinner human original authorship, they deserve weaker copyright protection. The insights extracted from data shall be fairly shared by all the filmmakers rather than monopolized by whoever discovered them. More version choices can also serve more audiences' tastes and promote competition among filmmakers.

[1023] DECHERNEY, *supra* note 125, at 56.
[1024] Warner Bros. Ent. v. X One X Prods., 644 F.3d 584, 597 (8th Cir. 2011).

D. CREATE A FAIR ENVIRONMENT FOR NEW AND SMALL PLAYERS

We are witnessing a critical moment for the motion picture industry. From history, we saw the old oligopolies who attempted to monopolize the motion picture market by "film invention" patents were ultimately displaced by the new oligopolies who attempted to monopolize the cross-media markets by "stories" and thus became bigger and stronger.[1025] Again, a rare opportunity arises a hundred years after the Supreme Court's *Kalem* decision. Some incumbents may become even larger than today if they can monopolize the motion picture content-generating AI models by controlling the copyrighted data. Or a new generation of winners may rise, just like Paramount and Universal took the places of Edison and Biograph.

Some observers wonder if Netflix will finally displace Hollywood because Netflix controls vast audience data.[1026] However, Hollywood has rushed to catch up with Netflix in the new wave of the streaming war.[1027] It is hard to say who will be the final winners. Most importantly, in this new wave of large language models and large image models, Hollywood can strike any rivals with their bulk of film data if copyright law stands on its side (**Figure 4**). Suppose we are not pleased to see a few large studios dominate the motion picture content-generating AI models. It is necessary to create a fair environment for everyone to use copyrighted film materials in the creation process.


[1025] *See supra* Chapter Five. I. B. 2.
[1026] *See* SMITH, *supra* note 139, at 178.
[1027] Erik Gruenwedel, *JustWatch: Prime Video Edges Netflix in Q2 U.S. Market Share*, MEDIAPLAYNEWS (July 6, 2023), https://www.mediaplaynews.com/justwatch-prime-video-edges-netflix-in-q2-u-s-market-share/ [https://perma.cc/GWB4-LG9R].

Figure 4 Data-Driven Filmmaking Dataset and the Uses
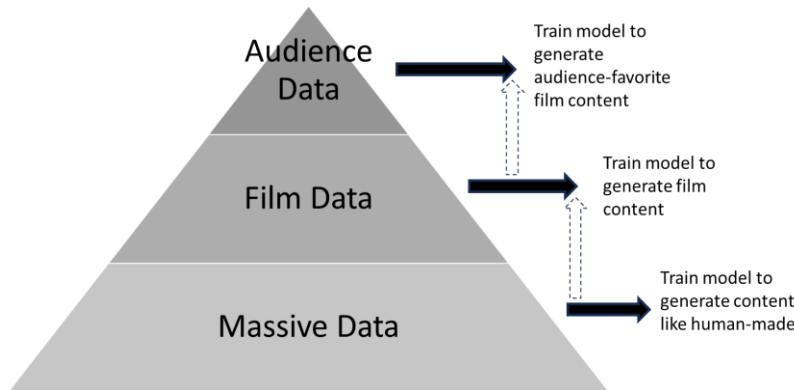
E. REDUCE JUDICIAL BURDEN TO INSPECT DATA-DRIVEN WORKS

Because any new work depends on others even if unconsciously, broad protection of intellectual property also creates a distinct possibility that the cost of litigation—old authors trying to get a "piece of the action" from current successes—will prevent or penalize the production of new works, even though the claims be rebuffed.[1028]

—Judge *Easterbrook*, *Nash v. CBS, Inc.*

As Chapter Five mentions, we must filter out the non-human expressions generated by machines powered by data to ascertain the scope of copyright case by case. As a data-driven work may only have a minimum human authorship, to detect a human expression within it would be like searching for a tiny pearl in a large ocean (Type 1: Machine → Human), extracting a handful of salt dissolved in tons of seawater (Type 2: Human → Machine), or identifying dolphin genes and whale genes from a Wholphin (Type3: Human

---

[1028] Nash v. CBS, Inc., 899 F.2d 1537, 1540–41 (7th Cir. 1990).

⇄ Machine). [1029] We know data-driven works may contain somewhat human original authorship, but finding it takes time and money. The courts will be asked to look into the complex creation process and thus increase litigation expenses. Consequently, many will surrender under the high litigation cost and long judicial process. Courts will also be exhausted. For people who want to use data-driven works, providing them with a lawful, efficient, and cheap way to access data-driven materials without litigation hurdles is necessary. It can also better save judicial resources.

III. PROPOSALS

As explained earlier, considering: (1) sustain the value of human writers' creativity and contribution, (2) motion pictures' creative freedom is limited, (3) provide more motion picture version choices for the audience, (4) create a fair environment for new and small players, and (5) reduce the judicial burden to inspect data-driven motion pictures, I propose as below (see **Table 6-2**):

Table 6-2 Proposals for Copyright Challenges of Data-Driven Works

STAGE	Creation Process		Final Work
COPYRIGHT VIOLATION RISK	Often involve massive film data		Often involve limited film data
PROPOSAL	Text and Data Mining Exception	Model Training Exception	Voluntary Negotiated License Backed Up by Compulsory License

---

[1029] *See supra* Chapter Five. III. B, Five. IV. B.

A. CREATION PROCESS: DATA PREPARING AND MODEL TRAINING SAFE HARBOR

1. Scheme

My first proposal is to add a copyright limitation or exception to using copyrighted materials in the motion picture creation process, including text and data mining ("TDM") and model training.

TDM aims to extract the patterns, trends, and correlations hidden in big data. These insights can inspire the creation of new content. In contrast, model training aims to build AI models to perform text and data mining tasks or generate new content. On the one hand, massive text and data are the premise of text and data mining. On the other hand, they are the admission ticket to access AI technology. Almost all film data and publicly available copyrighted materials can be the subjects of this limitation or exception. The amounts of data are so huge that it is practically or economically impossible to require the data miners or model developers to obtain all the copyright licenses.

Doctrinally, "copyright never assured authors even a limited monopoly over all forms of exploitation." [1030] There is no incontestable reason to oppose creating new limitations or exceptions for TDM and model training, and that is preferable compared to other options like fair use. Fair use may not be applicable for all types of TDM and model training activities, particularly for commercial purpose uses and when the number and

---

[1030] Jane C. Ginsburg, *Copyright and Control Over New Technologies of Dissemination*, 101 COLUM. L. REV. 1613, 1616 (2001).

portion of copyrighted film materials involved is large.[1031] Not to mention the inconsistent

ad hoc fair use analysis and legal uncertainty.

TDM and model training limitations or exceptions are also better than compulsory

license because establishing a compulsory license mechanism covering all available

copyrighted materials is hardly manageable. A functional film analytics platform or a

scriptwriting model demands massive data for TDM and model training. The large rate

adjudication and royalty distribution costs might not be justified in relation to copyright

owners' minor benefits.[1032]

2. Referenced National and Regional Legislation

Although the above limitations or exceptions have not been seen in the United

States, China, and Taiwan copyright laws, there have been laws in Japan (see **Appendix

E**),[1033] European Union (see **Appendix F**),[1034] and Singapore (see **Appendix G**).[1035]

Adding TDM limitations or exceptions do not surely mean all mining activities

would have infringed copyright. Council Directive (EU) 2019/790 indicates some activities

in relation to mere facts, not involving acts of reproduction, or complying with the

temporary reproduction exception are permitted under the copyright law.[1036] This new

exception aims to provide legal certainty to mining activities that involve making copies.

The EU directive covers "lawfully accessible works and other subject matter."[1037] By

---

[1031] Peter Henderson et al., *Foundation models and fair use*, ARXIV:2303.15715 (2023).

[1032] *See* Ginsburg, *supra* note 1030, at 1642-1645.

[1033] Chosakukenhō [Copyright Act], Law No. 52 of 2021, art.30-4, 47-5, translated in (Japanese Law Translation [JLT DS]), https://www.japaneselawtranslation.go.jp (Japan).

[1034] Council Directive 2019/790, art 3 & 4, 2019 O.J. (L 130), 92–125.

[1035] Copyright Act 2021 § 244 (No. 22 of 2021) (Sing.). *See also* Lim Daryl, *Life After Google v. Oracle: Three Reflections on a Theme*, 12 IP THEORY 41 (2022).

[1036] Council Directive 2019/790, Whereas clause (9).

[1037] *Id.* at art 4.

comparison, the Copyright Act of Japan does not require the legitimacy of works used for data analysis.[1038] The Copyright Act of Singapore even admits infringing works can be used for computational data analysis when necessary.[1039]

Moreover, if the TDM purposes are other than scientific research, the EU directive allows the right holders to opt out in an appropriate manner expressly. Foreseeably, this opt-out option will severely undermine the legislative purpose if we allow large studios to prevent rivals' text and data mining activities.[1040] By contrast, the Japan and Singapore Copyright Act do not provide right holders such an opt-out option. These would be better legislation considering potential film data hold-up problems.

It is critical to note that the EU's TDM limitation or exception can hardly extend to AI model training according to its definition.[1041] Article 30-4 of the Copyright Act of Japan does not mention model training either. But considering the purpose of model training is not "to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in that work," it is interpreted that model training can be covered by Article 30-4.[1042] By comparison, Section 243 of the Singapore Copyright Act indicates TDM in

---

[1038] Seiseikei AI No Katsuyō, Chosakukensha O Mamoru Tame No Aratana Kisei Ga Hitsuyōda（生成系 AI の活用、著作権者を守るための新たな規制が必要だ）[Utilization of generative AI, new regulations are necessary to protect copyright holders], Shūgīn Gīn Kīta Kashi (衆議院議員きいたかし) [House Representative Takashi Kii] (Apr. 24, 2023), https://kiitaka.net/21312/ [https://perma.cc/Y94X-DRQQ].
[1039] Copyright Act 2021 § 244 (Sing.).
[1040] More critics, please see Thomas Margoni & Martin Kretschmer, *A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology*, 71(8) GRUR Int'l 685 (2022).
[1041] Council Directive 2019/790, art 2 (2).
[1042] Chiteki Zaisan Senryaku Honbu Kenshō Hyōka Kikaku Īnkai Aratana Jōhōzai Kentō Īnkai (知的財産戦略本部検証・評価・企画委員会 新たな情報財検討委員会) [Intellectual Property Strategy Headquarters Verification/Evaluation/Planning Committee New Information Property Review Committee], AI Seiseibutsu To Chosakuken Nitsuite (AI 生成物と著作権について) [About AI products and copyright] 2 (Mar. 2023), https://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kousou/2023/dai2/siryou1.pdf [https://perma.cc/N8GD-MQN5].

Section 243(a) and indicates model training in Section 243(b) clearly and respectively (see **Table 6-3**). This is a more comprehensive legislation. Therefore, I suggest the United States, China, and Taiwan considering the Singapore model.

Table 6-3 The Comparison of EU Directive and Singapore Copyright Act

Article 2 (2) of Council Directive (EU) 2019/790	Section 243 of Singapore Copyright Act
'text and data mining' means any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations	In this Division, "computational data analysis", in relation to a work or a recording of a protected performance, includes — (a) using a computer program to identify, extract and analyse information or data from the work or recording; and (b) using the work or recording as an example of a type of information or data to improve the functioning of a computer program in relation to that type of information or data. *Illustration* An example of computational data analysis under paragraph (b) is the use of images to train a computer program to recognise images.

3. Review of International Treaties Obligations

Article 9(2) of the Berne Convention permits national legislation to exempt reproduction rights under certain conditions.[1043] Additionally, Article 10(2) of the World

---

[1043] Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as revised at Paris on July 24, 1971 and amended in 1979, S. Treaty Doc. No. 99-27 (1986) [hereinafter Berne Convention] § 9(2) provides: "It shall be a matter for legislation in the countries of the Union to permit the reproduction of such works in certain special cases, provided that such reproduction does not conflict with a normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author."

Copyright Treaty ("WCT") has a general limitation or exception clause to the rights conferred under the Berne Convention. But it is understood that Article 10(2) is a mere interpretation tool that does not change the scope of the limitations or exceptions permitted by the Berne Convention.[1044] In addition, Articles 1 through 21 of the Berne Convention (1971) are incorporated into the Agreement on Trade-Related Aspects of Intellectual Property Rights ("TRIPS Agreement"). [1045] Therefore, all WTO members' national legislation regarding limitations and exceptions to reproduction right shall comply with the so-called "three-step test" under Article 9(2) of the Berne Convention and Article 13 of the TRIPS agreement.[1046] The United States, China, and Taiwan are no exception as WTO members.

Under the "three-step test," any limitations or exceptions to reproduction rights shall: (1) confine to certain special cases, (2) not conflict with a normal exploitation of the work, and (3) not unreasonably prejudice the legitimate interests of the right holder. TDM and model training are confined to special cases. However, there may be wide disagreements with the second and third tests. To err on the side of caution, Article 30-4 of the Copyright Act of Japan provides that the data analysis limitation does not apply "if the action would unreasonably prejudice the interests of the copyright owner in light of the nature or purpose of the work or the circumstances of its exploitation." Similarly, whereas

---

[1044] WORLD INTELLECTUAL PROPERTY ORGANIZATION, GUIDE TO THE COPYRIGHT AND RELATED RIGHTS TREATIES ADMINISTERED BY WIPO AND GLOSSARY OF COPYRIGHT AND RELATED RIGHTS TERMS 212 (2004).
[1045] Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 U.N.T.S. 3; 33 I.L.M. 1197 (1994) [hereinafter TRIPS Agreement] § 9.1.
[1046] TRIPS agreement § 13 provides: "Members shall confine limitations or exceptions to exclusive rights to certain special cases which do not conflict with a normal exploitation of the work and do not unreasonably prejudice the legitimate interests of the right holder."

clause 6 of Council Directive (EU) 2019/790 states:[1047]

> The exceptions and limitations provided for in this Directive seek to achieve a fair balance between the rights and interests of authors and other rightholders, on the one hand, and of users on the other. They can be applied only in certain special cases that do not conflict with the normal exploitation of the works or other subject matter and do not unreasonably prejudice the legitimate interests of the rightholders.

Unfortunately, these redundant languages will increase uncertainties and create new questions in applying TDM limitations. Interestingly, Singapore seems very confident of its legislation because it does not put similar language in its related statute text. Some commentators argue that international policymakers shall enjoy fair freedom to regulate TDM because it is a "new category of copying that falls outside the scope of international copyright harmonization." [1048] TDM was unforeseeable when international diplomatic negotiations were wrapped up.

B. FINAL WORK: VOLUNTARILY NEGOTIATED LICENSE WITH BACKUP COMPULSORY LICENSE

Motion pictures aim to entertain the audience and are usually advertised. Motion pictures compete for audience attention, and copyright infringement is easily detected. Unlike using film materials for TDM and model training that do not compete with motion pictures' normal exploitation, using film materials in a final motion picture is completely different and thus requires discriminated treatment.

[1047] Council Directive 2019/790, Whereas clause (6), 2019 O.J. (L 130), 92–125.
[1048] Martin Senftleben, *Compliance of National TDM Rules with International Copyright Law: An Overrated Nonissue?*, 53 IIC, 1477 (2022).

1. Scheme

My second proposal is to create a voluntarily negotiated license backed up with a compulsory license after a period from the subject motion picture's first publication. In the event that a negotiated license fails, the users can seek a temporary compulsory license from competent agencies until a voluntarily negotiated license is reached after that. But duplicating a preexisting motion picture fixed by another without new human authorship is not allowed for a compulsory license. I will explain more details of the license framework (see **Table 6-4**) in the following sections.

Table 6-4 Motion Picture License Framework

Motion Picture	Initial Period (No compulsory license)	Royalty
High Human-Authorship	Long	High
Low Human-Authorship	Short	Low

- Voluntarily negotiated license (first) → compulsory license (backup)
- Single or blanket license allowed.
- Film artists guilds determine the degree of human-authorship.
- Credited film artists participate in the division of royalty fee

a. Justifications of the Backup Compulsory License

There are several reasons to support a backup compulsory license: (1) data-driven motion pictures' possible thin human original authorship deserving weaker copyright protection, (2) the limited creative freedom of motion pictures as public entertainment means, (3) inspiring more motion picture version choices for the audiences, and (4) saving

judicial resources. When a user doubts his or her use of film materials in their final motion pictures more than taking the ideas and their expressions are substantially similar to the copied film materials, instead of going to the court to argue the copied film materials' copyrightability and scope of copyright, they can choose the new mechanism to obtain a license in an easier, cheaper, and quicker way.

b. Mitigate the Blocking Effects of Long Copyright Terms

Moreover, the safety net compulsory license can mitigate the blocking effects of long copyright terms. Generally, in the United States, the copyright term of a motion picture made for hire endures for a term of 95 years from the year of its first publication; a motion picture not made for hire endures for a term consisting of the author's life and 70 years after the author's death.[1049] In China and Taiwan, the term of motion pictures' economic or property rights endure for 50 years after their first publication.[1050]

Some critics have long claimed that the current copyright term is overlong and propose a shorter term, like 14 years, 20 years, or 32.2 years.[1051] In addition, some scholars advocate a shorter term because filmmaking investment risk is reduced due to data analytics.[1052] However, the Berne Convention protects motion pictures' copyright at least 50 years after their first publication.[1053] Member countries cannot cut the copyright term to

---

[1049] 17 U.S.C. § 302.

[1050] *See* Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment) § 23; Zhuzuoquan Fa (著作權法) [Copyright Act (Taiwan)] § 34.

[1051] LAWRENCE LESSIG, FREE CULTURE: THE NATURE AND FUTURE OF CREATIVITY 292-293 (2005).

[1052] Raustiala & Sprigman, *supra* note 6, at 1604.

[1053] Berne Convention (1971) § 7(2).

less than a term of 50 years. Therefore, creating a voluntarily negotiated license backed up with a compulsory license might be a plausible alternative.

c. Terms and Conditions of Licenses

Any filmmakers seeking to exploit copyrighted film materials in their motion pictures and any copyright owners of film materials may negotiate and agree upon the terms and rates and the division of fees. Both sides may designate common agents to negotiate, agree to pay, or receive such royalty payments. These collective negotiations waive any antitrust liabilities. The licensed film materials could be a story, a script, a plot, a scene, sequences of events, special visual effects, or a character. This means if an author of literary or artistic work has authorized his or her work to be adapted into a first motion picture, his or her work will be subject to the above license. Single licenses or blanket licenses are allowed. Permitted new uses include remaking new motion pictures or preparing derivative motion pictures.

However, considering the high production costs of motion pictures, it is reasonable to reserve an initial period for film producers to recoup their investment without worrying the competition from the compulsory licensees. The length of the initial period can be left to Congress's discretion or the interested groups' consent. Except for musical works, film materials will be open to compulsory licenses after the initial period expires.

d. Incentives to Retain Human Artists

To encourage hiring human artists, I suggest granting a longer initial period (no compulsory license allowed) and a higher royalty rate to motion pictures created with high human authorship. Honoring the long tradition of film artists guilds in arbitrating over

screen credits and originality disputes,[1054] the writers and directors guilds will be given the authority to make verification thresholds and certify the degree of human authorship. Competent arbitrators like the Copyright Royalty Judges shall respect the film artists guilds' final decisions over the degree of human authorship. To honor human artists' creativity contribution, human film artists are entitled to participate in the division of royalty fee even though they do not own the copyright.

2. Referenced National and Regional Legislation

a. Musical Works Mechanical Reproduction Compulsory License

There has yet to be any known national and regional legislation granting compulsory licenses for reusing film materials in another motion picture. However, many national laws provide compulsory licenses of musical works for sound recording producers to make new sound recordings on certain conditions. For example, section 115 of the U.S. Copyright Act, Article 42 of the Copyright Law of China, and Article 69 of the Copyright Act of Taiwan. The following paragraphs will briefly illustrate the legislative history of section 115 of the U.S. Copyright Act.

Shortly after the U.S. Supreme Court's 1908 *White-Smith Music* decision held that making mechanical reproduction of musical compositions by perforated rolls did not constitute copying,[1055] section 1(e) of the Copyright Act of 1909 extended copyright to mechanical reproduction. Simultaneously, it conditioned such mechanical reproduction rights to a compulsory license.[1056] This compulsory license wished to prevent the sound

---

[1054] *See* DECHERNEY, *supra* note 125, at 88, 128.
[1055] White-Smith Music Pub. Co. v. Apollo Co., 209 U.S. 1, 18 (1908).
[1056] Copyright Act of 1909, Pub. L.60-349, 35 Stat. 1075.

recording market from being controlled by one company, the Aeolian Co., which allegedly

had made numerous long-term exclusive agreements with numerous music publishers in

the hope that new laws would recognize mechanical reproduction rights. [1057] Other

recording manufacturers did not oppose paying music composers in exchange for the right

to use music without fearing a potential mechanical reproduction monopoly.[1058]

Unquestionably, compulsory license of musical works was a political compromise

to settle different voices and conflicting interests of pro or against extending copyright

control groups. Since then, challenges to compulsory license of musical works have never

ceased due to the complaints of depriving music composers of creative control over their

works, dishonest accounting booking and production reports, royalty calculation basis, and

the royalty payment default of some irresponsible recording manufacturers.[1059] However,

assurance of equal access to music works and relative flourishing of the music industry in

countries with compulsory licenses compared to countries without compulsory licenses

keep the music mechanical reproduction compulsory license retained and evolving with

new recording technologies.[1060]

b. Special Considerations for Motion Picture Adaptations

Some commentators have long argued it was a historical mistake not to initially

impose a compulsory license for film adaptations by following the compulsory license

example of music mechanical reproduction.[1061] Of course, there are inherent differences

---

[1057] Harry G. Henn, *The Compulsory License Provisions of the U.S. Copyright Law*, 5 COPYRIGHT LAW REVISION STUDY 3 (1956).
[1058] *Id.*
[1059] *Id.* at 21-23.
[1060] *Id.* at 24.
[1061] *See* DECHERNEY, *supra* note 125, at 56.

between sound recordings and audiovisual recordings. Motion pictures are more complicated than sound recordings. Moreover, why cannot it extend to theatrical performances if a compulsory license for making motion pictures is allowed? Despite that, as discussed in Chapter Two and Chapter Three, we have reasons to believe motion picture studios are more likely to use data and AI technologies than theaters. It is good timing to rethink the old advice in the data-driven era.

But we must address some unique needs when designing compulsory licenses for motion pictures. Unlike a sound recording that usually loyally performs the musical work as the music sheet directs, it is not rare that a motion picture adapts a novel to something different but only keeps the title intact. Thus, my proposal permits film adaptations. Also, making motion pictures is more risky and costly than sound recordings. Therefore, I design an initial period disallowing compulsory licenses to let the copyright owners recoup their filmmaking investment as much as possible. These two characteristics make my proposal different from the music mechanical reproduction compulsory license.[1062] Although my proposal gives the filmmakers the right to adapt the film materials, it does not mean the licensees can do whatever they want. The moral right or similar legal concepts in American law will still constitute limitations.[1063]

3. Review of International Treaties Obligations

There are only two provisions in the Berne Convention that expressly sanction compulsory licenses, i.e., Article 11bis and Article 13(1). Unlike Article 13(1) of Berne permitting compulsory license to make sound recordings of musical works and attached

---

[1062] *See* LEAFFER, *supra* note 80, at 313.
[1063] *Id.* at 385.

lyrics after the authors of musical works and the lyrics have authorized others to make a first phonogram, [1064] Article 14(3) of the Berne Convention clearly prescribes: "The provisions of Article 13(1) shall not apply." It is understood compulsory licensing of musical works does not apply to reproducing music in films and "[p]ermission is always required."[1065] Therefore, compulsory licensing of using music works in a motion picture (not a sound recording but an audiovisual recording) is banned by Berne.[1066]

With respect to literary and artistic works and cinematographic works, Article 14 of Berne does not expressly permit or forbid compulsory licensing of them for cinematographic adaptation. However, "the existence of expressly-approved compulsory licenses" in Article 11bis and Article 13(1) of Berne "makes it difficult to infer the existence of an implied compulsory license" for using literary and artistic works and cinematographic works in cinematographic productions. [1067] Accordingly, compulsory license of literary and artistic works and cinematographic works for cinematographic production does not comply with Berne.

Historically, a similar Berne conflict issue happened to the old section 116 of the U.S. Copyright Act of 1976.[1068] The old section 116 allowed compulsory licensing of musical works on a jukebox for public performance was considered an impediment to U.S.

[1064] Berne Convention (1971) § 9(2) provides: "It shall be a matter for legislation in the countries of the Union to permit the reproduction of such works in certain special cases, provided that such reproduction does not conflict with a normal exploitation of the work and does not unreasonably prejudice the legitimate interests of the author."

[1065] WORLD INTELLECTUAL PROPERTY ORGANIZATION, GUIDE TO THE BERNE CONVENTION FOR THE PROTECTION OF LITERARY AND ARTISTIC WORKS (PARIS ACT, 1971) 84 (1978).

[1066] WORLD INTELLECTUAL PROPERTY ORGANIZATION, GUIDE TO THE COPYRIGHT AND RELATED RIGHTS TREATIES ADMINISTERED BY WIPO AND GLOSSARY OF COPYRIGHT AND RELATED RIGHTS TERMS 88 (2004).

[1067] Chapter III - Jukebox License, 10 COLUM.-VLA J.L. & ARTS 533, 534 (1985).

[1068] *Id.*

entry into Berne because Berne did not expressly sanction compulsory licensing of musical work for public performance.[1069] To harmonize Berne, section 4(a)(4) of the Berne Convention Implementation Act of 1988 introduced a new voluntarily negotiated license (numbered 17 U.S.C. §116A) backed up with the original compulsory license provisions of old section 116 in the event that voluntarily negotiated license did not reach.[1070] Although it looked like a delusional solution, Senate Report to the 1988 Amendment demonstrated the compatibility of section 116A with Berne as followed:[1071]

> With respect to the safety net provisions – which apply in the event a voluntary agreement cannot be reached or lapses – Mr. Oman testified that S. 1301 satisfies the requirements of the Berne Convention by elevating negotiated licenses between copyright owners and jukebox owners above the compulsory licenses in the current law. He noted that "some Berne Union countries do regulate organizations representing authors and copyright proprietors, including the setting of fees" and that the safety net provision can be "justified as analogous to regulation of collective societies." Thus, he concluded, S. 1301 was compatible with Berne.

Based on the same reason and U.S. jukebox precedent, I argue that a voluntarily negotiated license backed up with a compulsory license for using film materials in a new motion picture does not contravene Berne. It can be "justified as analogous to regulation of collective societies."[1072]

[1069] *See* LEAFFER, *supra* note 80, at 367.

[1070] *See id.* at 571.

[1071] S. REP. No. 100-352, at 27 (1988).

[1072] Section 116 A was renumbered to section 116 and the old section 116 (jukebox compulsory license) was thus completely repealed by the Copyright Royalty Tribunal Reform Act of 1993. But it is critical to note that "[i]n the event no such agreement is reached, however, the bill provides that copyright arbitration panels will be convened . . . . this will provide an adequate safeguard to jukebox operators that they will be able to continue to perform nondramatic musical works as in the past." *See* Copyright Royalty Tribunal Reform Act of 1993, Pub. L.103-198, 107 Stat. 2304; H.R. REP. No. 103-286, at 14 (1993). The copyright arbitration panels were later replaced by the Copyright Royalty Judges in accordance with the Copyright Royalty and Distribution Reform Act of 2004. *See* Copyright Royalty and Distribution Reform Act of 2004, Pub. L. 108–419, 118 Stat. 2341.

IV. CONCLUSION

Since the 1910s, Hollywood has measured audience preferences with rough industry-created methods. In the 1940s, scientific audience research led by George Gallup started to conduct film audience surveys with traditional statistical and psychological methods. However, the quantity, quality, and speed are limited. Things dramatically change in the internet age. The prevalence of digital data increases the convenience, instantaneousness, volume, width, and depth of collecting audience data and content data. Advanced data and AI technologies have also allowed machines to produce ideas or make human-like expressions. This brings new copyright challenges in the data-driven era.

No matter whether data and AI technologies are good or bad for human society, they shall be regulated by other legal mechanisms rather than being controlled by copyright.[1073] Copyright shall not be the gatekeeper of new technologies that do not impair the original uses of copyrighted works in the existing market. Given the high and uncertain copyright violation risks in the data-driven creation process, small and cautious motion picture producers will be kept out of data and AI technologies. However, it is unthinkable that a copyrighted work is created solely for TDM and machine learning purposes. I cannot see how making copies of these materials for TDM and model training will conflict with a normal exploitation of these works or unreasonably prejudice the legitimate interests of these right holders.

[1073] For example, the EU has proposed a comprehensive Artificial Intelligence Act to regulate the creation and use of AI. *See* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 [https://perma.cc/T65X-PHM9].

In order to ensure equal access to data and AI technologies, this study proposes to create a generous TDM and model training limitations or exceptions to reproduction and distribution rights and highly recommend the Singapore model. Some critics may argue creating such limitations or exceptions will encourage the use of data and AI and hence hurt human artists. This argument does not merit because the reality is that most film copyrights are controlled by large motion picture studios instead of film artists. Copyright can do nothing to prevent large studios from using data and AI technologies. On the contrary, copyright will empower these large studios to prevent rivals from accessing data and AI technologies to compete with them and become an even more powerful monopoly than today.

Unlike using film materials for TDM and model training in the creation process, using film materials in a final motion picture for the audience's entertainment is entirely different. This study proposes creating a voluntarily negotiated license backed up by a compulsory license for reusing film materials in new motion pictures. Although data-driven works and human-authored works appear identical, their natures are poles apart. In some extreme cases, data-driven works might be pure facts and deserve no copyright protection, depending on how we define human intellectual involvement. To incentivize motion picture producers to retain more human film artists, the degree of human original authorship should be a significant factor in deciding the compulsory license royalty rate and the length of the initial period that disallows compulsory licenses. The degree of human original authorship must be certified by film artists guilds. If motion picture producers want more copyright protection, they should retain film artists and get certification from film artists guilds.

From the conventional copyright perspective, such a proposal is undoubtedly radical. Some critics may even argue it is unconstitutional. However, this solution will sound reasonable if we treat data-driven work as a new category of work that differs from earlier ones. Recognizing the limited creative choices for motion pictures as public entertainment media, we must prevent their creative freedom from being blocked by ubiquitous data-driven works with thin human original authorship. This study argues national policymakers should enjoy broad discretion to qualify data-driven work's copyright protection. It would be too late to wait until the courts are overwhelmed by inspecting each data-driven work's core of protected material.

Appendix A: A Questionnaire for *Return of the Jedi* (1983)

SNEAK PREVIEW   RETURN OF THE JEDI   QUESTIONNAIRE

1) ☐ MALE   Under 10   11-15   16-20   21-25   26-35   36-45   46-55   56 & Over

    ☐ FEMALE   ☐   ☐   ☐   ☐   ☐   ☐   ☐   ☐

2) How would you rate this film? ☐ Excellent   ☐ Good   ☐ Fair   ☐ Poor

3) Have you seen STAR WARS? _____   THE EMPIRE STRIKES BACK? _____

4) List in order of preference your favorite characters (1-11):

    ☐ Luke Skywalker   ☐ Princess Leia   ☐ Han Solo   ☐ Darth Vader

    ☐ Chewbacca   ☐ Yoda   ☐ Lando Calrissian   ☐ C3PO   ☐ R2D2

    ☐ Ben (Obi-Wan) Kenobi   ☐ Boba Fett

5) List in order of preference your favorite new characters (1-9):

    ☐ The Emperor   ☐ Ewoks   ☐ Jabba the Hutt   ☐ Jabba's Palace Band

    ☐ Pig Guards   ☐ Nien Nunb (Lando's co-pilot)   ☐ Bib Fortuna (Jabba's

    right-hand man)   ☐ Admiral Ackbar (Commander of Rebel Fleet)

    ☐ Other_____

6) List in order of preference your favorite scenes (1-10):

    ☐ Jabba's Palace   ☐ Rancor Pit Monster   ☐ Yoda & Luke (Dagobah)

    ☐ Desert Sarlacc Pit & Skiff Battle   ☐ Luke & Ben in Forest

    ☐ Leia Finds the Ewok   ☐ Speeder Bike Chase   ☐ Luke Fights Vader

    ☐ Space Battle   ☐ Forest Battle

7) What did you especially like about this movie?_____

_____

8) Was there anything you disliked about this movie?_____

_____

9) What is your all-time favorite film?_____

10) How often do you go to the movies? ☐ Every week   ☐ 2-3 times a week

    ☐ Once a month   ☐ Less than once a month

11) Comments_____

_____

_____

A.54

243

Appendix B: A Questionnaire for *The Nines* (2007)

THE NINES

August 30, 2006

Did you like it?   YES   SORT OF   NOT REALLY   NO

Given a pair of magical scissors, is there anything you'd snip out?

Anything you'd like more of?

Were there any moments/scenes you particularly liked? (please list)

Were there any moments/scenes you particularly disliked, or felt didn't work? (please list)

Were there any moments you felt annoyed, frustrated or confused-in-a-bad-way by the movie? (please list)

How would you rate the following elements? You can elaborate in the margins, or the space below.
(1 = Excellent , 5 = Poor)

	← Better Worse →				
The ending	1	2	3	4	5
The music	1	2	3	4	5
The humor	1	2	3	4	5
The pace	1	2	3	4	5
The story	1	2	3	4	5
The beginning	1	2	3	4	5
The drama	1	2	3	4	5
Part One	1	2	3	4	5
Part Two	1	2	3	4	5
Part Three	1	2	3	4	5
Ryan Reynolds (Gary, Gavin, Gabriel)	1	2	3	4	5
Hope Davis (Sarah, Susan, Sierra)	1	2	3	4	5
Melissa McCarthy (Margaret, Melissa, Mary)	1	2	3	4	5
Elle Fanning (Noelle)	1	2	3	4	5

Any other comments or suggestions? We're all ears.

Thank you very, very much for filling this out. It's a big help.

Appendix C: ChatGPT Generated Script (Virtual Verdict)

Virtual Verdicts

June 27, 2023

Write a film script based on the words: law school students, online bar exam, covid pandemic, love.

Title: Virtual Verdicts

Genre: Drama/Romance

Logline:
In the midst of a global pandemic, a group of law school students prepares for an unprecedented online bar exam. As they navigate the challenges of studying remotely and the uncertainties of the COVID-19 era, two students find love and support in each other, giving them the strength to face their fears and fight for their dreams.

INT. LAW SCHOOL LIBRARY - DAY

The bustling law school library is filled with determined students pouring over textbooks, their faces hidden behind masks.

We focus on EMMA, a diligent and intelligent law student, and ADAM, a charismatic and charming law student.

EMMA, 23, sits alone at a table, scrolling through her laptop, researching online bar exam preparations. ADAM, 25, notices her from across the room.

ADAM
(whispering)
Hey, is this seat taken?

EMMA looks up, startled, but her eyes soften as she notices ADAM's warm smile.

EMMA
(smiling)
No, go ahead.

ADAM sits down, and they strike up a conversation about the challenges of studying during the COVID-19 pandemic and their fears about the online bar exam.

INT. EMMA'S APARTMENT - NIGHT

EMMA sits at her desk, surrounded by books and papers, feeling overwhelmed. Her laptop screen illuminates her determined face.

EMMA (V.O.)
This online bar exam is turning my life upside down. I don't know if I can do it.

Just then, her phone buzzes with a notification. It's a text from ADAM.

ADAM (TEXT)
You're stronger than you think. Believe in yourself. I believe in you.

EMMA smiles and takes a deep breath, finding comfort and inspiration in ADAM's message.

INT. VIRTUAL CLASSROOM - DAY

EMMA and ADAM, along with their fellow law school students, log into the virtual classroom for a bar exam preparation lecture. The professor, PROFESSOR WILSON, appears on the screen.

PROFESSOR WILSON
Good morning, everyone. Today, we'll be discussing contract law...

EMMA spots ADAM's video feed and catches his eye. They exchange encouraging smiles, their connection growing stronger.

INT. COFFEE SHOP - DAY

EMMA and ADAM sit across from each other at a cozy coffee shop, their textbooks and notes scattered on the table.

ADAM
You know, studying law during a pandemic feels like an uphill battle, but having you by my side makes it easier.

EMMA blushes, grateful for his support.

EMMA
I feel the same way. We can do this together.

They hold hands, united in their determination to overcome the challenges ahead.

INT. ONLINE EXAMINATION HALL - DAY

EMMA and ADAM sit in their respective homes, facing their laptops. The virtual bar exam begins, and tension fills the air.

EMMA (V.O.)
This is it. Months of hard work and sacrifices come down to this moment.

As they navigate through the exam questions, they exchange reassuring glances over video chat, their love and support fueling their determination.

INT. GRADUATION CEREMONY - DAY

EMMA and ADAM, dressed in their graduation gowns, stand among their law school peers. Their eyes meet, filled with pride and love.

DEAN WILLIAMS, a distinguished figure, addresses the graduates.

DEAN WILLIAMS

Today, we celebrate your resilience, your adaptability, and your unwavering commitment to the pursuit of justice. You have overcome extraordinary challenges to be here.

EMMA and ADAM hold hands tightly, a silent acknowledgment of the obstacles they've faced and the love that carried them through.

EXT. GRADUATION PARTY - NIGHT

EMMA and ADAM dance together, surrounded by their friends and loved ones, celebrating their accomplishments.

EMMA
(whispering)
I couldn't have done it without you.

ADAM
(kissing her forehead)
And I couldn't have done it without you. We make a great team.

They share a loving embrace, their journey as law school students in the midst of a pandemic now a chapter in their lives, forever marked by resilience, love, and triumph.

FADE OUT.

Appendix D: ChatGPT Generated Script (Love Amidst Bar Exam)

Love Amidst Bar Exam

July 1, 2023

Write a film script based on the words: law school students, online bar exam, covid pandemic, love.

Title: A New Verdict

Genre: Drama/Romance

Logline:
Amidst a global pandemic, a group of law school students navigates the challenges of preparing for an online bar exam. As they strive to achieve their professional dreams, they unexpectedly find love and support in each other, proving that even in the toughest times, love can still triumph.

INT. LAW SCHOOL LIBRARY - DAY

A group of LAW SCHOOL STUDENTS sits at a long table, studying books and legal documents. Laptops and online study materials are scattered everywhere. The atmosphere is tense.

JULIA (late 20s, determined but anxious) looks up from her notes and notices ALEX (early 30s, charismatic and supportive) sitting across from her.

JULIA
(stressed)
Alex, do you think we'll be ready for this online bar exam? The pressure is overwhelming.

ALEX
(sincerely)
I understand, Julia, but we've come this far. We're going to conquer this exam together. We have to believe in ourselves.

They share a brief moment of connection.

INT. VIRTUAL CLASSROOM - NIGHT

The students gather for an online class, their faces displayed in individual video windows. The professor explains the challenges of conducting the bar exam remotely due to the COVID PANDEMIC.

PROFESSOR
(through video)
Class, due to the ongoing pandemic, the bar exam will be administered online. It's going to be tough, but I believe in each and every one of you.

The students exchange worried glances but remain determined.

INT. VIRTUAL STUDY GROUP - DAY

Julia initiates a virtual study group with her classmates, including Alex, LUCAS (ambitious and focused), and EMMA (strong-willed and compassionate). They discuss legal concepts and share study tips.

EMMA
(excitedly)
Guys, I found a helpful online resource! It breaks down the complex concepts we're struggling with.

They all eagerly take notes, bonding over their shared determination to succeed.

EXT. PARK - DAY

Julia and Alex meet for a study break. They stroll through the park, enjoying a moment of respite from their intense preparation.

JULIA
(softly)
Alex, I appreciate your constant support. I couldn't have made it this far without you.

ALEX
(gently)
Julia, we've been through so much together. I want you to know that I've fallen in love with you, too. I believe in us, not just as study partners but as something more.

They share a heartfelt embrace, realizing their connection goes beyond the bar exam.

INT. VIRTUAL EXAM ROOM - DAY

The day of the online bar exam arrives. The students, including Julia and Alex, sit at their respective desks, laptops open. Anxiety fills the air as they prepare for the momentous test.

INT. LAW SCHOOL AUDITORIUM - DAY (MONTHS LATER)

A graduation ceremony takes place, celebrating the students' successful completion of law school and passing the online bar exam.

As Julia, Alex, Lucas, and Emma walk across the stage, they receive their diplomas and celebrate their hard-earned achievements.

INT. RECEPTION HALL - NIGHT

A celebration is in full swing. Music plays, and laughter fills the room as the law school graduates enjoy their success.

Julia and Alex dance together, their eyes full of love and pride for each other.

JULIA
(happily)
Who would've thought that love would bloom amidst the chaos of law school and a pandemic?

ALEX
(grinning)
It's true what they say: love conquers all. We faced incredible challenges together, and now we get to build a future side by side.

They share a kiss, surrounded by their friends who have become their chosen family.

FADE OUT.

Appendix E: 2018 Amendment to the Copyright Act (Japan)

著作権法の一部を改正する法律（平成 30 年法律第 30 号）

2018 Amendment to the Copyright Act (Japan)

（著作物に表現された思想又は感情の享受を目的としない利用）

(Exploitation without the Purpose of Enjoying the Thoughts or Sentiments Expressed in a Work)

第三十条の四 著作物は、次に掲げる場合その他の当該著作物に表現された思想又は感情を自ら享受し又は他人に享受させることを目的としない場合には、その必要と認められる限度において、いずれの方法によるかを問わず、利用することができる。ただし、当該著作物の種類及び用途並びに当該利用の態様に照らし著作権者の利益を不当に害することとなる場合は、この限りでない。

Article 30-4 It is permissible to exploit a work, in any way and to the extent considered necessary, in any of the following cases, or in any other case in which it is not a person's purpose to personally enjoy or cause another person to enjoy the thoughts or sentiments expressed in that work; provided, however, that this does not apply if the action would unreasonably prejudice the interests of the copyright owner in light of the nature or purpose of the work or the circumstances of its exploitation:

一著作物の録音、録画その他の利用に係る技術の開発又は実用化のための試験の用に供する場合

(i) if it is done for use in testing to develop or put into practical use technology that is connected with the recording of sounds or visuals of a work or other such exploitation;

二情報解析（多数の著作物その他の大量の情報から、当該情報を構成する言語、音、影像その他の要素に係る情報を抽出し、比較、分類その他の解析を行うことをいう。第四十七条の五第一項第二号において同じ。）の用に供する場合

(ii) if it is done for use in data analysis (meaning the extraction, comparison, classification, or other statistical analysis of the constituent language, sounds, images, or other elemental data from a large number of works or a large volume of other such data; the same applies in Article 47-5, paragraph (1), item (ii));

三 前二号に掲げる場合のほか、著作物の表現についての人の知覚による認識を伴うことなく当該著作物を電子計算機による情報処理の過程における利用その他の利用（プログラムの著作物にあつては、当該著作物の電子計算機における実行を除く。）に供する場合

(iii) if it is exploited in the course of computer data processing or otherwise exploited in a way that does not involve what is expressed in the work being perceived by the human senses (for works of computer programming, such exploitation excludes the execution of the work on a computer), beyond as set forth in the preceding two items.

（電子計算機による情報処理及びその結果の提供に付随する軽微利用等）

(Minor Exploitation Incidental to Computerized Data Processing and the Provision of the Results Thereof)

第四十七条の五 電子計算機を用いた情報処理により新たな知見又は情報を創出することによつて著作物の利用の促進に資する次の各号に掲げる行為を行う者（当該行為の一部を行う者を含み、当該行為を政令で定める基準に従つて行う者に限る。）は、公衆への提供又は提示（送信可能化を含む。以下この条において同じ。）が行われた著作物（以下この条及び次条第二項第二号において「公衆提供提示著作物」という。）（公表された著作物又は送信可能化された著作物に限る。）について、当該各号に掲げる行為の目的上必要と認められる限度において、当該行為に付随して、いずれの方法によるかを問わず、利用（当該公衆提供提示著作物のうちその利用に供される部分の占める割合、その利用に供される部分の量、その利用に供される際の表示の精度その他の要素に照らし軽微なものに限る。以下この条において「軽微利用」という。）を行うことができる。ただし、当該

公衆提供提示著作物に係る公衆への提供又は提示が著作権を侵害するものであること（国外で行われた公衆への提供又は提示にあつては、国内で行われたとしたならば著作権の侵害となるべきものであること）を知りながら当該軽微利用を行う場合その他当該公衆提供提示著作物の種類及び用途並びに当該軽微利用の態様に照らし著作権者の利益を不当に害することとなる場合は、この限りでない。

Article 47-5 (1)A person undertaking an action as set forth in one of the following items that contributes to facilitating the exploitation of a work by creating new knowledge or information through computerized data processing (this includes a person undertaking a part of such an action; limited to one doing so in accordance with the standards prescribed by Cabinet Order) may exploit a work that has been made available or presented to the public (this includes a work that has been made available for transmission; the same applies hereinafter in this Article) (hereinafter in this Article and Article 47-6, paragraph (2), item (ii) referred to as an "available or presented work") (limited to a publicized work or a work made available for transmission), in any way and to the extent considered to be necessary in light of the purpose of the action set forth in the relevant item, when exploiting it incidental to the undertaking of that action (limited to exploitation that is minor in light of the percentage it constitutes of the part of the available or presented work that has been provided for exploitation, the volume of the part of that work that has been provided for exploitation, the accuracy of indications made at the time it was provided for exploitation, and other elements; hereinafter in this Article referred to as "minor exploitation"); provided, however, that this does not apply if the person undertakes that minor exploitation knowing that the available or presented work's having been made available or presented to the public constitutes copyright infringement (for a work made available or presented to the public abroad, this means that the action would constitute copyright infringement if it took place in Japan), nor does it apply if the action would otherwise unreasonably prejudice the interests of the copyright owner in light of the nature or purpose of the available or presented work or the circumstances of its minor exploitation:

一電子計算機を用いて、検索により求める情報（以下この号において「検索情報」という。）が記録された著作物の題号又は著作者名、送信可能化された検索情報

256

に係る送信元識別符号（自動公衆送信の送信元を識別するための文字、番号、記号その他の符号をいう。）その他の検索情報の特定又は所在に関する情報を検索し、及びその結果を提供すること。

(i) using a computer to search for the title or author name of a work in which information that a person is searching for (hereinafter in this item referred to as "information being searched for") has been recorded, for the transmitter identification code (meaning the letters, numbers, symbols, or any other code by which the transmitter of an automatic public transmission is identified) associated with information being searched for that has been made available for transmission, or for any other information concerning the identification or location of information being searched for; and making the results of that search available;

二電子計算機による情報解析を行い、及びその結果を提供すること。

(ii) undertaking computerized data analysis and furnishing the results of that analysis;

三前二号に掲げるもののほか、電子計算機による情報処理により、新たな知見又は情報を創出し、及びその結果を提供する行為であつて、国民生活の利便性の向上に寄与するものとして政令で定めるもの

(iii) an action that Cabinet Order prescribes as contributing to increased convenience in the lives of the citizenry by creating new knowledge or information through computerized data processing and making the results of this available, beyond what is set forth in the preceding two items.

２前項各号に掲げる行為の準備を行う者（当該行為の準備のための情報の収集、整理及び提供を政令で定める基準に従つて行う者に限る。）は、公衆提供提示著作物について、同項の規定による軽微利用の準備のために必要と認められる限度において、複製若しくは公衆送信（自動公衆送信の場合にあつては、送信可能化を含む。以下この項及び次条第二項第二号において同じ。）を行い、又はその複製物による頒布を行うことができる。ただし、当該公衆提供提示著作物の種類及

び用途並びに当該複製又は頒布の部数及び当該複製、公衆送信又は頒布の態様に照らし著作権者の利益を不当に害することとなる場合は、この限りでない。

(2) A person that prepares to undertake an action set forth in one of the items of the preceding paragraph (limited to a person that collects, organizes, and provides information in preparation to undertake the action in accordance with the standards prescribed by Cabinet Order) may reproduce or make public transmissions of an available or presented work (or make the relevant work available for transmission, if such transmission is being made via an automatic public transmission; the same applies in this paragraph and Article 47-6, paragraph (2), item (ii)) or distribute copies thereof, to the extent considered to be necessary in order to prepare for minor exploitation under the preceding paragraph; provided, however, that this does not apply if the action would otherwise unreasonably prejudice the interests of the copyright owner in light of the nature or purpose of the available or presented work, the number of copies that would be reproduced or distributed, or the circumstances of the reproduction, public transmission, or distribution.

Appendix F: Directive (EU) 2019/790

Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC

Article 2

Definitions

For the purposes of this Directive, the following definitions apply:. . . .

(2) 'text and data mining' means any automated analytical technique aimed at analysing text and data in digital form in order to generate information which includes but is not limited to patterns, trends and correlations;

. . . .

Article 4

Exception or limitation for text and data mining

1. Member States shall provide for an exception or limitation to the rights provided for in Article 5(a) and Article 7(1) of Directive 96/9/EC, Article 2 of Directive 2001/29/EC, Article 4(1)(a) and (b) of Directive 2009/24/EC and Article 15(1) of this Directive for reproductions and extractions of lawfully accessible works and other subject matter for the purposes of text and data mining.

2. Reproductions and extractions made pursuant to paragraph 1 may be retained for as long as is necessary for the purposes of text and data mining.

3. The exception or limitation provided for in paragraph 1 shall apply on condition that the use of works and other subject matter referred to in that paragraph has not been expressly reserved by their rightholders in an appropriate manner, such as machine-readable means in the case of content made publicly available online.

4. This Article shall not affect the application of Article 3 of this Directive.

Appendix G: Copyright Act 2021 (Singapore)

Copyright Act 2021 (Singapore)

Copying or communicating for computational data analysis

244.—(1) If the conditions in subsection (2) are met, it is a permitted use for a person (X) to make a copy of any of the following material:

(a) a work;

(b) a recording of a protected performance.

(2) The conditions are —

(a) the copy is made for the purpose of —

(i) computational data analysis; or

(ii) preparing the work or recording for computational data analysis;

(b) X does not use the copy for any other purpose;

(c) X does not supply (whether by communication or otherwise) the copy to any person other than for the purpose of —

(i) verifying the results of the computational data analysis carried out by X; or

(ii) collaborative research or study relating to the purpose of the computational data analysis carried out by X;

(d) X has lawful access to the material (called in this section the first copy) from which the copy is made; and

Illustrations

(a) X does not have lawful access to the first copy if X accessed the first copy by circumventing paywalls.

(b) X does not have lawful access to the first copy if X accessed the first copy in breach of the terms of use of a database (ignoring any terms that are void by virtue of section 187).

(e) one of the following conditions is met:

(i) the first copy is not an infringing copy;

(ii) the first copy is an infringing copy but —

(A) X does not know this; and

(B) if the first copy is obtained from a flagrantly infringing online location (whether or not the location is subject to an access disabling order under section 325) — X does not know and could not reasonably have known that;

(iii) the first copy is an infringing copy but —

(A) the use of infringing copies is necessary for a prescribed purpose; and

(B) X does not use the copy to carry out computational data analysis for any other purpose.

(3) To avoid doubt, a reference in subsection (1) to making a copy includes a reference to storing or retaining the copy.

(4) It is a permitted use for X to communicate a work or a recording of a protected performance to the public if —

(a) the communication is made using a copy made in circumstances to which subsection (1) applies; and

(b) X does not supply (whether by communication or otherwise) the copy to any person other than for the purpose of —

(i) verifying the results of the computational data analysis carried out by X; or

(ii) collaborative research or study relating to the purpose of the computational data analysis carried out by X.

(5) For the purposes of this Act, the supply of copies of any material in circumstances to which this section applies —

> (a) is not to be treated as publishing the material (or any work or recording included in the material); and

> (b) must be ignored in determining the duration of any copyright in the material (or the included work).

BIBLIOGRAPHY

1. Treaties, Statutes

Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, 1869 U.N.T.S. 3; 33 I.L.M. 1197 (1994).

Berne Convention for the Protection of Literary and Artistic Works, Sept. 9, 1886, as revised at Paris on July 24, 1971 and amended in 1979, S. Treaty Doc. No. 99-27 (1986).

Chosakukenhō [Copyright Act], Law No. 52 of 2021, art.30-4, 47-5 (Japan).

Copyright Act 2021 § 244 (No. 22 of 2021) (Sing.).

Copyright Act of 1909, Pub. L.60-349, 35 Stat. 1075.

Copyright Royalty and Distribution Reform Act of 2004, Pub. L. 108–419, 118 Stat. 2341.

Copyright Royalty Tribunal Reform Act of 1993, Pub. L.103-198, 107 Stat. 2304.

Council Directive 2019/790, 2019 O.J. (L 130), 92–125.

Motion Picture Copyright Amendments, Pub. L. No. 62-303, 37 Stat. 488 (1912).

Zhonghua Renmin Gongheguo Zhuzuoquan Fa (中华人民共和国著作权法) [Copyright Law of the People's Republic of China] (2020 amendment).

Zhuzuoquan Fa (著作權法) [Copyright Act] (Taiwan).

Zhuzuoquan Fa Diwutiao Diyixiang Gekuan Zhuzuo Neirong Liqi (著作權法第五條第一項各款著作內容例示) [The Illustrated Contents of Each Kind of Works in Paragraph One, Article 5 of the Copyright Act] (Taiwan) Zhonghua Renmin Gongheguo Zhuzuoquan Fa Shishi Tiaoli (中华人民共和国著作权法实施条例) [Regulation for the Implementation of the Copyright Law of the People's Republic of China] (2013 Revision).

2. Cases, Orders, Opinions, and Decisions

Abdin v. CBS Broad. Inc., 971 F.3d 57 (2d Cir. 2020).

Am. MuKalem Co. v. Harper Bros., 222 U.S. 55 (1911).

Am. Mutoscope & Biograph Co. v. Edison Mfg. Co., 137 F. 262 (C.C.D.N.J. 1905).

Am. Mutoscope & Biograph Co. v. Edison Mfg. Co., 137 F. 262 (C.C.D.N.J. 1905).

Arnstein v. Porter, 154 F.2d 464 (2d Cir. 1946).

Astor-White v. Strong, 817 F. App'x 502 (9th Cir. 2020).

Axelbank v. Rony, 277 F.2d 314 (9th Cir. 1960).

Baker v. Selden, 101 U.S. 99, 103 (1879).

Baodexin Guoji Renshou Baoxian Gufen Youxian Gongsi Su Daduhui Guoji Renshou Baoxian Gufen Youxian Gongsi (保德信國際人壽保險股份有限公司訴大都會國際人壽保險股份有限公司) [Prudential Life Ins. Co. v. Metro. Life Ins. Co.], 98 Ming Zu Shang Zi No.16 (98 年度民著上字第 16 號) (Intell. Prop. Ct. Feb. 4, 2010) (Taiwan).

Beijing Shi Gaoji Renmin Fayuan Qinhaiz Huzuoquan Anjian Shenli Zhinan (北京市高级人民法院侵害著作权案件审理指南) [Beijing High People's Court Guidelines for the Trial of Copyright Infringement Cases] (promulgated by the Beijing High People's Ct., Apr. 20, 2018.

Beijing Xinlang Hulian Xinxi Fuwu Youxian Gongsi Yu Beijing Tianying Jiuzhou Wangluo Jishu Youxian Gongsi Buzhengdang Jingzheng Jiufen (北京新浪互联信息服务有限公司与北京天盈九州网络技术有限公司不正当竞争纠纷) [Beijing SINA Internet Info. Serv. Co. v. Beijing Tianying Jiuzhou Network Tech. Co.], 2020 Jing Min Zai No. 28 (2020 京民再 128 号) (Beijing Mun. High People's Ct. Sept. 23, 2020) [法宝引证码 CLI.C.115093220].

Benay v. Warner Bros. Entm't, Inc., 607 F.3d 620 (9th Cir. 2010).

Berkic v. Crichton, 761 F.2d 1289 (9th Cir. 1985).

Bucklew v. Hawkins, Ash, Baptie & Co., LLP., 329 F.3d 923 (7th Cir. 2003).

Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53 (1884).

Cain v. Universal Pictures Co., 47 F. Supp. 1013, 1017 (S.D. Cal. 1942).

Caituan Faren Laihe Wenjiao Jijinhui Su Lianhe Baike Dianzi Chuban Youxian Gongsi (財團法人賴和文教基金會訴聯合百科電子出版有限公司) [Laiho Cultural and Educ. Found. v. United Digital Publ'n Co.], 108 Tai Shang Zi No.1315 (108 年度台上字第 1315 號) (Sup. Ct. 2019) (Taiwan).

Carlini v. Paramount Pictures Corp., No. 21-55213, 2022 WL 614044 (9th Cir. Mar. 2, 2022).

Cavalier v. Random House, Inc., 297 F.3d 815 (9th Cir. 2002).

Childress v. Taylor, 945 F.2d 500 (2d Cir. 1991).

Chuanqi Wangle Youxi Gufen Youxian Gongsi Su Xingyu Hudong Yule Keji Gufen Youxian Gongsi (傳奇網路遊戲股份有限公司訴星宇互動娛樂科技股份有限公司) X-Legend Ent. Co. v. Loftstar Interactive Ent. Inc.), 107 Min Zhu Shang Yi Zi No.9 (107 年度民著上易字第 9 號) (Intell. Prop. Ct. Jan. 24, 2019) (Taiwan).

Chuck Blore & Don Richman Inc. v. 20/20 Advert. Inc., 674 F. Supp. 671 (D. Minn. 1987).

CK Co. v. Burger King Corp., 122 F.3d 1055 (2d Cir. 1995).

Computer Assocs. Int'l, Inc. v. Altai, Inc., 982 F.2d 693 (2d Cir. 1992).

Daniels v. Walt Disney Co., 958 F.3d 767 (9th Cir. 2020).

DC Comics v. Towle, 802 F.3d 1012, 1019 (9th Cir. 2015).

Design Data Corp. v. Unigate Enter., Inc., 847 F.3d 1169 (9th Cir. 2017).

Ding Fu Ting Weifan Zhuzuoquan Fa Anjian (丁富庭違反著作權法案件) [Republic of China v. Ding Fu Ting], 104 Xing Zhi Shang Yi Zi No. 95 (104 年刑智上易字第 95 號) (Intel. Prop. Ct. Mar. 10, 2016) (Taiwan).

Doe v. Github, Inc., No. 22-cv-06823-JST, 2023 U.S. Dist. LEXIS 86983 (N.D. Cal. May 11, 2023).

Edison v. Lubin, 119 F. 993 (C.C.E.D. Pa. 1903).

Edison v. Lubin, 122 F. 240 (3d Cir. 1903).

Emerson v. Davies, 8 F. Cas. 615 (C.C.D. Mass. 1845).

Esplanade Prods., Inc. v. Walt Disney Co., No. CV1702185MWFJCX, 2017 WL 5635024 (C.D. Cal. July 11, 2017).

Feist Publications, Inc. v. Rural Telephone Service Co., Inc., 499 U.S. 340 (1991).

Films, Inc. v. Time Warner Ent. Co., L.P., 462 F.3d 1072 (9th Cir. 2006).

Fu Xianjin Weifan Zhuzuoquan Fa Anjian (傅先進違反著作權法案件) [Republic of China v. Fu Xian Jin], 107 Xing Zhi Shang Su Zi No.1 (107 年度刑智上訴字第 1 號) (Intel. Prop. Ct. Dec. 27, 2018) (Taiwan).

FunkyTwentieth Century-Fox Film Corp. v. MCA, Inc., 715 F.2d 1327 (9th Cir. 1983).

Gaiman v. McFarlane, 360 F.3d 644 (7th Cir. 2004).

Gaoyang Su Jinse Shizu (Beijing) Yingshi Wenhua Youxian Gongsi Zhuzuoquan Quanshu Qinquan Jiufen (高阳诉金色视族（北京）影视文化有限公司著作权权属、侵权纠纷) [Gao Yang v. Golden Vision Co.], 2017 Jing 73 Min Zhong No. 797 (2017 京 73 民终 797 号) (Beijing IP Ct. Apr. 2, 2020) [法宝引证码 CLI.C.97510788].

Garcia v. Google, Inc., 786 F.3d 733 (9th Cir. 2015).

Gates Rubber Co. v. Bando Chem. Indus., Ltd., 9 F.3d 823 (10th Cir. 1993).

Guanyu Jixie Gufen Youxian Gongsi Weifan Zhuzuoquan Fa Anjian (冠昱機械股份有限公司違反著作權法案件) [Republic of China v. Guan-Yu Machine Co.], 105 Xing Zhi Shang Su Zi No. 42 (105 年度刑智上訴字第 42 號) (Intell. Prop. Ct. Apr. 16, 2020) (Taiwan).

Harper & Row Publishers, Inc. v. Nation Enterprises, 471 U.S. 539 (1985).

Helan Shang Endemol Nederland B.V. Zisu Quanneng Zhizuo Gufen Youxian Gongsi (荷蘭商 Endemol Nederland B.V.自訴全能製作股份有限公司) [Endemol Nederland

B.V. v. All Full Power Co.], 103 Xing Zhi Shang Yi No. 56 (103 年度刑智上易字第 56 號) (Intell. Prop. Sept. 3, 2015) (Taiwan).

Hoehling v. Universal City Studios, Inc., 618 F.2d 972 (2d Cir. 1980).

Houts v. Universal City Studios, Inc., 603 F. Supp. 26 (C.D. Cal. 1984).

Huang Pingying Su You Shujun (黃萍瑛訴游淑珺) [Huang Ping Ying v. You Shu Jun], 104 Min Zhu Su No.39 (104 年度民著訴字第 39 號) (Intell. Prop. Ct. Sept. 30, 2016) (Taiwan).

Idema v. Dreamworks, Inc., 90 F. App'x 496, 498 (9th Cir. 2003), as amended on denial of reh'g (Mar. 9, 2004).

Jiage Shipin Gufen Youxian Gongsi Su Tongfang Shengwu Keji Gufen Youxian Gongsi (佳格食品股份有限公司訴統芳生物科技股份有限公司) [Standard Foods Co. v. Top Greats Biotech Co.], 103 Tai Shang Zi No.1554 (103 年度台上字第 1544 號) (Sup. Ct. 2014) (Taiwan).

Jian Xuanzong Su Taiwan Aodaili Gufen Youxian Gongsi (簡玄宗訴台灣奧黛莉股份有限公司) [Jian Xuan Zong v. Sincerity Found. Mfg. Co.], 110 Min Zhu Su Zi No.106 (110 年度民著訴字第 106 號) (Intell. Prop. & Com. Ct. June 14, 2022) (Taiwan).

Jian YongCang Weifan Zhuzuoquan Fa Anjian (簡永倉違反著作權法案件) [Republic of China v. Jian Yong Cang], 104 Xing Zhi Shang Yi Zi No. 75 (104 年刑智上易字第 75 號) (Intell. Prop. Ct. May 13, 2016) (Taiwan).

Jiuye Qingbao Zixun Gufen Youxian Gongsi Su Yuantai Shuwei Keji Youxian Gongsi (就業情報資訊股份有限公司訴元太數位科技有限公司) [Career Consulting Co. v. Yuan Tai Tech., LLC.], 108 Min Zhu Shang Zi No. 12 (108 年度民著上字第 12 號) (Intell. Prop. Ct. June 18, 2020) (Taiwan).

Kelley v. Chicago Park Dist., 635 F.3d 290, 304 (7th Cir. 2011), Satava v. Lowry, 323 F.3d 805, 813 (9th Cir. 2003).

Kouf v. Walt Disney Pictures & Television, 16 F.3d 1042 (9th Cir. 1994).

KRISTINA KASHTANOVA, CORRESPONDENCE ID: 1-5GB561K (U.S. Copyright Off. Oct. 28, 2022).

Landsberg v. Scrabble Crossword Game Players, Inc., 736 F.2d 485 (9th Cir. 1984).

Longxing Keji Gufen Youxian Gongsi su Xuwei Wenhua Shiye Youxian Gongsi (龍行科技股份有限公司訴徐薇文化事業有限公司) [Long Xing technology Co. v. Xu Wei Culture Co.], 106 Tai Shang Zi No.2673 (106年度台上字第2673號) (Sup. Ct. Nov. 7, 2018) (Taiwan).

Marshall v. Yates, No. CV-81-1850-MML, 1983 WL 1148 (C.D. Cal. Oct. 26, 1983).

Meishang Meizhou Rentiantang Gufen Youxian Gongsi Su Jinfu Wanju Shiye Youxian Gongsi (美商美洲任天堂股份有限公司訴金富玩具實業有限公司) [Nintendo Co. v. Jin Fu Toy Ltd.], 110 Min Zhu Shang Zi No. 11 (110 年度民著上字第 11 號) (Intell. Prop. & Com. Ct. Nov. 18, 2021) (Taiwan).

Metcalf v. Bochco, 294 F.3d 1069 (9th Cir. 2002).

Miller v. Universal City Studios, Inc., 650 F.2d 1365 (5th Cir. 1981).

Montgomery v. NBC Television, 833 F. App'x 361 (2d Cir. 2020).

Narell v. Freeman, 872 F.2d 907 (9th Cir. 1989).

Naruto v. Slater, 888 F.3d 418, 426 (9th Cir. 2018).

Nichols v. Universal Pictures Corp., 45 F.2d 119 (2d Cir. 1930).

Reyher v. Children's Television Workshop, 533 F.2d 87 (2d Cir. 1976).

Rice v. Fox Broad. Co., 330 F.3d 1170 (9th Cir. 2003).

Rishang Wumu Shipin Gufen Youxian Gongsi Su Xinglin Shipin Gufen Youxian Gongsi (日商五木食品股份有限公司訴興霖食品股份有限公司) [Japanese Itsuki Foods Co. v. Sing-Lin Food Co.], 103 Tai Shang Zi No.1839 (103 年度台上字第 1839 號) (Sup. Ct. 2014) (Taiwan).

Salinger v. Random House, Inc., 811 F.2d 90 (2d Cir.), opinion supplemented on denial of reh'g, 818 F.2d 252 (2d Cir. 1987).

Sheldon v. Metro-Goldwyn Pictures Corp., 81 F.2d 49 (2d Cir. 1936).

Shenzun Shi Tengxun Jisuanji Xitong Youxian Gongsi Su Shanghai Yingxun Keji Youxian Gongsi Qinhai Zhu Zuoquan Ji Bu Hengdang Jingzheng Jiufen (深圳市腾讯计算机系统有限公司诉上海盈讯科技有限公司侵害著作权及不正当竞争纠纷) [Shenzhen Tencent Comput. Sys. Co. v. Shanghai Yingxun Tech. Co.], 2019 Yue 0305 Min Chu No. 14010 (2019 粤 0305 民初 14010 号) (Shenzhen Nanshan Dist. People's Ct. Dec. 24, 2019) [法宝引证码 CLI.C.97241331].

Shouying Chuangyi Gufen Youxian Gongsi Su Yuanjin Guoji Youxian Gongsi (首映創意股份有限公司訴原金國際有限公司) [Sofa Studio Co. v. Engine Studios LLC], 110 Min Zhu Shang Zi No. 4 (111 年度民著上字第 4 號) (Intell. Prop. & Com. Ct. Feb. 9, 2023) (Taiwan).

Sid & Marty Krofft Television Prods., Inc. v. McDonald's Corp., 562 F.2d 1157 (9th Cir. 1977).

Skidmore as Tr. for Randy Craig Wolfe Tr. v. Led Zeppelin, 952 F.3d 1051 (9th Cir. 2020).

Steven Thaler, Correspondence ID 1-3ZPC6C3; SR # 1-7100387071 (Copyright Rev. Bd. Feb. 14, 2022).

Stowe v. Thomas, 23 F. Cas. 201 (C.C.E.D. Pa. 1853).

Taiwan Intellectual Property Office, No. 1070110 Email (Jan. 10, 2018).

Taiwan Intellectual Property Office, No. 1081219 Email (Dec. 19, 2019).

Taiwan Intellectual Property Office, No. 1111031 Email (Oct. 31, 2022).

Taiwan Nuohua Gufen Youxian Gongsi Su Shengda Huaxue Zhiyao Gufen Youxian Gongsi (台灣諾華股份有限公司訴生達化學製藥股份有限公司) [Novartis (Taiwan) Co. v. Standard Chem. & Pharm. Co.], 105 Min Zhu Shang Zi No.4 (105 年度民著上字第 4 號) (Intell. Prop. Ct. Oct. 20, 2016) (Taiwan).

Thaler v. Perlmutter, No. CV 22-1564 (BAH), 2023 WL 5333236 (D.D.C. Aug. 18, 2023).

Tianyi Caijin Keji Fuwu Gufen Youxian Gongsi Su Zhongcheng Zixun Guanli Gufen Youxian Gongsi (天逸財金科技服務股份有限公司訴仲城資訊管理股份有限公司) [Vteam Financial Technology Service Co. v. Fortress System Co.], 102 Min Zhu Shang Zi No. 7 (102 年度民著上字第 7 號) (Intell. Prop. Ct. Dec. 22, 2016) (Taiwan).

TMTV, Corp. v. Mass Prods., Inc., 645 F.3d 464 (1st Cir. 2011).

Tremblay v. OPENAI, INC., 3:23-cv-03223 (N.D. Cal. 2023).

U.S. COPYRIGHT OFFICE, 37 CFR PART 202 COPYRIGHT REGISTRATION GUIDANCE: WORKS CONTAINING MATERIAL GENERATED BY ARTIFICIAL INTELLIGENCE, 88 FR 16190 (2023).

U.S. COPYRIGHT OFFICE, COMPENDIUM OF U.S. COPYRIGHT OFFICE PRACTICES (3d ed. 2021).

United States v. AT&T Inc., 310 F. Supp. 3d 161 (D.D.C. 2018).

United States v. Paramount Pictures, Inc., 334 U.S. 131 (1948).

Universal Pictures Co. v. Harold Lloyd Corp., 162 F.2d 354 (9th Cir. 1947).

Urantia Found. v. Kristen Maaherra, 114 F.3d 955 (9th Cir. 1997).

Van Lindberg on behalf of Kristina Kashtanova, Related Correspondence ID: 1-5GB561K, (U.S. Copyright Off. Feb. 21, 2023).

Walker v. Time Life Films, Inc., 784 F.2d 44 (2d Cir. 1986).

Wang Taichang Su Meishang Maigeluo Xier Guoji Gufen Youxian Gongsi Taiwan Fengongsi (王泰昌訴美商麥格羅．希爾國際股份有限公司台灣分公司) [Wang Tai Chang v. McGraw-Hill Education Co., Taiwan Branch], 104 Tai Shang Zi No.1251 (104 年度台上字第 1251 號) (Sup. Ct. 2015) (Taiwan).

Wangyin Guoji Gufen Youxian Gongsi Su Yinhe Xianshang Gufen Youxian Gongsi (網銀國際股份有限公司訴銀河線上股份有限公司) (Wanin Int'l Co., v. Galaxy Co.), 106 Min Zhu Su Zi No. 5 (106 年度民著訴字第 5 號) (Intell. Prop. Ct. Jan. 25, 2018) (Taiwan).

Warner Bros. Ent. v. X One X Prods., 644 F.3d 584 (8th Cir. 2011).

Warner Bros. Inc. v. Am. Broad. Companies, Inc., 654 F.2d 204 (2d Cir. 1981).

Warner Bros. Inc. v. Am. Broad. Companies, Inc., 720 F.2d 231 (2d Cir. 1983).

Warner Bros. Pictures v. Columbia Broad. Sys., 216 F.2d 945 (9th Cir. 1954).

Warner Bros. v. ABC, 654 F.2d 204 (2d Cir. 1981).

Wen Ruian Yu Zhoujing Deng Qinhai Zhuzuoquan Jiufen (温瑞安与周静等侵害著作权纠纷) [Wen Rui An v. Zhou Jing], 2017Jing 0105 Min Chu No.932 (2017 京 0105 民初 62752 号) (Beijing Chaoyang Dist. People's Ct. May 30, 2019) [法宝引证码 CLI.C.83245364].

White v. Twentieth Century Fox Corp., 572 F. App'x 475 (9th Cir. 2014).

White-Smith Music Pub. Co. v. Apollo Co., 209 U.S. 1, 18 (1908).

Williams v. Gaye, 895 F.3d 1106 (9th Cir. 2018).

Yang Wenhui Weifan Zhuzuoquan Fa Anjian (楊文輝違反著作權法案件) [Republic of China v. Yang, Wen Hui], 111 Tai Shang Zi No.2231 (111 年度台上字第 2231 號) (Sup. Ct., 2022) (Taiwan).

Yangshi Guoji Wangluo Youxian Gongsi Su Shanghai Juli Chuanmei Jishu Youxian Gongsi Zhuzuo Quan Qinquan Ji Buzhengdang Jingzheng Jiufen (央视国际网络有限公司诉上海聚力传媒技术有限公司著作权侵权及不正当竞争纠纷) [China Cent. Television Int'l Network Co. v. Shanghai SynaCast Media Tech. Co.], 2020 Hu Min Zhong No. 581 (2020 沪 73 民终 581 号) (Shanghai IP Ct. Apr. 20, 2021) [法宝引证码 CLI.C.410250170].

Zarya of the Dawn, Registration Number VAu001480196, https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?v1=1&ti=1,1&Search%5FArg=VAu001480196&Search%5FCode=REGS&CNT=25&PID=QYVn4FCjDLW8nt-grR26irPWTWWP&SEQ=20230819142409&SID=1.

Zeng Fanqi Su Zhonghua Dianshi Gufen Youxian Gongsi (曾繁祺訴中華電視股份有限公司) [Zeng Fan Qi v. Chinese Television Sys. Co.], 109 Min Zhu Su Zi No. 76 (109 年度民著訴字第 76 號) (Intell. Pro. Ct. Nov. 9, 2020) (Taiwan).

Zhang Xiaoyan Su Lei Xianhe Zhaoqi Sanding Aishuren Yinxiang Tushu Youxian Gongsi (张晓燕诉雷献和、赵琪、山东爱书人音像图书有限公司) [Zhang Xiaoyan v. Lei Xianhe, Zhao Qi, and Shandong Aishuren DVDs & Books Co., Ltd.], 2013 Min Shen Zi No. 1049 (2013 民申字第 1049 号) (Sup. People's Ct. 2014).

Zhiguan Keji Gufen Youxian Gongsi Su Heluo Youxi Youxian Gongsi (智冠科技股份有限公司訴河洛遊戲有限公司) [Soft-World Int'l Co. v. Heluo Games Co.], 108 Min Zhu Shang Zi No.3 (108 年度民著上字第 3 號) (Intell. Prop. Ct. July 30, 2010) (Taiwan).

Zhu Peixuan Weifan Zhuzuoquan Fa Anjian (朱珮萱違反著作權法案件) [The Republic of China v. Zhu Pei Xuan], 108 Xing Zhi Shang Su Zi No. 41 (108 年度刑智上訴字第 41 號) (Intell. Prop. Ct. Nov. 14, 2019) (Taiwan).

3. Legislative Materials

H. R. REP. No. 62-756 (1912).

H.R. 15263, 62nd Cong. (1911).

H.R. 20596, 62nd Cong. (1912).

H.R. REP. No. 103-286 (1993).

H.R. REP. No. 62-756 (1912).

Harry G. Henn, *The Compulsory License Provisions of the U.S. Copyright Law*, 5 COPYRIGHT LAW REVISION STUDY (1956).

Motion Picture Copyright Amendments, Pub. L. 62-303, 37 Stat. 488.

Ralph S. Brown, *Jr, The Operation of the Damage Provisions of the Copyright Law: An Exploratory Study*, 23 COPYRIGHT LAW REVISION STUDY (1958).

S. REP. No. 100-352 (1988).

Townsend Copyright Amendment: Complete File of Arguments Before the Committee on Patents, House of Representatives, on H.R. 15263 and H.R. 20596, Commencing January 24, 1912, 62d Cong. (1912).

4. Books and Reports

BERNARD MARR, BIG DATA IN PRACTICE: HOW 45 SUCCESSFUL COMPANIES USED BIG DATA ANALYTICS TO DELIVER EXTRAORDINARY RESULTS (2016).

BERTRAND MOULLIER, RIGHTS, CAMERA, ACTION! INTELLECTUAL PROPERTY RIGHTS AND THE FILMMAKING PROCESS 18 (2022).

BUTTERFIELD, A. & GERARD EKEMBE NGONDI, A DICTIONARY OF COMPUTER SCIENCE (7th ed. 2016).

CHANG YIYUN (張懿云), RESEARCH ON THE COPYRIGHT PROTECTION OF AUDIOVISUAL WORKS FINAL REPORT (視聽著作權利保護之研究期末報告) (Taiwan Intell. Prop. Off. 2011).

CHRISTOPHER BOOKER, THE SEVEN BASIC PLOTS: WHY WE TELL STORIES (2004).

DINA APPLETON & DANIEL YANKELEVITS, HOLLYWOOD DEALMAKING: NEGOTIATING TALENT AGREEMENTS FOR FILM, TV, AND NEW MEDIA (2nd ed. 2010).

DOMINIQUE HAUGHTON ET AL., MOVIE ANALYTICS: A HOLLYWOOD INTRODUCTION TO BIG DATA (2015).

ETHEM ALPAYDIN, MACHINE LEARNING (rev. & updated ed. 2021).

EVE LIGHT HONTHANER, THE COMPLETE FILM PRODUCTION HANDBOOK (4th ed. 2010).

FIELD CADY, DATA SCIENCE: THE EXECUTIVE SUMMARY: A TECHNICAL BOOK FOR NON-TECHNICAL PROFESSIONALS (2021).

HORTENSE POWDERMAKER, HOLLYWOOD: THE DREAM FACTORY (1950).

JOHN G. CAWELTI, ADVENTURE, MYSTERY, AND ROMANCE: FORMULA STORIES AS ART AND POPULAR CULTURE (1976).

LAWRENCE LESSIG, FREE CULTURE: THE NATURE AND FUTURE OF CREATIVITY (2005).

LEO A. HANDEL, HOLLYWOOD LOOKS AT ITS AUDIENCE: A REPORT OF FILM AUDIENCE RESEARCH (1950).

LILLIAN PIERSON, DATA SCIENCE FOR DUMMIES (2d ed. 2017).

MARSHALL A. LEAFFER, UNDERSTANDING COPYRIGHT LAW (7th ed. 2019).

MICHAEL C DONALDSON, CLEARANCE & COPYRIGHT: EVERYTHING THE INDEPENDENT FILMMAKER NEEDS TO KNOW (2d ed. 2003).

MICHAEL D. SMITH & RAHUL TELANG, STREAMING, SHARING, STEALING: BIG DATA AND THE FUTURE OF ENTERTAINMENT (2016).

NATIONAL COMMISSION ON NEW TECHNOLOGICAL USES OF COPYRIGHTED WORKS ("CONTU"), FINAL REPORT (1978).

NIKOLAUS REBER, FILM COPYRIGHT, CONTRACTS AND PROFIT PARTICIPATION (2000).

PETER BLOORE, THE SCREENPLAY BUSINESS: MANAGING CREATIVITY AND SCRIPT DEVELOPMENT IN THE FILM INDUSTRY (2012).

PETER DECHERNEY, HOLLYWOOD A VERY SHORT INTRODUCTION (2016).

PETER DECHERNEY, HOLLYWOOD'S COPYRIGHT WARS: FROM EDISON TO THE INTERNET (2012).

STUART J. RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH (4th ed. 2021).

SUSAN OHMER, GEORGE GALLUP IN HOLLYWOOD (2006).

THE MOTION PICTURE ASSOCIATION, INC, THEME REPORT (2019),

THE MOTION PICTURE ASSOCIATION, INC, THEME REPORT (2020).

U.S. COPYRIGHT OFFICE, SIXTY-EIGHTH ANNUAL REPORT OF THE REGISTER OF COPYRIGHTS FOR THE FISCAL YEAR ENDING JUNE 30, 1965 (1966).

U.S. PATENT AND TRADEMARK OFFICE, PUBLIC VIEWS ON ARTIFICIAL INTELLIGENCE AND INTELLECTUAL PROPERTY POLICY (2020).

VIDEO MINING (Azriel Rosenfeld et al. eds., 2003).

VIJAY KOTU & BALA DESHPANDE, DATA SCIENCE: CONCEPTS AND PRACTICE (2d ed. 2019).

WORLD INTELLECTUAL PROPERTY ORGANIZATION, GUIDE TO THE BERNE CONVENTION FOR THE PROTECTION OF LITERARY AND ARTISTIC WORKS (PARIS ACT, 1971) (1978).

WORLD INTELLECTUAL PROPERTY ORGANIZATION, GUIDE TO THE COPYRIGHT AND RELATED RIGHTS TREATIES ADMINISTERED BY WIPO AND GLOSSARY OF COPYRIGHT AND RELATED RIGHTS TERMS 88 (2004).

5. Articles

Atif Khan et al., *Movie Review Summarization Using Supervised Learning and Graph-Based Ranking Algorithm*, 2020 COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE (2020).

Chapter III - Jukebox License, 10 COLUM.-VLA J.L. & ARTS 533 (1985).

D. Daniel Sokol & Roisin Comerford, *Antitrust and Regulating Big Data*, 23 GEO. MASON L. REV. 1129 (2016).

Daniel L. Rubinfeld & Michal S. Gal, *Access Barriers to Big Data*, 59 ARIZ. L. REV. 339 (2017).

Edward Samuels, *The Idea-Expression Dichotomy in Copyright Law*, 56 TENN. L. REV. 321 (1989).

Felix M. Simon & Ralph Schroeder, *Big Data Goes to Hollywood: The Emergence of Big Data as a Tool in the American Film Industry*, SECOND INT'L HANDBOOK OF INT. RSCH. 10 (2019).

Harald Steck et al., *Deep Learning for Recommender Systems: A Netflix Case Study*, 42(3) AI MAG. 7, 9 (2021).

Hsieh Chengkang et al., *Convolutional Collaborative Filter Network for Video Based Recommendation Systems*, ARXIV:1810.08189 (2018).

Jake Linford, *Copyright and Attention Scarcity*, 42 CARDOZO L. REV. 143 (2020).

Jalal Rezaeenour et al., *Systematic review of content analysis algorithms based on deep neural networks*, 82 MULTIMEDIA TOOLS & APPLICATIONS 17879 (2023).

Jane C. Ginsburg, *Copyright and Control Over New Technologies of Dissemination*, 101 COLUM. L. REV. 1613 (2001).

Jane C. Ginsburg, *The Concept of Authorship in Comparative Copyright Law*, 52 DEPAUL L. REV. 1063 (2003).

Jon M. Garon, Content, *Control, and the Socially Networked Film*, 48 U. LOUISVILLE L. REV. 771 (2010).

Jooyoung Lee, et al., *Do Language Models Plagiarize?* WWW '23: PROCEEDINGS OF THE ACM WEB CONF. 3637 (2023).

Kal Raustiala & Christopher Jon Sprigman, *The Second Digital Disruption: Streaming and the Dawn of Data-Driven Creativity*, 94 N.Y.U. L. REV. 1555 (2019).

Leo A. Handel, *Hollywood Market Research*, 7(3) FILM RADIO & TELEVISION Q. 304 (1953).

Lim Daryl, *Life After Google v. Oracle: Three Reflections on a Theme*, 12 IP THEORY 41 (2022).

Luciano Floridi & Massimo Chiriatti, *GPT-3: Its Nature, Scope, Limits, and Consequences*, 30 MINDS & MACHINES 681, 684 (2020).

Marcel Marti et al., *CARDINAL: Computer Assisted Authoring of Movie Scripts*, 23RD INT'L CONF. ON INTELLIGENT USER INTERFACES 509 (2018).

Margaret A. Boden, *Artificial Intelligence and Creativity: A Contradiction in Terms?*, *in* THE PHILOSOPHY OF CREATIVITY: NEW ESSAYS 224 (Elliot Samuel Paul & Scott Barry Kaufman ed., 2014).

Marshall Leaffer, *Protecting Authors' Rights in a Digital Age*, 27 U. TOL. L. REV. 1 (1995).

Nadide Gizem Akgulgil Mutlu, *The future of film-making: Data-driven movie-making techniques*, 10(2) GLOBAL J. OF ARTS EDUC. 167 (2020).

Nantheera Anantrasirichai & David Bull, *Artificial Intelligence in the Creative Industries: A Review*, 55 ARTIF. INTELL. REV. 589 (2022).

Nina Baur et al., *The Quality of Big Data. Development, Problems, and Possibilities of Use of Process-Generated Data in the Digital Age*, 45(3) HIST. SOC. RES. / HIST. SOZIALFORSCHUNG 209 (2020).

OpenAI, *GPT-4 Technical Report*, ARXIV:2303.08774 (2023).

Pansy Nandwani & Rupali Verma, *A review on sentiment analysis and emotion detection from text*, 11(1):81 Soc. Network Analysis & Mining (2020).

Peter Henderson et al., *Foundation models and fair use*, ARXIV:2303.15715 (2023).

Peter Jaszi, Craig Joyce, Marshall Leaffer & Tyler Ochoa, *Statute of Anne: Today and Tomorrow*, 47 HOUS. L. REV. 1013 (2010).

Peter Lee, *Reconceptualizing the Role of Intellectual Property Rights in Shaping Industry Structure*, 72 VAND. L. REV. 1197 (2019).

Piotr Mirowski et al., *Co-Writing Screenplays and Theatre Scripts with Language Models: An Evaluation by Industry Professionals*, ARXIV 2209.14958 (2022).

Ronny Regev, *Hollywood Works: How Creativity Became Labor in the Studio System*, 17(3) ENTER. & SOC'Y 591, 595 (2016).

Rushit Sanghrajka et al., *Computer-Assisted Authoring for Natural Language Story Scripts*, 32 (1) AAAI CONF. ON ARTIFICIAL INTELLIGENCE 7811 (2018).

Shirish Jeble et al., *Role of Big Data in Decision Making*, 11(1) OPERATIONS AND SUPPLY CHAIN MANAGEMENT 36 (2018).

Shruti Saravanan & K. Sudha, *GPT-3 Powered System for Content Generation and Transformation*, 2022 5TH INT'L CONF. ON COMPUTATIONAL INTELLIGENCE AND COMM. TECHS (CCICT) 514 (2022).

Song Minzheong, *A Study on the Predictive Analytics Powered by the Artificial Intelligence in the Movie Industry*, 10(4) Int'l J. of Advanced Smart Convergence 72 (2021).

Stella Biderman et al., *Emergent and Predictable Memorization in Large Language Models*, ARXIV:2304.11158v2 (2023).

Stuart Bender & Billy Sung, *Data-driven creativity for screen production students: developing and testing learning materials involving audience biometrics*, 31(2) DIGITAL CREATIVITY 98 (2020).

Susan Ohmer, *Measuring Desire: George Gallup and Audience Research in Hollywood*, 43(½) Film and Video J. 3 (1991).

Thomas Margoni & Martin Kretschmer, *A Deeper Look into the EU Text and Data Mining Exceptions: Harmonisation, Data Ownership, and the Future of Technology*, 71(8) GRUR Int'l 685 (2022).

Tom B Brown et al., *Language Models are Few-Shot Learners*, ARXIV: 2005.14165 (2020).

Vanessa Bates Ramirez, *OpenAI's GPT-3 Wrote This Short Film—Even the Twist at the End*, SINGULARITY HUB (Oct. 23, 2020), https://singularityhub.com/2020/10/23/an-ai-wrote-this-short-film-and-its-sort-of-fascinating/ [https://perma.cc/5TU2-P977].

Vincent Fortuin et al., *InspireMe: Learning Sequence Models for Stories*, 32 (1) AAAI CONF. ON ARTIFICIAL INTELLIGENCE 7747 (2018).

Wendy J. Gordon, *Authors, Publishers, and Public Goods: Trading Gold for Dross*, 36 Loy. L.A. L. Rev. 159, 165 (2002).

Yao HsinAn (姚信安), *Cong Dianying Zhizuo Shiwu Lun Shiting Zhuzuo Quanli Guishu Yu Xingshi* (從電影製作實務論視聽著作權利歸屬與行使) [Attribution and Exercise of Copyrights of Audiovisual Works From Film Production Practices], *in* GENERAL REVIEW OF REPUBLIC OF CHINA'S COPYRIGHT LAW UNDER INTERNATIONAL COMPARISON I (國際比較下我國著作權法之總檢討上冊) 259 (2014).

Zechariah Chafee Jr., *Reflections on the Law of Copyright: I*, 45 COLUM. L. REV. 503 (1945).

6. The Internet and Other Nonprint Resources

1 Entertainment Industry Contracts P 2.02 (2023) (LEXIS).

1 Entertainment Industry Contracts P 2.04 (2023) (LEXIS).

13 Best Movie data sets for Machine Learning Projects, IMERIT (July 21, 2021), https://imerit.net/blog/13-best-movie-data-sets-for-machine-learning-projects-all-pbm/ [https://perma.cc/7BSB-AHVR].

2 Entertainment Industry Contracts Form 15-16 (2023) (LEXIS).

A HISTORY OF WGA CONTRACT NEGOTIATIONS AND GAINS, https://www.wga.org/the-guild/about-us/history/a-history-of-wga-contract-negotiations-and-gains [https://perma.cc/RN27-D2ES].

ABOUT DRAMATRON, https://deepmind.github.io/dramatron/details.html [https://perma.cc/YLW4-D9AL].

ABOUT US, DISNEYRESEARCH STUDIOS, https://studios.disneyresearch.com/about-us/ [https://perma.cc/5EC5-MKAN].

Adam Epstein, *No, Warner Bros. is not letting AI decide what movies it makes*, QUARTZ (Jan. 9, 2020), https://sg.news.yahoo.com/no-warner-bros-not-letting-152232931.html [https://perma.cc/7DTP-X2PA].

Adam T. Dean, The Paradox of Creativity and Business in Feature Hollywood Filmmaking: The Relationship Between Motion Picture Production and Budgeting, M.A. thesis 32 (August 2005) (University of North Texas, https://digital.library.unt.edu/ark:/67531/metadc4885/: accessed August 15, 2023).

Aiqiyi: 2019 nian Wangluo Dianying Hangye Baogao (爱奇艺：2019年网络电影行业报告), 199IT (Jan. 17, 2020), http://www.199it.com/archives/998277.html [https://perma.cc/7X5Q-W3S7].

Alberto Romero, *A New AI Trend: Chinchilla (70B) Greatly Outperforms GPT-3 (175B) and Gopher (280B)*, TOWARDS DATA SCIENCE (Apr. 11, 2022), https://towardsdata science.com/a-new-ai-trend-chinchilla-70b-greatly-outperforms-gpt-3-175b-and-go pher-280b-408b9b4510 [https://perma.cc/UQ88-BVZU].

Alex Dudok De Wit, *After A Decade, James Cameron Finally Admits That 'Avatar' Is Animated*, CARTOONBREW (Dec. 19, 2019), https://www.cartoonbrew.com/artist-rights/after-a-decade-james-cameron-finally-admits-that-avatar-is-animated-183990. html [https://perma.cc/LLP7-DQKE].

Anastasia Egorova (@anastasiaegorova440), YouTube (Apr. 26, 2017, 7:28 AM), https://www.youtube.com/watch?v=5qPgG98_CQ8&lc=UghZ5ElNx8N6y3gCoAE C [https://perma.cc/L8VP-SW53].

Andres Guadamuz, *Artificial intelligence and copyright*, WIPO MAGAZINE (Oct. 2017), https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html [https://perma. cc/WRW7-LWCT].

Andrew Liptak, *Rogue One's reshoots show how Disney saved the first standalone Star Wars movie*, VERGE (Jan 15, 2017), https://www.theverge.com/2017/1/15/14268 850/star-wars-rogue-one-reshoots-disney-gareth-edwards-tony-gilroy [https://perma. cc/53TP-9L59].

Annalee Newitz, *An AI wrote all of David Hasselhoff's lines in this bizarre short film*, ARSTECHNICA (Apr 25, 2017, 12:00 PM), https://arstechnica.com/gaming/2017/04/ an-ai-wrote-all-of-david-hasselhoffs-lines-in-this-demented-short-film/ [https://per ma.cc/CCZ3-BQKM].

Annalee Newitz, *Movie written by algorithm turns out to be hilarious and intense*, ARSTECHNICA (May 30, 2021, 11:55 AM), https://arstechnica.com/gaming/2021/ 05/an-ai-wrote-this-movie-and-its-strangely-moving/ [https://perma.cc/R9BK-EDE B].

Annie Mueller, *Why Movies Cost So Much to Make*, INVESTOPEDIA (July 31, 2022), https://www.investopedia.com/financial-edge/0611/why-movies-cost-so-much-to-make.aspx#the-bottom-line [https://perma.cc/HZJ4-NGW7].

APACHE HADOOP ON AMAZON EMR, AWS, https://aws.amazon.com/emr/features/hadoop/ [https://perma.cc/65QR-23BV].

Aran Davies, *Analytics and Movies: Predicting Movie Gross*, SOFY (2019), https://sofy.tv/blog/analytics-movies-predicting-movie-gross/ [https://perma.cc/J59Z-V87H].

Aran Davies, *How to use Largo.ai: Part 1 – Introduction*, SOFY (2021), https://sofy.tv/blog/use-largo-ai-part-1-introduction/ [https://perma.cc/6PYR-2XLS].

Aran Davies, *Largo Films Enters Data-Driven Moviemaking*, SOFY (2019), https://sofy.tv/blog/largo-films-enters-data-driven-moviemaking/ [https://perma.cc/L2RH-MZA7].

Ashley Rodriguez, *How Disney Plus, HBO Max, and NBC's Peacock streaming service stack up in the battle against Netflix*, INSIDER: MARKETS INSIDER (Sept. 17, 2019) https://markets.businessinsider.com/news/stocks/disney-plus-hbo-max-nbc-streaming-comparison-price-release-content-2019-7-1028409686 [https://perma.cc/PEM7-RNS7].

Benj Edwards, *Artist receives first known US copyright registration for latent diffusion AI art*, ARSTECHNICA (Sept. 22, 2022), https://arstechnica.com/information-technology/2022/09/artist-receives-first-known-us-copyright-registration-for-generative-ai-art/ [https://perma.cc/SBD3-GZAF].

Big Data and Hollywood: A Love Story, THE ATLANTIC, https://www.theatlantic.com/sponsored/ibm-transformation-of-business/big-data-and-hollywood-a-love-story/277/ [https://perma.cc/7QV5-PNMZ].

Borys Kit & Kim Masters, *Ron Howard Steps into Direct Han Solo Movie*, HOLLYWOOD REPORTER (June 22, 2017), https://www.hollywoodreporter.com/movies/movie-news/star-wars-han-solo-movie-ron-howard-steps-direct-1015674/ [https://perma.cc/9LWA-YSFP].

Brooks Barnes, *Solving Equation of a Hit Film Script, With Data,* N.Y. TIMES, (May 5, 2013) https://www.nytimes.com/2013/05/06/business/media/solving-equation-of-a-hit-film-script-with-data.html [https://perma.cc/HMU4-336B].

Chad Fitzgerald, *Cinema Law: Screenwriter Rights*, MOVIEMAKER (Dec. 14, 2009), https://www.moviemaker.com/cinema-law-screenwriter-rights-20091214/ [https://perma.cc/65DU-2PZY].

Chen Yanru, *Does China Back Copyrights for Automatic Photos from a Hot-Air Balloon?*, CHINA JUST. OBSERVER (Nov. 22, 2020), https://www.chinajusticeobserver.com/a/ does-china-back-copyrights-for-automatic-photos-from-a-hot-air-balloon [https://perma.cc/CD8N-NEFU].

Chiteki Zaisan Senryaku Honbu Kenshō Hyōka Kikaku Īnkai Aratana Jōhōzai Kentō Īnkai (知的財産戦略本部検証・評価・企画委員会 新たな情報財検討委員会) [Intellectual Property Strategy Headquarters Verification/Evaluation/Planning Committee New Information Property Review Committee], AI Seiseibutsu To Chosakuken Nitsuite (AI 生成物と著作権について) [About AI products and copyright] 2 (Mar 2023), https://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kousou/ 2023/dai2/siryou1.pdf [https://perma.cc/N8GD-MQN5].

Chris Heckmann, *What is New Hollywood? The Revolution of 1960s and '70s Hollywood*, STUDIO BINDER (May 17, 2020), https://www.studiobinder.com/blog/what-is-new-hollywood/ [https://perma.cc/FY47-5MXG].

Chris O'Brien, *How Cinelytic is using AI to help Hollywood reboot for the streaming wars*, VENTUREBEAT (May 26, 2020), https://venturebeat.com/business/how-cinelytic-is-using-ai-to-help-hollywood-reboot-for-the-streaming-wars/ [https://perma.cc/V8R2 -P2V3].

Christopher Schiller, *What is Film Packaging*? SCRIPT (Feb. 9, 2018), https://scriptmag. com/features/it-depends-what-is-film-packaging [https://perma.cc/VL3X-B7QK].

Data-Asisted Moviemaking with LargoAI, Largo, http://largofilms.ch/data-driven-movie making/ [https://perma.cc/38DK-EXHW].

Data-driven, Wikipedia, https://en.wikipedia.org/w/index.php?title=Data-driven&oldid= 1120149065 [https://perma.cc/6RLH-66WA].

Dataproc, Google Cloud, https://cloud.google.com/dataproc [https://perma.cc/4B8A-UBL X].

Dave Simpson, *GitHub Users Say Microsoft Ripped off Their Code for AI Tool*, Law360 (November 3, 2022, 9:15 PM EDT).

Debadrita Sur, *Netflix releases its first-ever bot-written horror film*, Best of Netflix (Oct. 8, 2021), https://best-of-netflix.com/netflix-releases-its-first-ever-bot-written-horror -film/ [https://perma.cc/RF9B-BC37].

Demi Ajayi, *How BERT and GPT models change the game for NLP*, IBM Blog (Dec. 3, 2020), https://www.ibm.com/blogs/watson/2020/12/how-bert-and-gpt-models-chan ge-the-game-for-nlp/ [https://perma.cc/HUV3-T8RL].

Digital Transformation Research Institute of Institute for Information Industry (資策會數 位轉型研究院), *AI, Dashuju, Juben: Kuaizhi Renkou de Women yu E de Juli Juben Shi Zenme Chansheng de* (AI、大數據、劇本：膾炙人口的「我們與惡的距離」 劇本是怎麼產生的？) DTRI (Apr. 14, 2020) https://ideas-dtri.iii.org. tw/%E6%9C%8D%E5%89%B5%E6%89%80%E7%A0%94%E7%99%BC%E6% 87%89%E7%94%A8/ai-writed-tv-series-script/ [https://perma.cc/W99V-SDYC]

Dillon Berjani, *How Netflix is making use of the "Long tail" theory to attract new subscribers*, Medium (Jan. 8, 2018) https://medium.com/@dillon.berjani/how-net flix-is-making-use-of-the-long-tail-theory-to-attract-new-subscribers-e4a96923ab6 d [https://perma.cc/6J6Q-8L6Q].

Dylan Brethour, *Sunspring Review*, Film & Television (June 23, 2016), https://www.cur atormagazine.com/dylan-brethour/sunspring-review/ [https://perma.cc/8CPN-3THC].

Early Motion Picture Productions, Library of Congress, https://www.loc.gov/ collections/edison-company-motion-pictures-and-sound-recordings/articles-and-ess ays/history-of-edison-motion-pictures/early-motion-picture-productions/ [https://pe rma.cc/3QL8-CH7X].

Eda Kavlakoglu, *NLP vs. NLU vs. NLG: the differences between three natural language processing concepts*, IBM Blog (Nov. 12, 2020), https://www.ibm.com/blogs/watson/2020/11/nlp-vs-nlu-vs-nlg-the-differences-between-three-natural-language-processing-concepts/ [https://perma.cc/K8P9-JK42].

Ehtsham Elahi et al., *Reinforcement Learning for Budget Constrained Recommendations*, Netflix TechBlog (Aug. 15, 2022), https://netflixtechblog.com/reinforcement-learning-for-budget-constrained-recommendations-6cbc5263a32a [https://perma.cc/GXU2-CC4Y].

Eileen Morley & Andrew Silver, *A Film Director's Approach to Managing Creativity*, Harv. Bus. Rev. (Mar. 1977), https://hbr.org/1977/03/a-film-directors-approach-to-managing-creativity [https://perma.cc/9KGY-SSEL].

Elizabeth Mixson, *Data Science at Netflix: How Advanced Data & Analytics Helps Netflix Generate Billions*, AI, Data & Analytics Network (Mar. 30, 2021), https://www.aidataanalytics.network/data-science-ai/articles/data-science-at-netflix-how-advanced-data-analytics-helped-netflix-generate-billions#:~:text=Using [https://perma.cc/BC9H-8WL6].

Elle Love, *Why can't Hollywood come up with new ideas for movies?*, Gateway (Nov. 1, 2019), http://unothegateway.com/why-cant-hollywood-come-up-with-new-ideas-for-movies/ [https://perma.cc/H55N-4HHH].

Eric Avidon, *Data drives the decisions behind making movies*, TechTarget (June 12, 2020), https://searchbusinessanalytics.techtarget.com/feature/Data-drives-the-decisions-behind-making-movies [https://perma.cc/U2R7-GUHP].

Erik Gruenwedel, *JustWatch: Prime Video Edges Netflix in Q2 U.S. Market Share*, MediaPlayNews (July 6, 2023), https://www.mediaplaynews.com/justwatch-prime-video-edges-netflix-in-q2-u-s-market-share/ [https://perma.cc/GWB4-LG9R].

Erin Pearson, *The Basics to Making a Low Budget Film*, Topsheet (Sept. 27, 2019, https://topsheet.io/blog/basics-to-making-a-low-budget-film [https://perma.cc/FA9F-ES9C].

284

Everett C. Ladd, *The Trials of Election Polling: Election Poll—1948 and Today*, https://www.trumanlibrary.gov/sites/default/files/1948Campaign_Trials.pdf [https://perma.cc/4EJT-VZGQ].

Fotis Georgiadis, *The Future Is Now: Sami Arp of Largo AI on How Their Technological Innovation Will Shake up the Tech Scene*, MEDUIUM (Apr. 27, 2022), https://medium.com/authority-magazine/the-future-is-now-sami-arp-of-largo-ai-on-how-their-technological-innovation-will-shake-up-the-c5fa3cd3a755 [https://perma.cc/DQ2E-G8CQ].

Frequently Asked Questions, Cinelytic, https://www.cinelytic.com/faq/[https://perma.cc/9Z5V-M6SC].

Govind Dheda, *Is Chat GPT Plagiarism? Unveiling the Truth Behind Generated Content*, OPENAIMASTER (May 24, 2023), https://openaimaster.com/is-chat-gpt-plagiarism/ [https://perma.cc/F63J-HAL3].

GPT-4, OPENAI (March 14, 2023), https://openai.com/research/gpt-4 [https://perma.cc/H4HD-GA96].

Hannah Shaw-Williams, *Black Mirror: Bandersnatch Choices Guide - Best Decisions & All Outcomes Explained*, SCREEN RANT (Dec. 28, 2018), https://screenrant.com/black-mirror-bandersnatch-choices-guide-best-decisions/ [https://perma.cc/C2WY-GT4M].

HEAVEN'S GATE, IMDB, https://www.imdb.com/title/tt0080855/ [https://perma.cc/83CC-HRZE].

Hilary Swett, *The Screen Writers' Guild: An Early History of the Writers Guild of America*, WRITERS GUILD FOUNDATION (2020), https://www.wgfoundation.org/screenwritersguild-history [https://perma.cc/S3GZ-7VP5].

Hillel Aron, *Microsoft and GitHub ask court to scrap lawsuit over AI-powered CoPilot*, COURTHOUSE NEWS (May 4, 2023), https://www.courthousenews.com/microsoft-and-github-ask-court-to-scrap-lawsuit-over-ai-powered-copilot/[https://perma.cc/E5ML-V98Y].

285

How Artificial Intelligence Can Help To Pitch A Movie, Largo, http://largofilms.ch/how-artificial-intelligence-can-help-to-pitch-a-movie/ [https://perma.cc/2TUQ-KU5K].

Huang, Chieh, "Where is the Author: the Copyright Protection for AI-Generated Works" (2022). *Maurer Theses and Dissertations*. 107. https://www.repository.law.indiana.edu/etd/107.

Huang, Chieh, WHERE IS THE AUTHOR: THE COPYRIGHT PROTECTION FOR AI-GENERATED WORKS, 107 MAURER THESES AND DISSERTATIONS (2022), https://www.repository.law.indiana.edu/etd/107.

Institute for Information Industry(資策會), *Buzhi Xie Juben Zicehui Dashuju Fenxijishu Yanjiu Shehui Baitai* (不只寫劇本 資策會大數據分析技術研究社會百態),WWW.III.ORG.TW (Jan. 29, 2021), https://www.iii.org.tw/Press/NewsDtl.aspx?nsp_sqno=2341&fm_sqno=14 [https://perma.cc/5QJ3-X5YX].

Interview with Nadira Azermai, CEO of ScriptBook, *Writing new chapters in the film business: Storytelling meets AI*, EUROPEAN BUSINESS, https://www.european-business.com/interviews/scriptbook/writing-new-chapters-in-the-film-business-storytelling-meets-ai [https://perma.cc/QBV3-9TU8].

INTRODUCING CHATGPT, OPENAI (NOV. 30, 2022), HTTPS://OPENAI.COM/BLOG/CHATGPT/ [https://perma.cc/8XAZ-MYMP].

James Somers, *The Man Who Would Teach Machines to Think*, THE ATLANTIC (Nov. 2013) https://www.theatlantic.com/magazine/archive/2013/11/the-man-who-would-teach-machines-to-think/309529/?fbclid=IwAR0Gvjha5PxS5p47x0pbIEp31KhdS2wmVfCkPpWTuG50wSa84ok96CyizB4 [https://perma.cc/PPQ4-TJXS].

James Vincent, *Hollywood is quietly using AI to help decide which movies to make*, VERGE (May 28, 2019, 11:10 AM EDT), https://www.theverge.com/2019/5/28/18637135/hollywood-ai-film-decision-script-analysis-data-machine-learning [https://perma.cc/2LLK-4WBF].

Jan Hendrik Kirchner et.al., *New AI classifier for indicating AI-written text*, OPENAI (Jan. 31, 2023), https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text [https://perma.cc/JR82-BBRZ].

Jason P. Frank, *The 2023 WGA Strike for Dummies*, Vulture (updated June 30, 2023), https://www.vulture.com/article/wga-strike-2023.html [https://perma.cc/BJ75-SLM 6].

Jennifer Maas, *Netflix's Ted Sarandos Is Feeling 'Better and Better' About $17 Billion Content Budget: 'We're Spending at About the Right Level'*, Variety (Oct. 18, 2022, 3:53 PM PT), https://variety.com/2022/tv/news/netflix-content-spend-17-billion-subscriber-growth-1235407818/ [https://perma.cc/3XTZ-GNU2].

John Markoff, *Computer Wins on 'Jeopardy!': Trivial, It's Not*, N.Y. Times (Feb. 16, 2011), https://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html.

John R. Smith, *IBM Research Takes Watson to Hollywood with the First "Cognitive Movie Trailer,"* IBM Think Blog (Aug. 31, 2016), https://www.ibm.com/blogs/think/20 16/08/cognitive-movie-trailer/ [https://perma.cc/RX6A-ZCRE].

JP, *Data-Driven Film Making – Cinema's Future?*, Harvard Business School: Digital Initiative (Nov. 22, 2015) https://digital.hbs.edu/platform-digit/submission/data-driven-film-making-cinemas-future/ [https://perma.cc/JY4F-7BHD].

Julia Mas-Guindal, *The Motion Picture Industry: Critical Issues Concerning Moral Rights and Authorship*, JoltDigest (Dec.31, 2011), https://jolt.law.harvard.edu/digest/the-motion-picture-industry-critical-issues-concerning-moral-rights-and-authorship [htt ps://perma.cc/3MX8-DC7T].

Julia Stoll, *Most expensive Netflix original series 2022, by production cost per episode*, Statista (Oct 19, 2022), https://www.statista.com/statistics/1249573/most-expen sive-netflix-original-series-production-cost-per-episode/ [https://perma.cc/DRX8-V GEM].

Kaitlin Thomas, *Everything We Know About Apple's Streaming Service, Apple TV+*, TVGuide (Mar. 30, 2020) https://www.tvguide.com/news/apple-tv-plus-streaming-service/ [https://perma.cc/2ZWS-EVWF].

Kasey Moore, *Netflix Originals Now Make up 50% of Overall US Library*, What's on Netflix (Aug. 24, 2022, 9:45 AM EST), https://www.whats-on-netflix.com/news/

50-of-netflixs-library-is-now-made-of-netflix-originals/ [https://perma.cc/SNN5-XP 4F].

Kris Holt, *Deepmind Created An AI Tool That Can Help Generate Rough Film and Stage Scripts*, ENGADGET (Dec. 9, 2022 4:23 PM), https://www.engadget.com/deepmind-ai-tool-film-theater-scripts-dramatron-212328597.html [https://perma.cc/GV7G-V6 VE].

Kwame Opam, *Does Disney want its directors to have creative freedom?*, VERGE (June 23, 2017), https://www.theverge.com/2017/6/23/15861162/disney-marvel-star-wars-lucasfilm-han-solo-director-creative-freedom [https://perma.cc/HVG9-AFTQ].

Lauren Berg, *ChatGPT Is 'Trained' Using Copyrighted Books, Authors Say*, LAW360 (June 29, 2023, 9:56 PM EDT), https://www.law360.com/articles/1694688/chatgpt-is-trained-using-copyrighted-books-authors-say[https:[//perma.cc/QG8A-EH9Q].

Lauren Goode, *AI Made a Movie—and the Results Are Horrifyingly Encouraging*, WIRED (June 11, 2018, 3:00 PM), https://www.wired.com/story/ai-filmmaker-zone-out/ (last visited Aug. 17, 2023).

Lexus Europe, *Lexus ES | Driven by Intuition | A film made by Artificial Intelligence | Lexus Europe*, YOUTUBE (Nov. 19, 2018), https://www.youtube.com/watch?v=-iaBJ5rqOdg&t=10s.

Li Xiazhi (李夏至), *Wuda Zhixian Shiqizhong Jieju Hudongju Guanzhong Maizhangma?* (五大支线十七种结局，互动剧观众买账吗？) BEIJING DAILY (Li Xiazhi (李夏至), *Wuda Zhixian Shiqizhong Jieju Hudongju Guanzhong Maizhangma?* (五大支线十七种结局，互动剧观众买账吗？) BEIJING DAILY (北京日报) (June 25, 2019), http://m.xinhuanet.com/ent/2019-06/25/c_1124665831.htm [https://perma.cc/5PV8-YJ5S].)

Lian Yiting(連以婷), Womenyue Weihe You Suanmin Meiti Juese Muhou Zicehui Dashuju Tuandui de Gong Zuo Jiemi (《我們與惡》為何有酸民、媒體角色？幕後資策會大數據團隊的工作揭密！), TECHNEWS (June 24, 2019, 2:00 PM), https://technews.tw/2019/06/24/the-world-between-us/ [https://perma.cc/4D5Q-AG72].

Lillian Rizzo, *NBCUniversal CEO Jeff Shell says movie business benefiting from new release model*, CNBC (Oct. 4, 2022, 11:56 AM), https://www.cnbc.com/2022/10/04/nbcuniversal-ceo-jeff-shell-says-movie-business-is-better-with-new-release-model.html [https://perma.cc/P6WP-4S6Z].

Lin Ruiwan (林瑞婉), *Women yu Dashuju de Juli* (我們與大數據的距離), PTS ACTION Q. (May 31, 2018), https://medium.com/@PTS_quarterly/%E6%88%91%E5%80%91%E8%88%87%E5%A4%A7%E6%95%B8%E6%93%9A%E7%9A%84%E8%B7%9D%E9%9B%A2-69ecfce2f2dc [https://perma.cc/U7LB-UCMQ].

M. Tim Jones, *Unsupervised learning for data classification*, IBM (Dec. 4, 2017). https://developer.ibm.com/articles/cc-unsupervised-learning-data-classification/#:~:text=In%20unsupervised%20learning%2C%20an%20algorithm,for%20tasks20like%20anomaly%20detection [https://perma.cc/WS68-VU2M].

Madeline Riley et al., *AVA: The Art and Science of Image Discovery at Netflix*, NETFLIX TECH. BLOG (Feb. 7, 2018), https://netflixtechblog.com/ava-the-art-and-science-of-image-discovery-at-netflix-a442f163af6 [https://perma.cc/5FM6-QDN5].

Manage your big data needs in an open-source platform, AZURE, https://azure.microsoft.com/en-us/products/hdinsight/#overview [https://perma.cc/RG4R-HYCG].

Manori Ravindran, *Netflix: Humans, not machines, greenlight our shows*, TBI (Mar. 11, 2019), https://tbivision.com/2019/03/11/netflix-humans-not-machines-greenlight-our-shows/ [https://perma.cc/5SSE-BQPW].

Mansoor Iqbal, *Netflix Revenue and Usage Statistics (2023)*, BUSINESSOFAPPS (May 24, 2023), https://www.businessofapps.com/data/netflix-statistics/ [https://perma.cc/A2ZS-DGLY].

Mark Brown, *Scientists uncover formula for box office movie success*, GUARDIAN (July 24 2018), https://www.theguardian.com/film/2018/jul/24/uk-scientists-uncover-formula-box-office-movie-success [https://perma.cc/XK9G-EQXG].

Marlon Mosley, *Heaven's Gate: UA and Michael Cimino's demise*, MWMBLOG (Oct. 26, 2019), https://mwmblog.com/2019/10/26/heavens-gate-ua-and-michael-ciminos-demise/ [https://perma.cc/85PR-J8M9].

Matthias Bastian, *Deepmind's "Dramatron" Can Write Film and Theater Scripts*, DECODER (Dec. 10, 2022), https://the-decoder.com/deepminds-dramatron-can-write-film-and-theater-scripts/ [https://perma.cc/73QV-BQTF].

Maximilian Schreiner, *Deepmind: Artificial Intelligence is far from being fed up*, DECODER (Apr. 23, 2022), https://the-decoder.com/deepmind-artificial-intelligence-is-far-from-being-fed-up/ [https://perma.cc/HP7J-Q8FQ].

Mehul Gupta, *Tokenization algorithms in Natural Language Processing (NLP)*, MEDUIUM (Mar. 31, 2020), https://medium.com/data-science-in-your-pocket/tokenization-algorithms-in-natural-language-processing-nlp-1fceab8454af [https://perma.cc/PZ4M-KX2N].

Melody Dye et al., *Supporting content decision makers with machine learning*, NETFLIX TECH. BLOG (Dec. 10, 2020), https://netflixtechblog.com/supporting-content-decision-makers-with-machine-learning-995b7b76006f [https://perma.cc/N6BB-DQ7T].

Michael Krigsman, *"Moneyball" for Movies: Data and Analytics at Legendary Entertainment*, CKOTALK (Feb. 09, 2018), https://www.cxotalk.com/episode/moneyball-movies-data-analytics-legendary-entertainment [https://perma.cc/B9T5-6EK9].

Michael Mrbullwinkle & Eric Urban, *Learn how to work with the ChatGPT and GPT-4 models*, AZURE (May 05, 2023), https://learn.microsoft.com/en-us/azure/cognitive-services/openai/how-to/chatgpt?pivots=programming-language-chat-completions [https://perma.cc/U4G5-8W25].

Miguel Campo-Rembado & Sona Oakley, *How 20th Century Fox uses ML to predict a movie audience*, GOOGLE CLOUD (Oct. 29, 2018), https://cloud.google.com/blog/products/ai-machine-learning/how-20th-century-fox-uses-ml-to-predict-a-movie-audience [https://perma.cc/HG8J-YJR7].

MOVIE BUDGET AND FINANCIAL PERFORMANCE RECORDS, THE NUMBERS, https://www.the-numbers.com/movie/budgets [https://perma.cc/ETL6-KJ4G].

Netflix de Hudongshi Jiemu yu Dianying (Netflix 的互動式節目與電影), NETFLIX, https://help.netflix.com/zh-tw/node/62526 [https://perma.cc/Q2R7-D83T].

Netflix Is A Joke, *The First Horror Movie Written Entirely By Bots*, YOUTUBE (OCT. 6, 2021), https://www.youtube.com/watch?v=WZzbxNoMjGM.

Nicholas Barber, *Heaven's Gate: From Hollywood disaster to masterpiece*, BBC (Dec. 4, 2015), https://www.bbc.com/culture/article/20151120-heavens-gate-from-hollywoo d-disaster-to-masterpiece [https://perma.cc/33WF-T6Q4].

Nickalus Rupert, *Jurassic Park: How the Dinosaurs were Created (Is It All CGI?)*, SCREENRANT (Dec. 07, 2020), https://screenrant.com/jurassic-park-movie-dinosaurs -create-cgi-effects-explained/ [https://perma.cc/KDY6-5UY9].

Origins of Motion Pictures, LIBRARY OF CONGRESS, https://www.loc.gov/collections/ edison-company-motion-pictures-and-sound-recordings/articles-and-essays/history- of-edison-motion-pictures/origins-of-motion-pictures/ [https://perma.cc/PD7M-272 P].

OXFORD DICTIONARIES. OXFORD UNIVERSITY PRESS, https://premium.oxforddictionaries. com/us/.

Paul Duncan, FACEBOOK, (Apr. 9, 2021, 12:24 PM), https://www.facebook.com/kers hed/photos/a.2625273730832567/4539457882747466/?locale=hi_IN (last visited Aug. 15, 2023).

Peter Caranicas, *Artificial Intelligence Could One Day Determine Which Films Get Made*, VARIETY (JULY 5, 2018), https://variety.com/2018/artisans/news/artificial-intelligen ce-hollywood-1202865540/ [https://perma.cc/C7YG-3WNL].

Peter Suderman, *How Marvel built such an impressive movie universe*, VOX (MAY 9, 2016), https://www.vox.com/2016/5/9/11595344/marvel-cinematic-universe-captain-ameri ca-avengers [https://perma.cc/K5FG-THHC].

Peter White, *Netflix Comedy Chief Tracey Pakosta on Netflix Is A Joke Festival, Multicamera Conundrum & Laughter Pipeline*, DEADLINE (Apr. 28, 2022, 1:28 PM), https://deadline.com/2022/04/netflix-comedy-tracey-pakosta-netflix-is-a-joke-1235 010699/ [https://perma.cc/9GFX-SGCH].

Press Release, Hulu, Hulu acquires the Video Genome Project Technology (Nov. 15, 2016), https://press.hulu.com/hulu-acquires-the-video-genome-project-technology/ [https://perma.cc/Y4LE-YQ6C].

Press Release, Joint Statement from WGA & DGA Presidents Regarding Max Credit Changes (May 24, 2023), https://www.wga.org/news-events/news/press/joint-statement-from-wga-and-dga-presidents-regarding-max-credit-changes [https://perma.cc/7L4E-N3WQ].

PROMPTS, MIDJOURNEY, https://docs.midjourney.com/docs/prompts[https://perma.cc/4J29-YUJZ].

Raf, *What are tokens and how to count them?*, OPENAI, https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them[https://perma.cc/3X6Z-TEWT].

Reece Medway, *Lexus Europe Creates World's Most Intuitive Car Ad with IBM Watson*, IBM THINK BLOG (Nov. 19, 2018), https://www.ibm.com/blogs/think/2018/11/lexus-europe-creates-worlds-most-intuitive-car-ad-with-ibm-watson/ [https://perma.cc/2TJH-PSMU].

Renee Laroche-Rheaume, *Goodbye, Green Screen — Here's the Tech That's Changing Film Forever!* GOODNESS-EXCHANGE (Nov. 20, 2020), https://goodness-exchange.com/the-volume-mandalorian-storytelling/ [https://perma.cc/4G3J-LNT3].

Rethinking the Long Tail Theory: How to Define 'Hits' and 'Niches', KNOWLEDGE AT WHARTON (Sept. 16, 2009), https://knowledge.wharton.upenn.edu/article/rethinkg-the-long-tail-theory-how-to-define-hits-and-niches/ [https://perma.cc/WL2P-KD9U].

Rob Thomas, *IBM Watson: Reflections and Projections*, IBM THINK BLOG (Oct. 10, 2019), https://www.ibm.com/blogs/think/2019/10/what-is-watson/ [https://perma.cc/UN3P-XGN2].

Robert W. Welkos, *From the Archives: 'Heaven's Gate': The film flop that reshaped Hollywood*, L.A. TIMES (June 12, 2004), https://www.latimes.com/entertainment/movies/la-et-mn-heavens-gate-flop-archive-20040612-snap-story.html [https://perma.cc/F279-CQ6G].

Roger Ebert, *Cameron retains his crown*, ROGEREBERT.COM (Dec. 11, 2009), https://www.rogerebert.com/reviews/avatar-2009 [https://perma.cc/CPJ9-TAM4].

Rudie Obias, *11 Movies That Changed Because of Test Audiences*, MENTAL FLOSS (Apr. 4, 2016), https://www.mentalfloss.com/article/64467/11-movies-changed-because-test-audiences [https://perma.cc/PB8J-T9TE].

Sam Machkovech, *This wild, AI-generated film is the next step in "whole-movie puppetry,"* ARSTECHNICA (JUNE 11, 2018, 12:01 PM), https://arstechnica.com/gaming/2018/06/this-wild-ai-generated-film-is-the-next-step-in-whole-movie-puppetry/ [https://perma.cc/E39B-UA9C].

Seiseikei AI No Katsuyō, Chosakukensha O Mamoru Tame No Aratana Kisei Ga Hitsuyōda (生成系 AI の活用、著作権者を守るための新たな規制が必要だ) [Utilization of generative AI, new regulations are necessary to protect copyright holders], Shūgīn Gīn Kīta Kashi (衆議院議員きいたかし) [House Representative Takashi Kii] (Apr. 24, 2023), https://kiitaka.net/21312/ [https://perma.cc/Y94X-DRQQ].

Shane Atchison & Jason Burby, *Big data and creativity: What we can learn from 'House of Cards'*, THE NEXT WEB (Mar. 20, 2016) https://thenextweb.com/insider/2016/03/20/data-inspires-creativity/ [https://perma.cc/8HTV-X9Y5].

Shift to Projectors and the Vitoscope, LIBRARY OF CONGRESS, https://www.loc.gov/collections/edison-company-motion-pictures-and-sound-recordings/articles-and-essays/history-of-edison-motion-pictures/shift-to-projectors-and-the-vitoscope/ [https://perma.cc/QD3F-U6UJ].

Spam filter, *This is what happens when an AI-written screenplay is made into a film*, GUARDIAN (June 10, 2016), https://www.theguardian.com/technology/2016/jun/10/artificial-intelligence-screenplay-sunspring-silicon-valley-thomas-middleditch-ai?CMP=share_btn_tw [https://perma.cc/XA5C-FN5J].

Stephanie Prange, *Big Data Revolution*, MEDIAPLAYNEWS (Sept. 3, 2018), https://www.mediaplaynews.com/big-data-revolution/ [https://perma.cc/ELY9-TJVX].

Stephen Follows, *How many screenwriters does it take to write a movie?* STEPHEN FOLLOWS (Feb. 24, 2020), https://stephenfollows.com/how-many-screenwriters-does-it-take-to-write-a-movie/ [https://perma.cc/8GW6-DHY4].

Stephen Follows, *The prevalence of sequels, remakes and original movies*, STEPHEN FOLLOWS (Apr. 30, 2018), https://stephenfollows.com/the-prevalence-of-sequels-remakes-and-original-movies/ [https://perma.cc/ED9F-RRNA]

Stephen Galloway, *The New Power of "Yes": Who Actually Has Greenlight Authority at the Movie Studios*, HOLLYWOOD REPORTER (Feb. 15, 2017), https://www.hollywood reporter.com/news/general-news/new-power-yes-who-actually-has-greenlight-auth ority-at-movie-studios-976236/ [https://perma.cc/C4K9-BHZT].

Sven Mikulec, *'Jaws': The Groundbreaking Summer Blockbuster that Changed Hollywood, and Our Summer Vacations, Forever*, CINEPHILIA & BEYOND, https:// cinephiliabeyond.org/jaws-groundbreaking-summer-blockbuster-changed-ho llywood-summer-vacations-forever/ [https://perma.cc/M93H-H8FW].

Tatiana Siegel, *Warner Bros. Signs Deal for AI-Driven Film Management System*, HOLLYWOOD REPORTER (Jan. 8, 2020, 12:23 PM), https://www.hollywoodreporter. com/business/business-news/warner-bros-signs-deal-ai-driven-film-management-sy stem-1268036/ [https://perma.cc/N99Z-QH7U].

The IBM Punched Card, IBM100, https://www.ibm.com/ibm/history/ibm100/us/en/icons/ punchcard/ [https://perma.cc/XX8S-9WST].

The Video Genome Project, https://thevideogenomeproject.com/ [https://perma.cc/M9AP-BKE4].

Thomas Wick, *Demand More Original Films, Not Remakes, From Hollywood*, THE PITT NEWS (Mar. 19, 2017) https://pittnews.com/article/117993/opinions/demand-origi nal-films-not-remakes-hollywood/ [https://perma.cc/XR8F-6VVB].

Tiffany Hsu & Steven Lee Myers, *Another Side of the A.I. Boom: Detecting What A.I. Makes*, N.Y. TIMES (May 18, 2023), https://www.nytimes.com/2023/05/18/techn ology/ai-chat-gpt-detection-tools.html?searchResultPosition=1 [https://perma.cc/45 BX-7GW5].

Tim Wu, *Netflix's Secret Special Algorithm Is a Human*, NEW YORKER (Jan. 27, 2015) https://www.newyorker.com/business/currency/hollywoods-big-data-big-deal [https ://perma.cc/DR32-E5JH].

Vanessa Bates Ramirez, *OpenAI's New Text Generator Writes Even More Like a Human*, SINGULARITY HUB (June 18, 2020), https://singularityhub.com/2020/06/18/openais-new-text-generator-writes-even-more-like-a-human/ [https://perma.cc/8V2L-8K95].

WebGPT: Improving the factual accuracy of language models through web browsing, OpenAI (Dec.16, 2021), https://openai.com/research/webgpt [https://perma.cc/3X2 E-9BH3].

WGA Negotiations—Status as of May 1, 2023, WGA ON STRIKE, https://www.wgacon tract2023.org/the-campaign/wga-negotiations-status-as-of-5-1-2023[https://perma.c c/LLV7-8JHT].

WHAT IS ARTIFICIAL INTELLIGENCE (AI)?, IBM, https://www.ibm.com/cloud/learn/what-is-artificial-intelligence [https://perma.cc/427Y-QK3H].

Writers Guild of America West (@WGAWest), Tweeter (Mar 22, 2023, 8:47 PM), https://twitter.com/WGAWest/status/1638643544977195008?ref_src=twsrc%5Etf w%7Ctwcamp%5Etweetembed%7Ctwterm%5E1638643544977195008%7Ctwgr% 5Eac8451a069a50959ba745ad0e69e4c0a0845201f%7Ctwcon%5Es1_&ref_url=htt ps%3A%2F%2Fvariety.com%2F2023%2Fbiz%2Fnews%2Fwriters-guild-artificial-intelligence-proposal-1235560927%2F [https://perma.cc/A5WU-NUTQ].

Writers Guild of America, *Why We Strike*, https://www.wgacontract2023.org/member-voices/why-we-strike [https://perma.cc/ZUG3-QD87].

Writers Guild of American West, *How the Major Hollywood Talent Agencies Put Their Interests Ahead of Their Clients' Interests*, https://www.wga.org/uploadedfiles/ members/member_info/agency_agreement/wga_no_conflict_no_interest_19.pdf [ht tps://perma.cc/FK6D-MBB2].

Writers Guild of American West, *Members Vote Yes on Screen Credits Rule Change*, WGAW (Nov. 15, 2021), https://www.wga.org/members/membership-information/ members-vote-yes-on-screen-credits-rule-change [https://perma.cc/G96Z-38TC].

Writers Guild of American West, *Screen Credits Manual*, https://www.wga.org/contracts/credits/manuals/screen-credits-manual.

Writers Guild of American West, *The Truth About Agency Studios*, Youtube (Nov. 11, 2019), https://www.youtube.com/watch?v=EaXQ84Hn6_Y&t=4s.

Writers Guild of American, *Screen Credits Referendum 2021*, https://www.wga.org/uploadedfiles/the-guild/elections/screen_credits_explainer.pdf [https://perma.cc/6ZNR-LQKZ]

Yang Wenjun (楊文君), *Dazhaishidai Bei Ping Xieshi dao Jiuxin Dashuju Jiachi de Juben Weihe Neng Ruci Tieqie* (大債時代被評寫實到揪心 大數據加持的劇本為何能如此貼切？) Radio Taiwan Int'l (Feb. 15, 2021, 9:35 AM), https://www.rti.org.tw/news/view/id/2091158 [https://perma.cc/CYJ3-C7D4].

Ye Guanyin (葉冠吟), *Dashuju Zhugong Taiju Chuangzuo Xiankai Xinxie* (大數據助攻台劇創作掀開新頁), Cent. News Agency (Taiwan) (Mar. 14, 2021, 12:20 PM), https://www.cna.com.tw/news/firstnews/202103145002.aspx [https://perma.cc/W6Q6-M5U3].

Yuan Bo (袁博), *Cong Jinxiu Weiyang Shexian Chaoxi Kan Zhuzuoquan Qinquan Xin Leixing* (从《锦绣未央》涉嫌抄袭看著作权侵权新类型) [From the Suspected Plagiarism of "Jinxiu Weiyang" See New Type of Copyright Infringement], China Press Publ'n Radio Film &Television J (中国新闻出版广电报) (Feb. 23, 2017, 1:45 PM), http://www.ezhicai.com/news/66347.htm [https://perma.cc/D4ZE-NZSW].

Yuri Baranovsky, *Six Tips for a Successful Test Screening*, Frame.io (July 6, 2017), https://blog.frame.io/2017/07/06/successful-test-screening-tips/ [https://perma.cc/E742-GVXX].

Zack Sharf, *'Black Mirror: Bandersnatch' Contains 5 Endings and 1 Trillion Story Combos, Changes the Way Netflix Streams*, IndieWire (Dec. 28, 2018), https://www.indiewire.com/2018/12/black-mirror-bandersnatch-endings-one-trillion-story-combinations-netflix-streaming-1202031075/ [https://perma.cc/K4TY-JSZJ].

Zaynab Zlitni, *Is Hollywood running out of ideas?* UNIV. DAILY KANSAN (Feb. 27, 2023), https://www.kansan.com/opinion/is-hollywood-running-out-of-ideas/article_39977 4fe-b6ed-11ed-be6a-bfaab38dc00a.html [https://perma.cc/R4J6-D67L].

Zhou Bo (周波), *Rengong Zhineng Yu Zhuzuoquan Baohu — Zhongguo Fayuan De Sifa Shijian* (人工智能与著作权保护 —中国法院的司法实践) [Artificial Intelligence and Copyright Protection —Chinese Courts' Judicial Practice], https://www.wipo. int/export/sites/www/about-ip/en/artificial_intelligence/conversation_ip_ai/pdf/ms_ china_1_zh.pdf [https://perma.cc/42R2-YVRG].