

Exploring Marvel's IP Marketing Strategies in the Digital Era: A Comparative Analysis of Social Media Promotion and Audience Engagement

Yibo Huang^{1,*†}, Xuhui Rong^{2,†}, Yufan Du³, Yujia Mu⁴

¹*Department of the University of Waikato Joint Institute at Hangzhou City University, Hangzhou City University, Hangzhou, China*

²*Department of Economics, University of California, Santa Cruz, US*

³*Department of Marketing, University of New South Wales, Sydney, Australia*

⁴*Minhang Crosspoint Academy at Shanghai Wenqi Private Middle School, Shanghai, China*

**Corresponding Author. Email: 32411117@stu.hzcu.edu.cn*

†These authors contributed equally to this work and should be considered as co-first authors

Abstract. With the rapid growth of social media, the marketing approach of the entertainment and film industry has transferred from one-way promotion to audience participation and community interaction. This study examines how Marvel's digital marketing strategies on social media stimulate audience engagement and contribute to the development of its intellectual property (IP). Using YouTube as a primary data source, this study collects 317,730 comments from 16 videos, among which 13 are about Marvel and 3 are about DC, published within ten years. Topic modeling was conducted using Machine Driven Classification of Open-ended Responses (MDCOR), while Sentiment and Emotion Network Analysis (SENA) was used to analyse audience sentiment and emotions. The findings show that Marvel generates a more diverse online discussion landscape than DC, covering emotional responses, official trailers, character and plot discussion, actor-related content and some offline activities. Marvel's marketing attracted the audience to participate through Easter Eggs, trailers, and fan-oriented content. Besides, the result of SENA revealed that anticipation, joy, surprise, and trust are the dominant positive emotions among audiences, while negative emotions remain relatively limited. These findings suggest that effective IP marketing is not only promoting films, but motivating audiences emotion, making them interpret, discuss, and produce related content. Therefore, this study can provide a data-driven reference for other companies that are seeking to build a fan community and promote intellectual property.

Keywords: Social Media, Data Mining, Marvel, Marketing, Entertainment and Film Industry

1. Introduction

In the contemporary global economy, the entertainment and film industry plays a vital role in driving GDP growth and creating job opportunities. According to the Motion Picture Association (MPA), the

global combined theatrical and home/mobile entertainment market reached \$99.7 billion in 2021, which had a 24% increase compared to 2020. Meanwhile, the global theatrical box office rebounded to \$21.3 billion, with international markets outside North America contributing nearly 79% of total ticket sales. Furthermore, the recent research from MPA pointed out that the American film and television industry supports 2.01 million jobs, pays out \$202 billion in total wages, and includes more than 162,000 businesses. These data reveal the vast economic scale and global reach of the contemporary film and entertainment industry.

Media fusion and the development of Intellectual Property (IP) have been the core engines driving the growth of entertainment industries. There are several particularly outstanding entertainment and film production companies that have independent brand appeal and core IP. For example, Lucasfilm, DC Studios, Pixar Animation Studios, Studio Ghibli, and Marvel. Among them, Marvel, an entertainment company that was on the brink of bankruptcy in the late 1990s, finally constructed the massive Marvel Cinematic Universe (MCU) and evolved into one of the world's most commercially valuable entertainment empires. The main reason for this legendary transformation is not only the success of content production, but also the marketing of IP and the implementation of a profound filmic strategy. So, it is an important example in the present era.

With the popularity of digital media increasing, the simple mode of transmission, which lacks interaction with the audience, cannot satisfy the highly connected consumer market. Besides, with the development of social media, the dissemination speed of entertainment and film companies has accelerated, causing some IPs to become extremely famous on social media and bringing profit to the companies. Most importantly, the audience can post comments and discuss with each other instantly on social media; film and television promotion is no longer dominated by hard advertisements but uses community interaction and emotion for marketing. YouTube, which is the world's largest video and social platform, which has approximately 2.5 to 2.7 billion monthly active users. A large number of well-known entertainment and film companies are stationed here. They post some official trailers, promotional videos and "Easter Eggs" on YouTube to broadcast and market their IP. Therefore, this study adopts big data-driven quantitative and text mining, using YouTube as the core data source.

The aim of this research is to identify and understand the marketing strategies employed by Marvel, which transformed it from a company on the verge of bankruptcy into a multibillion-dollar enterprise. Exploring how this type of marketing can be applied to the current market and social media environment, and what profound significance it holds for the development of an entertainment film company. Providing highly valuable paradigm references for modern enterprises in areas such as IP operation and crisis transformation. Truly successful marketing is no longer just letting consumers know what the producers are selling, but rather inspiring consumers to organize themselves and actively participate in the construction of a brand or IP. By releasing official trailers, embedding "Easter Eggs," and some forward-looking clues in films, Marvel has successfully broken down the barrier between producers and consumers. This strategy immensely stimulates the audience's desire for exploration and curiosity, driving them to form a community on social media. Prompted by curiosity, the audience autonomously engage in plot inference, character analysis and make some creation, posting on the social media platforms will once again trigger intense discussions, which can trigger large secondary marketing through the social media. This form of marketing is based on the interaction of the audience, raising the market's expectation of the sequel and transforming consumers into transmitters. Hence, this study will use topic modeling and sentiment analysis to analyze the comments from YouTube, then explore this type of marketing in the film and entertainment industry.

2. Literature review

2.1. The long-term ecosystem of Marvel

With the rapid development of digital media, particularly the emergence of social media, the way of audience engagement in films has changed greatly. Interactive marketing has become more important in the film industry. Marvel, as a leading film franchise, provides a useful business case to illustrate how audience engagement can bring marketing value. Following its financial crisis in the 1990s, Marvel gradually shifted from relying on character licensing to independent film production and greater control over its intellectual property. Marvel's bankruptcy crisis in the 1990s has been attributed to several factors, including oversupply in the comic book market, underperformance in its diversified businesses, and reliance on a high-debt strategy [1]. Therefore, Marvel started to figure out a new way to create value. Specifically, Marvel shifted from a character licensor to an independent film producer after securing financing of US\$525 million in 2005, which made Marvel less dependent on external film producers [2]. This strategic change in Marvel has brought control over the character development and economic returns, as Marvel has become from an IP licensor to an active film producer [3]. Indeed, the stunning business performance of Iron Man and The Incredible Hulk proved that the strategic transformation generated significant value for Marvel, and film production became the most important source of Marvel's revenue [4]. Additionally, the development of Marvel's IP has been further expanded due to the acquisition of Disney in 2009. With the assistance of Disney's resources on film distribution, brand licensing, product development and international networks, Marvel can create consistent business value through multiple channels and different markets, and no longer relies on the only value of an individual film [5].

This type of business model can go beyond Marvel's own films and products, and extend to collaborations with other commercial brands. When a commercial brand is aligned with a superhero image, its placement in a film can fit naturally into the narrative, increase brand visibility, and create additional value for Marvel's IP [6]. This suggests that the business value of Marvel's IP derives from both its own products and collaborations with other brands. It is important to note that this type of transformation in the business model is consistent with the change in the film industry. The film industry comprises producers, distributors, cinemas, streaming platforms, and other supporting sectors that together form an interconnected ecosystem [7]. Based on current research and business cases, companies in today's film industry that want to make profits in the long term need to do three things: own their IP, distribute their content to larger markets, and continuously develop their IP through different channels.

Having completed its transition from a licence holder to an independent film producer, and having gained access to broader distribution channels through Disney, the key challenge facing Marvel is no longer merely how to produce successful films, but how to continuously integrate these resources and develop the long-term value of its IP. From this perspective, Marvel's development can be further explained through the concepts of dynamic capabilities and cross-media storytelling. Dynamic capabilities emphasise an organisation's ability to adapt to environmental changes by reconfiguring internal resources and external partnerships, which aligns with Marvel's ongoing adjustments from licensing and production to multi-channel commercialisation [8]. Cross-media storytelling can extend a single narrative across films, television series, comics, and games [9]. Further, digital platforms facilitate the spread of content and attract more audiences. Such cross-platform content can encourage audiences to engage with narratives across multiple works and prolong online discussion [10]. Digital platforms can make film and television productions more accessible to audiences while also enabling consumers to participate in content production and

dissemination [11, 12]. Consequently, Marvel's long-term IP ecosystem can be understood as the result of the combined effects of the company's ability to integrate resources, create cross-media content, and sustain audience engagement.

2.2. The roles of trailers and Easter eggs in audience engagement

Within Marvel's IP ecosystem, trailers and Easter eggs not only fulfill traditional promotional functions but also serve distinct narrative and interactive purposes. Film trailers are a key promotional medium for shaping consumer expectations; their typical structure usually introduces characters and settings, presents a conflict or plot twist, accelerates toward the end, and hints at a potential conclusion [13]. In this way, a trailer can give information to the audience, but it does not reveal the most important parts of the story. As a result, the audience will want to know what will happen next. Considering the trailer of Marvel, it will leave suspense and connect the prior and future story plots. This would encourage the audience to discuss the contents of the film [14]. In a film, Easter eggs function similarly. To speak more specifically, Easter eggs connect different works by hiding some clues. They make audiences be willing to analyse plots. They will thus continue discussing Marvel's stories. From this perspective, audiences do not engage with just one film, but are involved in a film universe which keeps expanding [15]. The audiences seek connections between different works, look for and identify the Easter eggs within them. They can then feel more engaged in the film with more immersion.

Trailers and Easter eggs give enough effective information for the audience to engage in, while social media further helps the spread of such engagement. Audiences are no longer merely passive recipients of information but active participants [16]. For Marvel's audience, this type of engagement simply means interactive behaviours including commenting and sharing. Also, it may include user-generated content like reaction videos, character analyses, and plot discussion. An analysis of dialogue in Marvel and DC films found that the language used in these films is highly dependent on context [17]. It suggests that audiences often need to interpret character interactions within the certain context of certain work. This might explain to some extent the reason why audiences continue to analyse characters and film plots after having regarded the films, as well as trailers and Easter eggs. This study has to admit that this evidence is not direct, as the study focused on film dialogue rather than audience behaviour itself. It can be noticed that by providing incomplete but interconnected narrative information, Marvel's trailers and Easter eggs shift audiences from the place of passive viewing to the state of active interpretation and even plot prediction. Afterwards, they further encourage them to post-create and re-create content. It can be seen that trailers and Easter eggs can be regarded as one of the most vital mechanisms for the formation of stronger audience engagement.

2.3. Marketing value created by audience engagement

The development of digital media makes the relationship between audiences and film and television content away from the act of viewing itself as before. The digital economy has increased consumer engagement in the production, marketing, and dissemination of entertainment content [12]. The internet and social media have also extended the duration of audience engagement with film and television content [11]. Audiences used to be merely a component of entertainment consumption, but now they themselves have acquired brand and marketing value. Consumers can collectively reshape brand meaning by creating and spreading their own stories about a brand [18]. Marvel fans speculate plot, analyse trailers and Easter eggs, express and disseminate these views via user-originated

content including comments and reaction videos, and in this process they naturally recreate brand meaning without awareness. That's to say, consumers do not merely passively receive content from Marvel, but also actively interpret and recreate the given content. They participate in this way in the recreation of Marvel's brand meaning. Social media provides a stage for this co-creation process, as well as an ongoing basis. Social media platforms can function as a virtual third space [19]. In this space, users gather, interact, and form emotional attachments. These emotional connections make consumers more willing to choose certain brands and more loyal to the brand. This process can further enhance brand value at the consumer level. Research on Facebook and YouTube communities related to Avengers: Endgame provides more direct evidence regarding Marvel: comments and conversations can strengthen fans' emotional connections to Marvel characters [20]. The sharing of information on storylines and the entire narrative universe is quite helpful. These emotional connections further help to keep longer duration for fan engagement, and their loyalty to the Marvel brand also benefits from it.

From a marketing perspective, sustained consumer engagement is significant, because the value consumers create is not limited to direct purchases. Consumers also generate indirect value by influencing other consumers [21]. In a social media context, this type of indirect value can be generated through user-generated content, such as comments, reaction videos and plot analyses, as well as through dissemination behaviours such as sharing and recommendations. When such content and behaviours contain fans' evaluations, opinions or recommendations regarding Marvel works, and are disseminated to other users, they may form electronic word-of-mouth (eWOM) and expand the reach of Marvel content and influence other consumers. The value of eWOM lies not only in the dissemination of information but also in the generation of consumer feedback. User-generated content, such as online reviews, can broaden the reach of information about films and help production companies understand audience opinions [22]. Consequently, the marketing value of fan discussions, sharing and content creation is primarily reflected in the following: fan engagement can, first and foremost, strengthen their relationship with the Marvel brand through the co-creation of brand meaning and the forging of emotional connections. At the same time, reviews, shares and other user-generated content may also generate eWOM, thereby expanding brand dissemination and further generating indirect consumer value.

2.4. Youtube comments as evidence of audience engagement

The current research indicates that audience engagement can manifest as discussions and content creation, and that these behaviours may in turn generate marketing value through eWOM; therefore, to understand these specific manifestations, it is necessary to further observe the actual expressions of the audience on social media. YouTube, as one of the leading social media platforms, provides a valuable platform to analyse audiences' comments. These comments display directly what audiences discuss and what emotions audiences have. However, the number of comments on YouTube is massive, and there are no uniform formats, which make it hard to read and summarise all these comments manually. Natural language processing and sentiment analysis can assist in identifying whether comments are positive, negative, or neutral [23]. Accordingly, YouTube comments can show us how audiences feel about Marvel's videos or other related content, such as whether audiences feel excited, disappointed or dissatisfied. Nevertheless, it should be noted that these comments do not necessarily mean audiences dislike Marvel's films or the Marvel brand if they post something negative. That is why researchers should consider other audiences' behaviours on YouTube when assessing audience engagement. Participation on YouTube takes many forms, including watching videos, posting and reading comments, liking, sharing, and uploading content;

audiences also participate for different reasons, so not all behaviours represent the same attitude [24]. Therefore, audience engagement on YouTube should be regarded as the result of multiple behaviours and motivations, which can bring measurement issues. A high number of views indicates that a video has been seen by many people or gained attention, but it does not necessarily indicate deep audience engagement [25]. Therefore, the analysis of audience engagement should not only rely on the number of views, but also consider the number of likes, comments and shares, as well as the contents of audience discussions.

Besides deciding which indicator to use to measure audience engagement, whether the data collected from YouTube can truly reflect audience discussion should be considered. Researchers should not randomly select or classify videos and comments when analysing YouTube comments [26]. Instead, the whole analytical procedure should be clearly explained so that others can replicate the process with the same method. Specifically, researchers should explain which videos are analysed, when the comments are collected, and what criteria are used to classify them. If the comments are classified manually, two or more researchers can classify the same comments and then compare the classification results to examine whether they are similar. If the classification results are consistent, this suggests that the criteria are relatively reliable. In addition to researchers' own sampling and coding procedures, YouTube's platform mechanisms can also affect the research sample. YouTube can influence which videos and comments users ultimately see through factors such as personalised settings, search keywords, and channel characteristics [27]. Accordingly, the data collected from social media only represent part of the content on YouTube and may not represent the whole platform. To make it clear of the data collection process, researchers should clearly explain the process of sampling, searching and screening.

After explaining the selection of the sample, researchers still need to consider how to protect YouTube users' privacy and how to collect comments with a consistent method. Although YouTube comments are publicly available, users' privacy still needs to be protected [28]. The comments data should be anonymised for that reason. Researchers should not retain directly the names and account information of users because they are private. They must be careful when quoting a complete original comment. The reason is that others may trace specific users from the comments with the aid of a search engine. And based on these ethical requirements, researchers should collect comment data with a consistent method. The method should be able to be replicated by others. Researchers can automatically obtain video and comment data through the YouTube Data API. The Videos: list endpoint can retrieve public information, including the video title, channel, publication date, and number of views [29]. The *CommentThreads: list* endpoint can retrieve comments under a specific video using its video ID [30].

Marvel maintains audience interest through the long-term development of its IP. It then uses trailers and Easter eggs to stimulate audience discussion and content recreation. These activities can help audiences shape brand loyalty and help spread Marvel-related content. In this way, they create marketing value. This engagement can even be directly observed through social media comments. Comment data provides information for the research on audience engagement. As a key platform for Marvel's official content, YouTube's comments help researchers easily identify the main topics discussed by viewers. Also, the emotional sentiments they express can be retrieved. Platform metrics such as view counts, likes, and the number of comments provide supplementary insights into audience engagement. However, these data only reflect audience reactions on the YouTube platform. It is not applicable for the broader case on Marvel's overall commercial success.

3. Data and study design

3.1. Data collection

The video selection is focused on the trailers, easter eggs, events, and interviews. The research team collected 16 videos from YouTube (Table 1). Thirteen of them talked about Marvel Studios, and three of them talked about DC [31-46]. As of the time of data acquisition (Pacific Daylight Time, August 9, 2026), 317,730 comments were collected from 16 videos using RStudio (version 5.3), including 275,380 parent comments and 42,350 responses. After a discussion, the research team decided to keep the emoji and transfer it to the text version to keep the original emotion of the posters. Due to the huge number of videos and comments, the videos were distributed to 2 companies, Marvel and DC. There were 163,402 comments in DC and 154,328 comments in Marvel. The collected data was exported into a ".csv" file for further analysis.

Table 1. YouTube videos included in the dataset.

No.	Focus and type	Video	Source
1	Marvel Event	SDCC 2024 Hall H: Marvel Studios Panel (full video)	[31]
2	Marvel Documentary	How Bankrupt Marvel Risked Its Main Characters to Save Itself	[32]
3	DC Announcement	Chapter 1: Gods and Monsters	[33]
4	DC Trailer	Superman: Official Teaser Trailer	[34]
5	Marvel Easter eggs	Doctor Strange in the Multiverse of Madness: Iron Man Deleted Scene and Marvel Easter Eggs	[35]
6	Marvel Interview	Avengers: Endgame Cast Full Roundtable Interview on Stan Lee and More	[36]
7	Marvel Documentary	Marvel: The Rise from Bankruptcy to a Multi-Billion Dollar Franchise	[37]
8	DC Interview	Joaquin Phoenix on Playing Joker and Exclusive Outtake	[38]
9	Marvel Documentary	The Story of Marvel Studios: The Making of the Marvel Cinematic Universe	[39]
10	Marvel Announcement	The X-Men Will Return	[40]
11	Marvel Event footage	Robert Downey Jr.'s Entrance to the Walt Disney Studios CinemaCon Presentation	[41]
12	Marvel Trailer	Avengers: Doomsday: Official Trailer	[42]
13	Marvel Film clip	Iron Man Saves the Earth and Dies: Avengers: The Finale, Part 4	[43]
14	Marvel Easter eggs	Fantastic Four: First Steps Breakdown: Marvel Easter Eggs You Missed	[44]
15	Marvel Trailer	Spider-Man: Brand New Day: Official Trailer	[45]
16	Marvel Interview	Robert Downey Jr. Discusses His Iron Man Legacy	[46]

Note. Source numbers correspond to the numbered references in the manuscript.

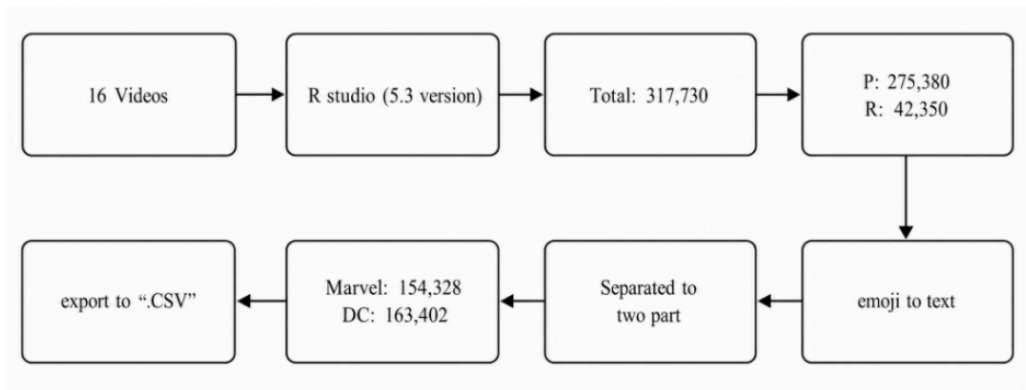


Figure 1. Data collection process

3.2. Study design

After data collection, the research team moved to data analysis. Machine-Driven Classification of Open-Ended Responses (MDCOR) and Sentiment and Emotion Network Analysis (SENA) were used for topic modeling and Sentiment analysis.

MDCOR is a program developed and provided by Professor Manuel. (Figure 2) Open-ended responses were analyzed using MDCOR, a computer-assisted framework that combines topic modeling, text classification, and researcher-led qualitative interpretation. MDCOR was selected because it can classify large collections of text systematically while retaining links between the original responses, respondent identifiers, assigned topics, and other variables in the dataset. The analysis followed a mixed-methods logic in which computational classification and human interpretation were treated as complementary components. MDCOR was used to identify patterns in the text, but the substantive meaning and naming of those patterns remained the responsibility of the research team.

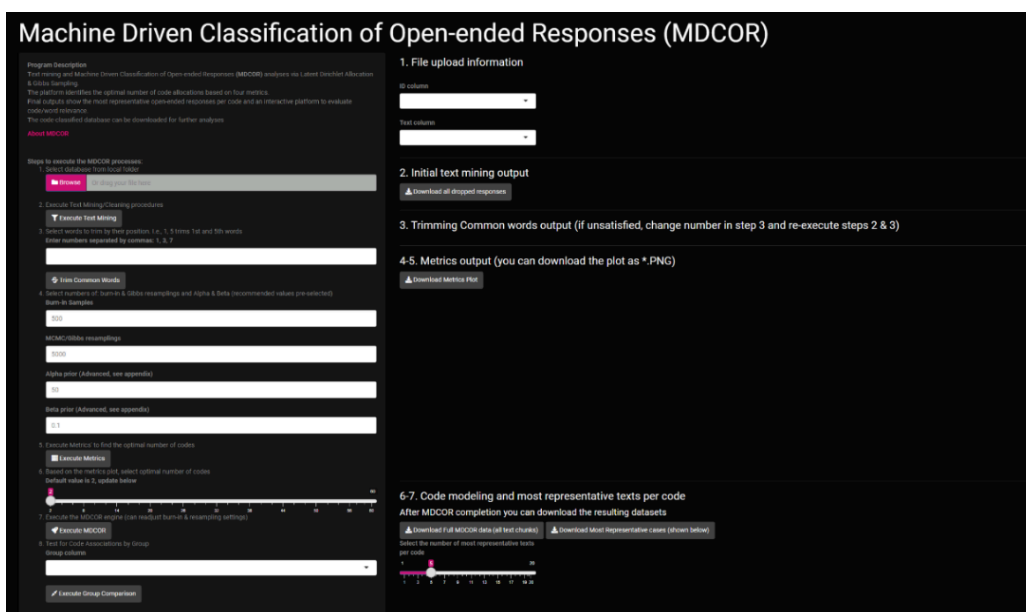


Figure 2. MDCOR

MDCOR uses topic modeling to identify recurring patterns of words that tend to appear together across responses. Rather than beginning with a predetermined coding scheme, the procedure examines the full collection of responses and identifies latent groups of related language. These groups provide an initial set of machine-generated topics, referred to in MDCOR as codes. This approach is particularly useful when the range of issues expressed by participants is not known in advance or when the volume of responses makes entirely manual coding impractical.

Several possible topic solutions were examined before selecting the final model. MDCOR provides multiple diagnostic measures to help researchers evaluate the number of topics (Figure 3) that offers a useful balance between excessively broad categories and unnecessarily fragmented ones. The final number of topics was selected by considering agreement among these measures, the distinctiveness of the resulting word patterns, the coherence of representative responses, and the substantive usefulness of the solution for the research questions. This decision was therefore informed by both statistical evidence and qualitative judgment.

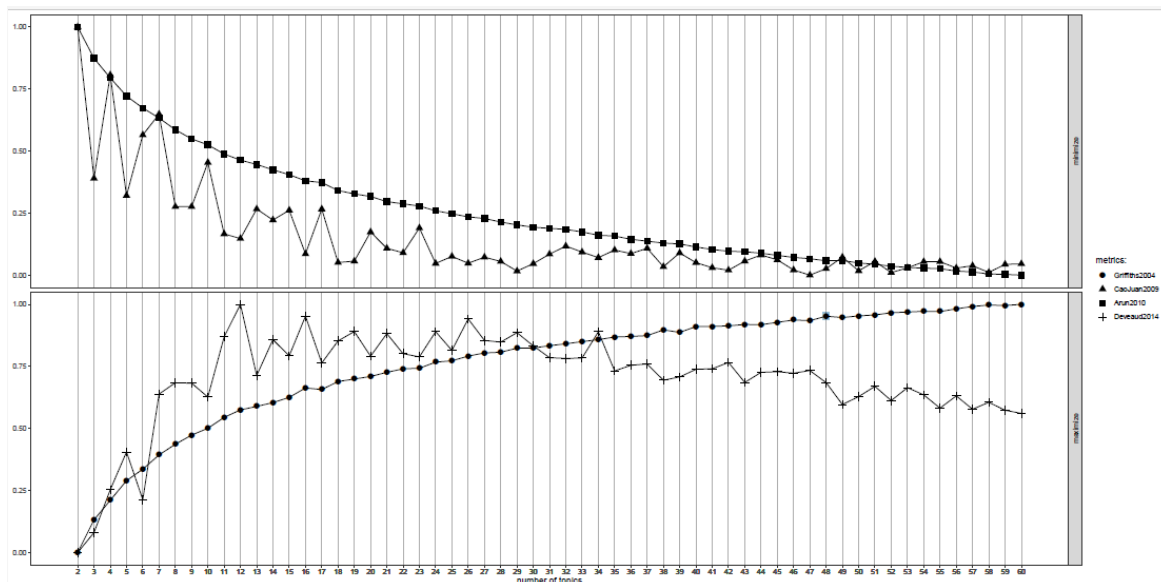


Figure 3. Metric plot

After the 12 topic solutions were selected, MDCOR classified the responses according to their estimated association with each topic. The software produced a classified dataset linking each response and respondent identifier to the relevant machine-generated code. It also generated summaries of the words most strongly associated with each code and identified responses that were especially representative of each topic.

The research team reviewed the prominent words and representative responses for every machine-generated code. At least 20 representative responses per code were examined to determine the central meaning, boundaries, and internal variation of each category. Codes were then assigned descriptive labels that reflected participants' language and the study context. This interpretive review was essential because machine-generated topics identify patterns of textual similarity but do not, by themselves, explain the social or substantive meaning of those patterns. Where a topic contained several related concerns, these distinctions were documented in analytic notes rather than forcing the responses into an overly narrow interpretation.

Following classification, chi-squared tests of independence (Figure 4) were used to examine whether topic membership differed across respondent groups. For each comparison, a contingency

table was constructed using the number of responses classified within each topic and group.

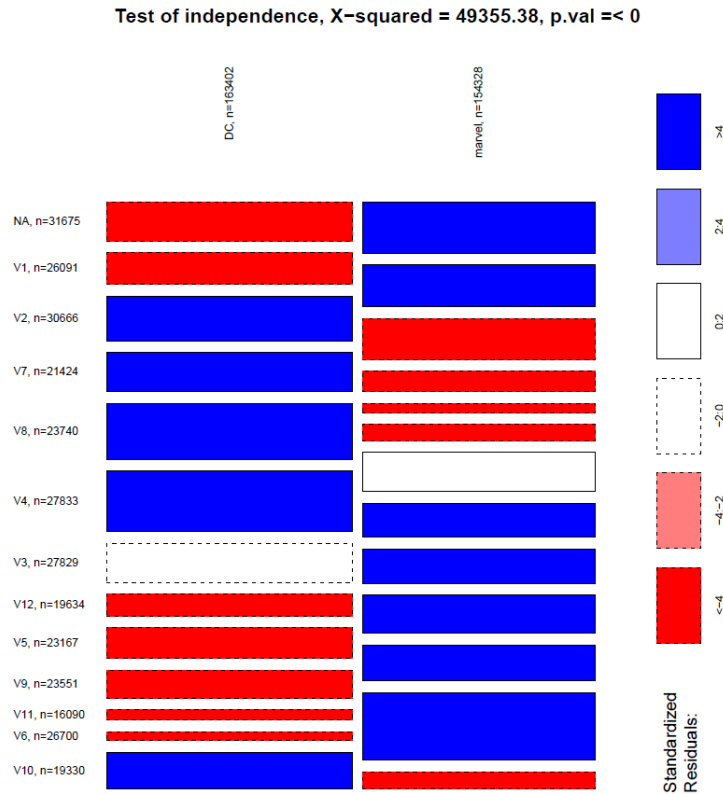


Figure 4. Chi-squared topic and group

The chi-squared test assessed whether the observed distribution of topic classifications differed from the distribution that would be expected if topic membership and group membership were unrelated. MDCOR supports these post-classification group-to-code comparisons and provides visual summaries of the resulting associations.

SENA is another program provided by Professor Manuel. SENA is used to objectively analyze the large amount of individuals' emotions and sentiments. To analyse sentiment (positive and negative) and emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust), SENA uses the National Research Council Canada (NRC) Emotion Lexicon [47]. However, its outputs do not consider participants' attributes and experiences, and they cannot represent the consistency or contradiction of emotions (i.e., emotional entropy) [47]. Emotional entropy measures the consistency or inconsistency of emotions in a text. Its value ranges from 0 to 1: 0 indicates that the emotions are entirely positive or entirely negative, whereas 1 indicates high emotional inconsistency. The two equations used to calculate emotional entropy are described in [47].

$$frequencies = \frac{sentiment_i}{\Sigma sentiment} \quad (1)$$

$$entropy = - \Sigma(frequencies \star \log_2(frequencies)) \quad (2)$$

The research team exported the 20 most representative comments in MDCOR, and SENA would use the NRC Emotion Lexicon to text mining, in order to match each word with 2 sentiments and 8

emotions. Next, SENA would use the equations this study mentioned to calculate the emotional entropy. After that, SENA generated 3 things, WordCloud, sentiment distribution diagram, and QAP.

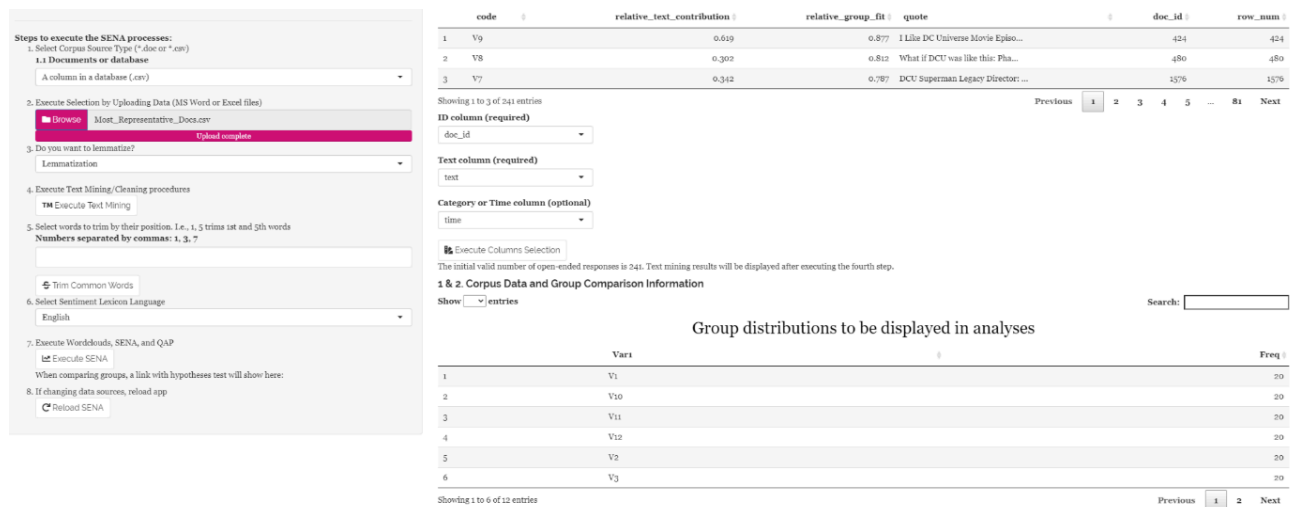


Figure 5. SENA

As briefly described above, the 20 most representative comments can be analyzed by the NRC Emotion Lexicon, which analyzes whether there are sentiment variations of these comments. Alternatively, this database can also be analyzed in SENA following a different approach that involves machine learning.

Our analytic strategy consisted of merging SENA with machine learning text classification and then analyzing whether the resulting machine-learned topics or classes vary in their distribution of emotions and sentiments. Specifically, this study first determined machine-driven text classes via topic modeling as implemented by MDCOR and then analyzed whether there exist variations in the distribution of these sentiments and emotions as a function of the machine-learned classes. This approach enabled us to ask whether machine-learned classes are configured by different distributions of sentiments and emotions as well as the level of emotional entropy those classes contain. From this perspective, SENA may be easily integrated with existing software and with machine learning text classification applications.

Note that a description of the topic modeling via latent Dirichlet allocation is beyond the scope of this paper. For now, let us simply describe that the result of implementing MDCOR with 5000 MCMC/Gibbs resamplings and 500 burn-in samples yielded the following seven topics as the optimal solution of this text classification learning process: V1. Emotional Shock, V2. Debate between MCU/DCU, V3. Interpretive Fan Discussion, V4. Symbol/Emoji, V5. Reactions voting, V6. Prediction about Avengers: Doomsday, V7. affirmation on fans, V8. Fans' opinions, V9. the positive emotion (Love), V10. Criticism of Superman, V11. Prediction about Spider-Man: Brand New Day, V12. Celebration about something.

MDCOR and SENA were run locally on Windows 11, allowing the text to be analyzed without uploading participant responses to an external service. This local workflow was used to support confidentiality and was especially important for responses containing potentially sensitive information. Only de-identified results and aggregated group comparisons re used in reporting. Interactive outputs and representative quotations were reviewed for disclosure risk before dissemination, particularly when subgroup sizes were small.

4. Result

This research mainly focuses on comparing the digital marketing performance of Marvel and DC, the two most influential superhero IPs in the global film and television industry. Based on real comment data collected from the YouTube platform, this study uses multiple common quantitative research tools, including chi-square test, mosaic plot analysis, LDA topic modeling, and QAP network density test. By analyzing audience preference, online discussion content, and network communication structure, this research aims to find out why the two similar entertainment brands show completely different online popularity and fan growth. All statistical results in this paper are significant and reliable, and they do not come from accidental data errors. The detailed analysis of experimental results is explained clearly below.

4.1. Audience preference differences based on mosaic plot and chi-square test

In order to scientifically verify whether Marvel and DC have different audience groups, this study divided all online users into thirteen independent audience subgroups and conducted a chi-square independence test. The final statistical result shows that the chi-square value equals 49355.38 and the p-value is less than 0.001. In statistics, a p-value lower than 0.05 means the result is highly significant. Therefore, the audience preference difference between Marvel and DC is real and stable, rather than a random coincidence.

The mosaic plot (Figure 4) is used to visually display the audience preference distribution of all subgroups. In this figure, dark blue blocks represent that the group's preference is higher than the average expected level, while dark red blocks mean the group's preference is lower than expected. According to the visual data, eight out of thirteen audience subgroups show an obvious preference for Marvel. These users are mostly young Internet users, short-video viewers, and casual entertainment audiences. Most of them do not have long-term comic reading experience. They prefer relaxing, interesting, fast-updating content that fits social media culture.

On the other hand, four audience subgroups clearly tend to support DC. These users are mainly senior comic fans and long-term loyal movie audiences. They value story depth, dark narrative style, and independent character building. Only one subgroup shows neutral attitudes and does not have obvious bias toward either IP.

The difference in audience preference fully reflects the two companies' totally different digital marketing strategies. Marvel adopts an open and active online marketing method. It continuously releases various short videos, movie clips, official trailers, behind-the-scenes footage, and actor interview content on social platforms. Moreover, Marvel allows and encourages fan-made creations and cooperates with many online influencers to expand its online exposure. With its huge multiverse story system, Marvel maintains long-term and stable online popularity and successfully attracts a large number of new young audiences outside traditional fan circles.

In contrast, DC's digital marketing style is much more conservative and closed. It focuses more on dark, serious, and standalone superhero stories. DC does not actively produce lightweight social media content and has strict restrictions on fan derivative works. Although this strategy successfully maintains high loyalty among old core fans, it fails to attract ordinary young viewers who prefer casual and entertaining online content. As a result, DC's fan group is stable but lacks new blood and online activity.

4.2. Online public opinion features analyzed by LDA topic modeling

In order to systematically summarize what the audience talks about online, this study used the LDA topic model (Figure 6) to automatically classify and cluster massive YouTube comment data. After repeated model testing and screening, twelve stable and effective public opinion topics were extracted. These topics can be divided into four major categories: audience emotional feelings after watching, evaluation of work quality, discussion of characters and plots, and various topics derived from online marketing activities. The clustering result is highly consistent with real Internet discussion situations, which proves the validity of the model output.

For Marvel, the extracted public opinion topics are diverse, active, and entertainment-oriented. Most audience discussions focus on exciting trailer previews, interesting character personalities, wonderful movie scenes, short video highlights, and positive emotional experiences while watching. In addition, many discussions are related to fan creation, online trend topics, and social media sharing behaviors. These lightweight and lively topics are very suitable for young people's fragmented reading and viewing habits on the Internet. This fully proves that Marvel's digital marketing strategy is extremely successful in stimulating audience interaction, improving user participation, and creating continuous online heat.

For DC, the overall style of public opinion topics is completely different. Most comments are concentrated on professional and in-depth content, including the comparison between movie plots and original comic books, the rationality of plot adaptation, the analysis of character settings, and the discussion of story themes. There are almost no casual, popular, or marketing-driven discussion topics on social media. It shows that DC's online content is too professional and serious to spread widely on public social platforms. Its online influence is limited within small professional fan groups and cannot achieve large-scale viral communication.

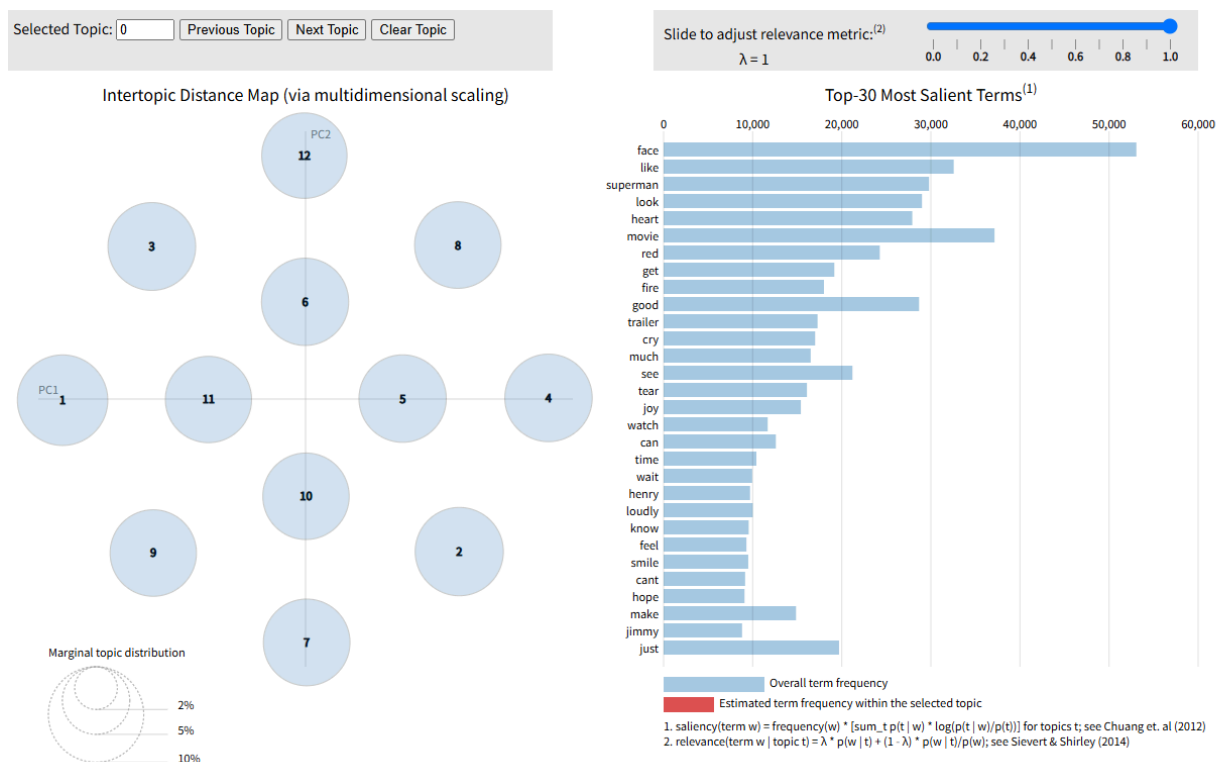


Figure 6. LDA topic modeling

Through LDA topic analysis, this study can reasonably explain the essential reason for the audience gap between the two IPs. Marvel relies on rich and diverse digital marketing content to continuously attract new casual audiences and expand its fan base. DC relies on high-quality and in-depth story content to retain old fans, but it cannot effectively expand its influence among the general public.

4.3. Communication structure differences verified by QAP density test

To further explore the difference in online communication ability between Marvel and DC, this research adopted the QAP network density test. Unlike traditional statistical methods, the QAP test can effectively solve the autocorrelation problem of network data and ensure that the experimental results are accurate and statistically credible.

The QAP density visualization (Figure 7) clearly shows the structural gap between the two public opinion networks. Marvel's online discussion network has extremely high node density and tight connection relationships. Different discussion topics can interact with each other, drive new discussions, and form a complete and closed online communication loop. User comments, topic reposting, and fan content spreading are very frequent and active. Therefore, Marvel's marketing topics can quickly spread across different social fields, continuously attract public attention, and maintain long-term online popularity.

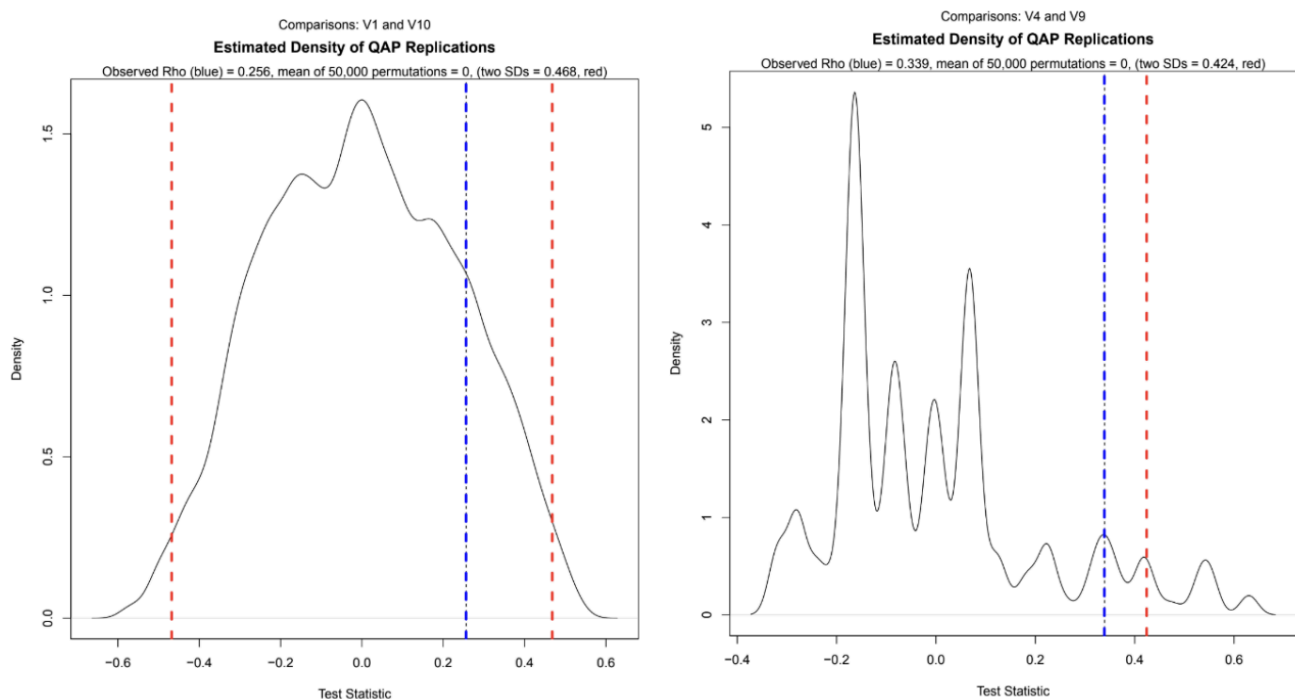


Figure 7. The QAP density visualization (2 of 66)

On the contrary, DC's public opinion network structure is relatively sparse and scattered. Most discussions are isolated and only focus on the content of movies and comics themselves. There are almost no cross-topic connections, secondary spreading paths, or viral derivative discussions. The entire communication system lacks expansion ability. Its online discussion heat is always limited inside the fixed core fan circle and cannot break through to attract broader public audiences. The QAP test result provides strong quantitative evidence to prove that Marvel has obvious advantages

in digital marketing and online communication, while DC has clear shortcomings in social media dissemination.

4.4. Comprehensive research conclusion

Combining all the above statistical tests, visual figures, and topic analysis results, all core findings of this study are statistically significant and reliable. The data fully proves that Marvel and DC have formed two completely different digital marketing models and audience operation systems. Marvel adopts an open, entertainment-oriented, and social-media-friendly digital marketing strategy. It successfully breaks the limitations of traditional fan circles, continuously attracts young new audiences, and maintains active, diverse, and vigorous online public opinion. In contrast, DC uses a closed, in-depth, and content-oriented marketing model. It maintains extremely high loyalty among core old fans, but its online spreading ability is weak, and it cannot effectively expand its influence in the mass Internet environment.

To conclude, this research successfully quantifies the different audience effects and communication results brought by different digital marketing strategies. It clearly explains why Marvel and DC have completely different online development trends. This study provides objective experimental data and practical suggestions for the future IP operation, social media promotion, and digital marketing optimization of superhero works and other similar entertainment IP projects.

5. Discussion

This study explored how Marvel's social media marketing strategies stimulate audience discussion and how audience responses on YouTube can reflect the effectiveness of these strategies. By analyzing 317,730 comments from 16 videos, including 13 Marvel-related videos and 3 DC-related videos. Combining MDCOR for topic modeling and SENA for sentiment and emotion analysis, this study provides an audience-centered perspective on Marvel's digital marketing strategies.

The result of the findings firstly answers that Marvel's trailers, Easter eggs, connected narratives, and fan-oriented activities contribute to diverse forms of audience discussion. This study used MDCOR to identify the discussion of the audience concerning trailers, characters and plots, emotional reactions, actors, fan activities, and social media sharing. This finding is consistent with previous research on transmedia storytelling, which shows that different media texts can contribute to a larger narrative system and that interconnected stories can encourage audiences to engage across media [9, 10]. Our findings extend this perspective by showing how these narrative connections are reflected in actual audiences' discussions on YouTube. Moreover, this research indicates that Easter eggs can function as a mechanism of audience participation by encouraging audiences to search for connections between hidden information and the storyline [15]. Besides, the results also discover that the audience would discuss clues, characters, future development, and relationships between different Marvel stories. Therefore, the value of Easter Eggs is not only for the existing core fans, but also for creating discussion topics that can help the audience understand actively and connect different parts of the Marvel IP

Secondly, according to SENA analysis, positive emotions occupy a prominent position among the audience, among which anticipation, joy, surprise, and trust stand out; however, negative emotions are limited. Besides, Marvel-oriented audience groups also show more obvious positive emotions. These results are similar to the research emphasizing the importance of emotional responses and engagement in social media and digital consumer behavior [19, 20]. At the same time, the results show that this study should understand emotions combined with discussion behavior. The QAP

analysis further suggests that marketing activities involving connected narratives, fan activities and continuing content are associated with stronger communication relationships. It supports the opinion that social media engagement involves more than passive exposure. Previous research likewise indicates that YouTube engagement should not be understood only through basic indicators such as views and that audience interaction is important within YouTube environments [24, 25]. Our research expands these opinions, showing the relationship among discussions and engagement.

The main contribution of this study is that it connects several areas of research that have often been examined separately. Existing literature has discussed social media marketing, transmedia storytelling, Marvel's IP development, YouTube engagement, audience participation, and digital consumer behavior. For instance, existing research has highlighted the importance of creative social media content for consumer engagement and examined how social media marketing activities can influence consumer responses [16, 21]. Research has also examined transmedia storytelling and interconnected narratives, as well as engagement within YouTube environments [9, 10, 24, 25]. However, these perspectives have not necessarily explained how specific IP marketing strategies are reflected in the discussion, feelings and connection. This study combined the comments from YouTube with MDCOR and SENA analysis to address this gap. User-generated content can help explain audiences and digital media participation [22]. This research discussed how audiences respond to, understand and spread the information instead of treating the audience only as receivers. So, this study contributes to the existing literature by shifting the analysis from a primarily company-centered perspective toward an audience-centered perspective. Instead of asking only what Marvel communicates through its marketing, the study also examines the audience's reaction after encountering them

Furthermore, the findings have practical meanings for entertainment and film companies and IP managers. First of all, the results reflect that social media marketing can not only focus on promoting related information. Instead, the marketing staff should design content that creates opportunities for audience interpretation and discussion; trailers can stimulate expectation, and Easter Eggs can motivate curiosity. This is consistent with broader research on the role of creative social media content [16]. In addition, the results of Marvel reveal that different forms of content may be suitable for different audience groups, which indicates that IP marketing can combine accessible social media content with deeper narrative material in order to attract new audiences while maintaining existing audience interest. Furthermore, the companies can get feedback on their content through comment analysis, which may be helpful for their further development.

There are still several limitations in this research that should be considered. The data of this study focuses on YouTube. Although YouTube provides a substantial amount of publicly available audience discussion, user behavior varies across platforms. Research on social media and YouTube indicates that platform characteristics can influence how users communicate and interact [24-26]. So, the results of the research cannot be directly applied to other social media platforms, such as TikTok and Instagram. In the second place, this study is based on a selected video sample, and the number of videos related to DC is significantly less than that of those related to Marvel. The dataset contains 317,730 comments, but the selected videos cannot fully represent every aspect of their marketing strategies, especially DC. Lastly, the study analyzes publicly available comments. However, users who choose to post comments may differ from audiences who only watch, like, or share content. Consequently, the comments analyzed in this study should not be considered a complete representation of the entire Marvel or DC audience.

Overall, this research suggests that the effectiveness of social media marketing by entertainment and film companies is not only dependent on content promotion. It is also related to the

opportunities for audiences to participate, which are created by content. Therefore, effective IP marketing can transform promotion into participation. When audiences actively discuss and spread the content, they become a part of the dissemination process surrounding the IP, providing entertainment and film companies with a practical opportunity to generate economic benefits.

Acknowledgement

Yibo Huang and Xuhui Rong contributed equally to this work and should be considered co-first authors.

References

- [1] Calandro, J. Jr. (2016). Distressed M&A and corporate strategy: Lessons from Marvel Entertainment Group's bankruptcy. Revised draft, Gabelli Center for Global Security Analysis, Fordham University.
- [2] Marvel Enterprises, Inc. (2005). Marvel launches independently financed film slate with closing of \$525 million non-recourse credit facility. Media release, 6 September.
- [3] Johnson, D. (2012). Cinematic destiny: Marvel Studios and the trade stories of industrial convergence. *Cinema Journal*, 52(1), 1–24. <https://doi.org/10.1353/cj.2012.0108>
- [4] Marvel Entertainment, Inc. (2009). Annual report 2008: Form 10-K. Filed 27 February 2009, U.S. Securities and Exchange Commission.
- [5] The Walt Disney Company. (2009). Disney to acquire Marvel Entertainment. Media release, 31 August.
- [6] Papp-Váry, Á. F., & Rönky, Á. (2025). Building brands with superheroes, building superheroes with brands: The brands of Iron Man and Captain America in the Marvel Cinematic Universe. *Businesses*, 5(2), Article 19. <https://doi.org/10.3390/businesses5020019>
- [7] Motion Picture Association. (2024). The American motion picture and television industry: Creating jobs, trading around the world. 6 March.
- [8] Florea, D. L., & Munteanu, C. C. (2020). Lone wolf or social monkey? The role of marketing outsourcing in the development of second-order marketing competences. *Industrial Marketing Management*, 86, 191–200. <https://doi.org/10.1016/j.indmarman.2019.11.017>
- [9] Scolari, C. A. (2009). Transmedia storytelling: Implicit consumers, narrative worlds, and branding in contemporary media production. *International Journal of Communication*, 3, 586–606.
- [10] Webster, L., Leaver, T., & Sandry, E. (2022). Transmedia storytelling during the COVID-19 pandemic: Marvel's WandaVision and Zack Snyder's Justice League. *First Monday*, 27(7). <https://doi.org/10.5210/fm.v27i7.11784>
- [11] Pardo, A. (2012). Digital Hollywood: How Internet and social media are changing the movie business. In *Handbook of Social Media Management: Value Chain and Business Models in Changing Media Markets* (pp. 327–347). Springer. https://doi.org/10.1007/978-3-642-28897-5_19
- [12] Currah, A. (2006). Hollywood versus the Internet: The media and entertainment industries in a digital and networked economy. *Journal of Economic Geography*, 6(4), 439–468. <https://doi.org/10.1093/jeg/lbl006>
- [13] Finsterwalder, J., Kuppelwieser, V. G., & de Villiers, M. (2012). The effects of film trailers on shaping consumer expectations in the entertainment industry—A qualitative analysis. *Journal of Retailing and Consumer Services*, 19(6), 589–595. <https://doi.org/10.1016/j.jretconser.2012.07.004>
- [14] Jiang, L. (2024). Research on Marvel Studios' product marketing strategy in the new media environment. *SHS Web of Conferences*, 181, 04009. <https://doi.org/10.1051/shsconf/202418104009>
- [15] Watts, J., & Wing, H. (2026). Harvesting Easter eggs: An exploratory study of enjoying transnarrative media. *PLOS ONE*, 21(2), e0341588. <https://doi.org/10.1371/journal.pone.0341588>
- [16] Ashley, C., & Tuten, T. (2015). Creative strategies in social media marketing: An exploratory study of branded social content and consumer engagement. *Psychology & Marketing*, 32(1), 15–27. <https://doi.org/10.1002/mar.20761>
- [17] Forchini, P., & Forchini, P. (2021). Movie discourse: Marvel and DC Studios compared. In E. Friginal & J. A. Hardy (Eds.), *The Routledge Handbook of Corpus Approaches to Discourse Analysis* (pp. 183–201). Routledge. <https://doi.org/10.4324/9780429259982-12>
- [18] Swaminathan, V., Schwartz, H. A., Menezes, R., & Hill, S. (2022). The language of brands in social media: Using topic modeling on social media conversations to drive brand strategy. *Journal of Interactive Marketing*, 57(2), 255–277. <https://doi.org/10.1177/10949968221088275>

- [19] Dwivedi, A., Johnson, L. W., Wilkie, D. C. & De Araujo-Gil, L. (2019). Consumer emotional brand attachment with social media brands and social media brand equity. *European Journal of Marketing*, 53(6), 1176–1204. <https://doi.org/10.1108/EJM-09-2016-0511>
- [20] Atarama-Rojas, T. & Feijoo, B. (2023). Social audience and emotional bonding in Marvel's transmedia phenomenon: An exploration of Peruvian digital communities. *International Journal of Communication*, 17, 3054–3071.
- [21] Li, F., Larimo, J. & Leonidou, L. C. (2021). Social media marketing strategy: Definition, conceptualization, taxonomy, validation, and future agenda. *Journal of the Academy of Marketing Science*, 49(1), 51–70. <https://doi.org/10.1007/s11747-020-00733-3>
- [22] Yang, J. & Yecies, B. (2016). Mining Chinese social media UGC: A big-data framework for analyzing Douban movie reviews. *Journal of Big Data*, 3, 3. <https://doi.org/10.1186/s40537-015-0037-9>
- [23] Bhuiyan, H., Ara, J., Bardhan, R. & Islam, M. R. (2017). Retrieving YouTube video by sentiment analysis on user comment. *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, 474–478. <https://doi.org/10.1109/ICSIPA.2017.8120658>
- [24] Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior*, 66, 236–247. <https://doi.org/10.1016/j.chb.2016.09.024>
- [25] Wu, S., Rizoio, M.-A., & Xie, L. (2018). Beyond views: Measuring and predicting engagement in online videos. *Proceedings of the International AAAI Conference on Web and Social Media*, 12(1), 434–443. <https://doi.org/10.1609/icwsm.v12i1.15031>
- [26] Yun, G. W., Allgayer, S., & Park, S.-Y. (2020). Mind your social media manners: Pseudonymity, imaginary audience, and incivility on Facebook vs. YouTube. *International Journal of Communication*, 14, 3418–3438.
- [27] Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to "ranking cultures": Investigating the modulation of visibility in YouTube search results. *Convergence*, 24(1), 50–68. <https://doi.org/10.1177/1354856517736982>
- [28] Tanner, J. P., Takats, C., Lathan, H. S., Kwan, A., Wormer, R., Romero, D., & Jones, H. E. (2023). Approaches to research ethics in health research on YouTube: Systematic review. *Journal of Medical Internet Research*, 25, e43060. <https://doi.org/10.2196/43060>
- [29] Google. (2026a). Videos: List. *YouTube Data API documentation*, Google for Developers.
- [30] Google. (2026b). CommentThreads: List. *YouTube Data API documentation*, Google for Developers.
- [31] blackfilmandtv. (2024, July 27). *SDCC 2024 Hall H - Marvel Studios Panel (full video)* [Video]. YouTube. <https://www.youtube.com/watch?v=KMNrVGuoYkc>
- [32] Cheddar. (2018, August 14). *How bankrupt marvel risked its main characters to save itself - cheddar examines* [Video]. YouTube. <https://www.youtube.com/watch?v=PSaoKQCWAbI>
- [33] DC. (2023, January 31). *Chapter 1 - gods and monsters | DC studios | DC*. YouTube. <https://www.youtube.com/watch?v=wY8XcmrIujE>
- [34] DC. (2024, December 19). *Superman | official teaser trailer*. YouTube. <https://www.youtube.com/watch?v=uhUht6vAsMY>
- [35] Emergency Awesome. (2022, May 10). *Doctor strange multiverse of madness iron man deleted scene and marvel easter eggs* [Video]. YouTube. <https://www.youtube.com/watch?v=COkQVvie0jc>
- [36] Entertainment Weekly. (2019, April 10). 'Avengers: Endgame' cast full roundtable interview on Stan Lee & more | entertainment weekly [Video recording]. In *YouTube*. <https://www.youtube.com/watch?v=ECdQ6OV-rkU>
- [37] FD Finance. (2024, June 28). *Marvel: The rise from bankruptcy to a multi-billion dollar franchise | Inside the storm | FD Finance* [Video]. YouTube. <https://www.youtube.com/watch?v=59ToHTFma3A>
- [38] Jimmy Kimmel Live. (2019, October 2). Joaquin Phoenix on playing Joker + exclusive outtake [Video recording]. In *YouTube*. <https://www.youtube.com/watch?v=tZNBy-BdM2A>
- [39] Marvel Entertainment. (2021, September 14). *The story of Marvel Studios: The making of the Marvel Cinematic Universe* [Video]. YouTube. <https://www.youtube.com/watch?v=dBmRPPnrm1s>
- [40] Marvel Entertainment. (2026a, January 6). *The X-Men will return | Avengers: Doomsday in theaters December 18, 2026*. YouTube. <https://www.youtube.com/watch?v=kH1XlwHQv9o>
- [41] Marvel Entertainment. (2026b, April 17). *Robert Downey Jr.'s entrance to the Walt Disney Studios CinemaCon Presentation* [Video]. YouTube. <https://www.youtube.com/watch?v=XRVP9YTLMc0>
- [42] Marvel Entertainment. (2026c, July 20). *Avengers: Doomsday | Official trailer | In theaters December 18* [Video]. YouTube. <https://www.youtube.com/watch?v=irVNGjRFZGk>
- [43] Marvel universe vs Thanos army: Iron man saves the Earth and dies / avengers: The finale - part 4. (n.d.). *Www.Youtube.Com*. Retrieved 25 August 2026, from <https://www.youtube.com/watch?v=-NRm5XZfV7A>

- [44] ScreenCrush. (2025, July 25). *Fantastic four: First steps breakdown - marvel easter eggs you missed!* [Video]. YouTube. https://www.youtube.com/watch?v=fF_zF0KR3Tk
- [45] Spider-Man. (2026, March 18). *Spider-man: Brand new day – official trailer (HD)*. YouTube. <https://www.youtube.com/watch?v=8TZMtslA3UY>
- [46] Stern, H. (2019, April 26). Robert Downey Jr. discusses his iron Man Legacy [Video recording]. In *YouTube*. <https://www.youtube.com/watch?v=Sdj-zpcaqeE>
- [47] González Canché, M. S., & Zheng, K. (2025). Sentiment and emotion network analysis—SENA: An analytic framework and no-code software to integrate natural language processing, linguistic analysis, and data science with qualitative and mixed-methods research. *International Journal of Qualitative Methods*, 24. <https://doi.org/10.1177/16094069251346929>