# Using LDA Topic Modeling to Understand Regrowth Factors of the Chinese Gaming Industry in the COVID-19 Era: Current Situation, Future and Predicament

Yi-Qian Han[1], Won-Jun Jeong[1], Gi-Sung Oh[1],
Seok Hee Oh[2] and Taeg-Keun Whangbo[1,*]

[1]*Department of Computer Engineering, Gachon University, Seongnam-si, Gyeonggi-do, 13120, Republic of Korea*
[2]*Korea creative content agency, Daejeon gwangyeoksi, 34863, Republic of Korea*
*E-mail: 971008@gachon.ac.kr; tp04045@gachon.ac.kr; eustia@gachon.ac.kr; tkwhangbo@gachon.ac.kr; shoh@kocca.kr*
*\*Corresponding Author*

## Abstract

The gaming industry, which was among the industries least affected by the COVID-19 outbreak, exhibited positive growth trends during the COVID-19 period. This paper explores the impact of COVID-19 on the gaming industry by analyzing news texts from 2020 to 2022 using text mining and LDA (latent Dirichlet allocation) topic classification, and visualizing charts. The study focuses on three themes, namely the current situation, the future, and possible problems of China's game industry in the post-epidemic era.

The findings of this study suggest that the development of the game industry during the COVID-19 outbreak prompted the government to regulate policies and promote the transformation of game companies, which had a positive impact on the development of China's game industry. However,

this study also found that due to the effects of COVID-19, society and the government have increased their focus on the time management of underage game users, which poses a significant challenge to the games industry.

This paper recommends improvements from three perspectives, namely society, policy, and enterprise, with the aim of contributing to the long-term development of China's game industry.

**Keywords:** Topic modeling, games industry, COVID-19, text mining, latent Dirichlet allocation, news text.

## 1  Introduction

The COVID-19 pandemic significantly impacted people's social lives, high-lighting the need for continued economic recovery and development in the post-epidemic period, which is a common challenge for the international community. The widespread and rapid spread of the virus has severely impacted economies, businesses, and people's lives, exacerbating concerns about the development of the real economy. However, the digital economy, as the least affected economic system, has experienced growth during the pandemic, particularly in the game industry, where China's actual sales reached US$46.6 billion in 2021, representing an increase of US$2.6 billion compared to 2020, and a year-on-year increase of 6.4% [1].

This research employs LDA (latent Dirichlet allocation), a model that serves as an analysis tool for mining the hidden development topics of news texts, to analyze news texts related to the Chinese game industry during the COVID-19 period from 2020 to 2022. The study aims to identify the relevant topics related to the game industry, analyze the current situation and policies of the Chinese game industry, and explore prospects and potential problems. Additionally, the paper presents three suggestions for the future development of the game industry in the post-epidemic era.

## 2  Background Theory

Topic models are methods of modeling articles to obtain abstract topics. This method can automatically find the hidden semantic topics in the text data [2]. The topic model is used in text mining [3]. In data science: many scholars rely on the analytical capabilities of LDA in patent collation and academic paper analysis, and organizing big data for scientific research materials to classify and analyze the literature and the industry. For example, some scholars

have constructed the LDA institution-topic model to realize the connection between the subject of patent knowledge and the institution, providing more valuable information for industry dynamic tracking [4]. Wang et al. analyzed the graphene industry patents and papers based on the LDA topic model, established a comparative study of integrating linked data and carried out visualization processing to improve the efficiency of LDA analysis [5]. Scholars such as Wang selected academic papers in different fields, proposed a scientific literature classification method through the LDA topic model, and verified the validity of the LDA topic model in the literature [6].

The topic model is used in journalism and social media commentary. Scholars have applied it to Twitter text analysis and found that the LDA topic model in the dataset can effectively discover hidden topics and perform sentiment analysis [7]. Liu collected news media reports and used LDA analysis to study the positive role of media reports in the COVID era [8]. Zhang extract and classify social media comments related to hot news by LDA topic for analysis [40]. Some scholars collect and define the characteristics of new concepts by analyzing news, which helps to deepen the understanding of the metaverse field [9]. Kumar examines the literature on renewable energy projects through LDA; the content, from the perspective of social science, studies the necessary factors for industrial success [10]. This paper uses news text for LDA analysis as the research method. First, we visualize the distribution of news keywords and topics, and analyze topics from a social science perspective by combining charts with actual literature. Next, we analyze the impact of the COVID era on China's game industry, and put forward suggestions for improvement from the three views of enterprise, government, and media. Finally, we provide a reference for the game industry to adapt to the development of the post-epidemic era.

## 3 Research Methods

### 3.1 Research Frame

This study utilizes a web crawling technique to collect news articles for analysis. After data cleaning, the corpus is tokenized and stop words are removed. Subsequently, LDA modeling is employed to reveal latent topics within the news articles. The data processing flow is illustrated in Figure 1. The primary objective of this research is to investigate the impact of COVID-19 on China's gaming industry and to uncover hidden themes in the articles. The study's results are categorized into three sections: the
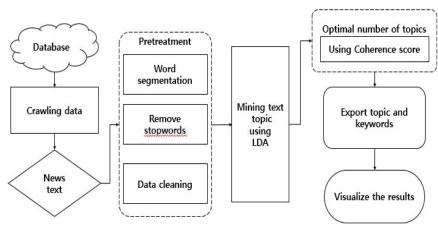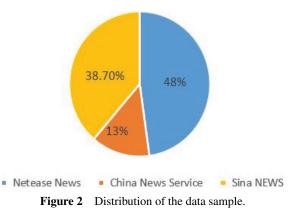
**Figure 1**    Research framework of news text mining based on LDA.

present situation, future prospects, and potential challenges of the industry. Finally, the study's conclusion summarizes the findings and provides relevant literature references to further enhance researchers' understanding in related fields.

## 3.2  Data Acquisition

The sources of the news data in this study are Chinese news websites, namely "China News," "Sina News," and "NetEase News." These websites were selected for specific reasons: China News is the official news media in China and provides valuable insights into government policies; Sina News has extensive coverage of news in China, making it a comprehensive source of information; and NetEase News is one of the most prominent news websites on the Chinese internet, providing a reflection of public opinion. The news data used in this study covers the period from December 2020 to August 2022, and was collected using web crawlers. The Python-based "bs4(Beautiful Soup 4)" package was used to implement the web crawler framework [11]. The crawled data include the news release time, news URL, news search keywords, news title, and news text. The use of web crawlers provides a high degree of flexibility, allowing the data to be obtained and stored according to its unique characteristics. Moreover, it facilitates data preprocessing [12].

This article utilizes the keyword "game" as the primary search term for web crawling, while also including "medical," "economy," and "COVID-19

**Figure 2**    Distribution of the data sample.

epidemic" as additional keywords to extract relevant news articles. The resulting dataset comprises 23,478 news texts, the distribution of which is visualized in Figure 2.

## 3.3  Data Cleaning

After acquiring the local corpus using the method discussed in the previous section, the news headlines and texts cannot be used directly, and data preprocessing is necessary. In this study, the data undergoes several stages of processing, including data cleaning, word segmentation, and stop word removal, to remove factors that can affect the accuracy of the experiment. Only after these steps can LDA analysis be conducted on the textual data.

## 3.4  Text Segmentation

This study adopts a text analysis approach that employs words, sentences, and punctuation as fundamental units of analysis. The text segmentation method employed is based on the TextRank algorithm [13], which divides the text into smaller units such as words and sentences and utilizes a graph model and voting mechanism to extract the salient components in the text, allowing for keyword extraction and summarization of information. The data sample used in the study consists of numerous medium and long articles, making text segmentation a challenging task. The TextRank algorithm effectively handles the large volume of text for segmentation without the need for prior training on multiple documents. Additionally, using TextRank enhances the efficiency of LDA analysis, as demonstrated in the comparison table of LDA analysis times in samples containing the TextRank algorithm (Table 1).

**Table 1**    Comparison of using the Textrank algorithm and without using it

| Algorithm | Text | Time |
|---|---|---|
| TextRank | … Appropriately/balancing /videogame's/difficulty /essential/provide/players/pleasant experience/ study/Korean/scientists/developed/novel/approach/ dynamic difficulty/adjustment/players' emotions/ estimated/using/in-game/data/difficulty/level/ tweaked/accordingly/maximize/player/satisfaction… | 15 min |
| Non-TextRank | … Appropriately/balancing/ videogame/difficulty/ essential/provide/players/with/pleasant/experience/ recent/study/Korean/scientists/developed/novel /approach/dynamic/difficulty/adjustment/where/ players/emotions/estimated/using/in/game/data/and the/difficulty/level/ tweaked/accordingly/maximize /player/satisfaction… | 25 min |

This study utilized the "jieba" library to conduct the word segmentation process. "Jieba" is a Python-based natural language segmentation tool [14]. It supports the TextRank analysis function, which enables text segmentation, keyword extraction, word relationship building, and TextRank algorithm calculation using a graph model. The stop word list was also utilized to retain nouns during the segmentation process, which increased the frequency of certain words and phrases while preventing the loss of part of speech due to inadequate word frequency during word segmentation.

## 3.5  LDA (Latent Dirichlet Allocation)

This research employs a topic model called LDA (latent Dirichlet allocation) [15] to analyze the corpus of documents. The model determines the topic of each document in the corpus through probability distributions. The process of generating a document in the LDA model involves:

- Setting the total number of words in the article ($\alpha$), the total number of topics in the article ($\beta$) and the number of all topics ($K$).
- Generating the topic distribution ($O_m$) of the document ($m$) from $\alpha$ and generating the topic ($z_{mj}$) of the *j*th word of this document ($m$) from the topic distribution ($O_m$).
- Generating the word distribution ($\varphi_k$) in topic ($z_{mj}$)from $\beta$.
- Sampling the final word ($w$) from the word distribution ($\varphi_k$).
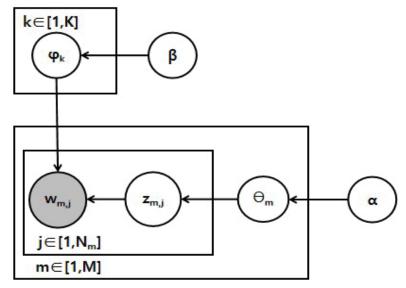- Repeating the above operation until word number is the total number of words ($\alpha$).

**Figure 3** Graph model structure of LDA.

**Table 2** The variable names and definition

| Variable | Definition |
|---|---|
| $M$ | Notes the number of documents |
| $N$ | Number of words in document $m$ |
| $K$ | Topic number |
| $\alpha$ | Parameter of the Dirichlet prior to the per-document topic distributions (the higher the number, the more topics may be included in each document) |
| $O_m$ | Topic distribution for document $m$ |
| $z_{mj}$ | Topic for the $j$th word in document $m$ |
| $\beta$ | Parameter of the Dirichlet prior to the per-topic word distribution (the higher, the more words in each topic) |
| $\varphi_k$ | Word distribution for topic $k$ |
| $w_{mj}$ | Specific words |

However, LDA runs in the reverse process of the above steps, where it speculates which topics the analyzed documents have and the probability of each topic appearing in each article (topic distribution) [16]. Figure 3 illustrates the procedure of the LDA algorithm and Table 2 defines the variable names. A document can contain multiple topics, with keywords generated by one of the topics. The LDA topic model is a Bayesian probabilistic model consisting of three layers: words, topics, and articles [17].

**Table 3**    Using time and results

| Algorithm | Example of Results | Time |
|---|---|---|
| LDA Mallet | Metaverse/Tencent/videogames/virtual reality /VR/contents/assets/league of legends/blockchain /experience | 50 min |
| Gensim LDA | Meta/Tencent/virtual/reality/VR/ chain/welfare/player/assets/league/experience | 37 min |

We use environment of the LDA model in this study as follows. The experimental setting is as follows: Intel i72.8GHz quad-core processor, 16.0GB memory, Windows 10 operating system, and coding by Python.

In this article, the LDA algorithm is implemented using the "Mallet" version. "Mallet LDA" is a Java-based NLP machine learning algorithm package that includes a fast and scalable implementation of Gibbs sampling, an optimized method for determining the topic hyperparameters of a document, and a tool for predicting the topics of new documents with a trained model. It can be accessed using Python [18]. Table 3 compares the generation time and topics with corresponding keywords of articles using both "Mallet LDA" and "Gensim LDA model" [19].
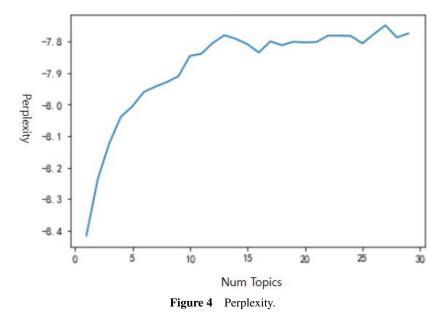
We can conclude that although the analysis speed of "LDA mallet" is slower, the consistency between the generated topics and keywords is better.

## 3.6 Topic Accuracy

In the previous section, we utilized LDA to infer the topics of news samples. However, the model requires the determination of an appropriate number of topics to run. To this end, this section introduces the training of the model and how to determine its parameters for this study. Currently, the most common methods to obtain the optimal number of topics include human judgment-based methods, perplexity, coherence score, and other techniques.

Human judgment-based methods rely on the expertise of specialists to compare the results of LDA under different topic numbers, resulting in personal subjective opinions. If the sample size is small, such methods can perform well. The accuracy of personal subjective analysis is high, but its efficiency is low, making it suitable for analyzing small text sets. However, due to the large number of articles in this study, the method based on manual analysis is inefficient and not suitable.
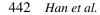
In this study, we do not consider perplexity as a criterion for selecting the optimal number of topics since it is a metric used to evaluate the effectiveness

**Figure 4**   Perplexity.

of probabilistic language models. Despite the general notion that a lower perplexity value indicates a better predictive accuracy of the model [20], differences in results could be due to variations in the data. However, with different data, the correct calculation of perplexity may be obtained. The perplexity values for this model are presented in Figure 4.

A coherence score is a metric used to evaluate the coherence of text or language models. Topic coherence observes the set of words in the topics generated by the modeland confirmation the metric through connection between words similarity scores in each topic [58]. Topic coherence is a numerical value used to measure whether words within a topic are coherent. Topic coherence measures the coherence of a single topic by measuring the semantic similarity between high-scoring words in a topic, helping to improve semantic understanding of that topic. Words are expressed as vectors through their co-occurrence relationships, and semantic similarity is the cosine similarity between word vectors. Higher coherence scores indicate that the text is more easily understood and well-structured, while lower coherence scores indicate that the text is fragmented, unclear, or disjointed [21].

The coherence measure used in this study was the $C_v$ method, which measures the probability of topics' words appearing together in the corpus. The mathematical derivation process of $C_v$ is explained in detail in this
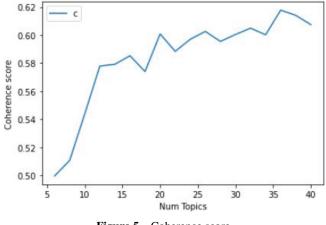
**Figure 5**    Coherence score.

paper, with three equations used to describe it [59].

$$\vec{v}(W') = \left\{ \sum_{w_i \in W'} NPMI(w_i, w_j)^{\gamma} \right\}_{j=1,\ldots,|W|} \tag{1}$$

$$NPMI(w_i, w_j)^{\gamma} = \left( \frac{\log \frac{P(w_i, w_j)+\varepsilon}{P(w_i) \cdot P(w_j)}}{-\log(P(w_i, w_j)) + \varepsilon} \right)^{\gamma} \tag{2}$$

$$\phi_{S_i}(\vec{u}, \vec{w}) = \frac{\sum_{i=1}^{|W|} u_i \cdot w_i}{\|\vec{u}\|_2 \cdot \|\vec{w}\|_2}. \tag{3}$$

This article uses topic coherence to determine the optimal number of topics. The higher the topic consistency value, the better the model. Figure 5 shows that the model has the highest topic consistency when the number of topics is 37.

We used LDAvis to reflect the distribution of topics in the sample [56]. In Figure 6 these circles overlap, and their centers are determined by computed topic distance [54].

## 4 Experiment Results

Given in the tables are the top 10 keywords per topic rank, sorted in order of importance. We sorted seven topics into three main themes according

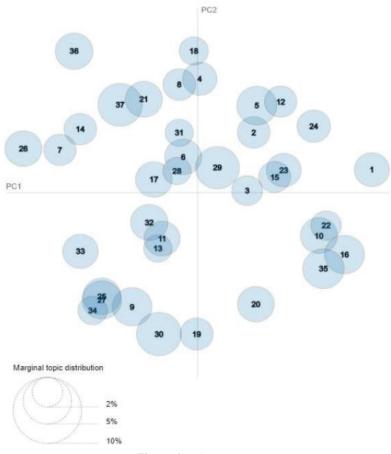Intertopic Distance Map (via multidimensional scaling)
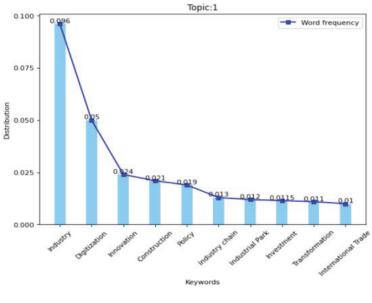


**Figure 6**   Distance map.

to their content about the regrowth factors of the China game industry in the COVID-19 era. We arranged them in descending order of proportion in the sample. The theme of policy mainly revolves around the game industry policy and the current status of the game industry. The theme of industry transformation mainly discusses the development status and direction of Chinese game companies. The theme of society and media presents two development problems: minors being exposed to online game addiction and regulatory measures limiting industrial development. Table 4 gives the topic composition and keywords. The results and semantics of the Chinese–English

**Table 4**    Theme 1: Policy

| Theme 1 | Topic | Keywords | Distribution |
|---|---|---|---|
| Policy | Topic 1 | Industry/digitizaton/innovation/construction/policy/ industry chain/industrial park/investment/transformation/international trade | 7.5% |
| | Topic 2 | Digital/digital economy/digitizing/bigdata/industry/smart/transfor mation/ internet/city/construction | 4.5% |
| | Topic 6 | Billion/increase/year-on-year/revenue/financial report/business/COVID-19/game/report/market | 3.7% |

translation may differ from the actual ones. The Chinese results have been modified according to the semantics.

Theme 1 contains three topics, of which the topic of opinions related to industrial policy is the most significant among the 31 issues (Figure 6), accounting for 7.5% of the sample, as shown in Table 4. Firstly, this topic reflects the Chinese government's concern for the economy during the epidemic. We can see in Figure 7 the Chinese government's supportive measures for the industry, such as the establishment of industrial parks and industrial reform. The second theme reflects the direction of economic policies. The keywords revolve around specific policy guidelines, featuring digitalization, intelligence, and Internet transformation. This information shows that the current high-tech development and digital industry are essential economic means for China's economic development during the COVID-19 period. The third theme has seen much growth, with keywords such as COVID-19, revenue, etc indicating that the game industry was thriving in the COVID-19 period.

Theme 2 also comprises three topics. The most extensively discussed topic is the new direction of the industry resulting from the policy changes in the COVID-19 era and enterprise transformation. The impact of the pandemic has caused businesses to display high interest in emerging technologies such as VR and metaverse, which is reflected in the keyword distribution in Table 5 and the distribution of keywords in the aggregate is shown in Figure 8. Topic 4 revolves around the support of games in healthcare during the COVID-19 period. The final topic, Topic 7, highlights the Chinese game industry's overseas development. Under the theme of corporate transformation, the development ideas and future directions of Chinese game companies in the COVID-19 period are depicted.

**Figure 7**   Continued

**Figure 7**    Theme 1 keyword distribution.

**Table 5**    Theme 2: Industrial transformation

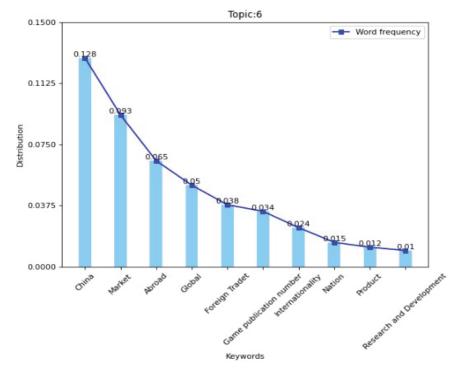| Theme 2 | Topic | Keywords | Distribution |
|---------|-------|----------|--------------|
| Industrial transformation | Topic 3 | Metaverse/user/digitizing/platform/technology/experience/transformation/internet/VR/contents | 3.5% |
| | Topic 4 | Healthcare/product/digital therapeutics/game development/china/co-op/digitizing/software/mental illness/ophthalmology | 3.2% |
| | Topic 7 | China/market/abroad/global/foreign trade/game publication number/internationality/nation/product/research and development | 3.1% |

Theme 3 consists of only one topic, Topic 5, which indicates that social factors also affect the development of the game industry (Table 6). Although this topic comprises only one topic, it cannot be ignored; family education and media participation in this topic is significant (Figure 9), with Topic 5

**Figure 8**    Continued
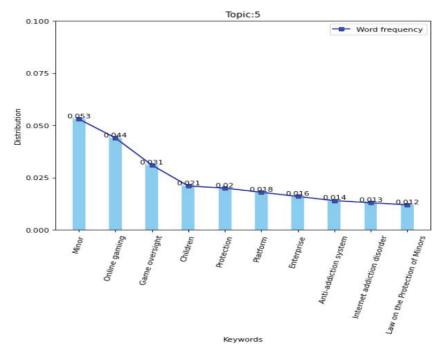
**Figure 8**    Theme 2 keyword distribution.

**Table 6**    Theme 3: community and social support

| Theme 3 | Topic | Keywords | Distribution |
|---|---|---|---|
| Community and social support | Topic 5 | Minor/online game/game oversight/children/protection/platform/enterprise/anti-addiction system/Internet addiction disorder/law on the protection of minors | 6.5% |

accounting for 6.5% of the total sample. The third theme mainly reflects that the game industry in China has been criticized at the social and community levels due to the influence of policies. The emphasis is on the internet addiction of minors and the policy protection of minors. Internet addiction among minors is becoming a significant social problem in China, reflecting social opinions.

This section presents an LDA theme analysis of game industry-related samples, resulting in seven themes that are categorized into three main themes related to the re-growth factors of the China game industry during

**Figure 9**    Theme 3 keywords distribution.

the COVID-19 era. These three main themes are: (1) policy, (2) enterprise transformation, and (3) social and media. In the subsequent discussion, we will examine these themes using numerous articles and papers to demonstrate how they have impacted the development of the Chinese game industry in the COVID-19 era.

## 5  Conclusion and Discussion

This study conducted an LDA analysis to identify the themes of the current state of game enterprise development during the COVID-19 pandemic. The three main themes identified were policy, enterprise transformation, and social and community support.

- The first theme, "policy," highlights the positive impact of COVID-19 on the gaming industry and how government policies promoting the digital economy have attracted investors and stimulated future development. Despite the absence of rating entities in China, active government policies are crucial to the industry's growth.

- The second theme, "industrial transformation," highlights how the game industry has leveraged technology to enable innovation and expansion into more social areas. For instance, cross-border collaboration between game companies and the medical industry was used to develop SaMD content as an example.
- The third theme, "community and social support," addresses the challenges that the Chinese e-sports industry is facing due to the minor protection policy, as well as how the government's stance towards the game industry affects its growth. While the policy poses limitations, it also motivates companies to enhance their game products and facilitate the development of high-quality games.

## 5.1 Policy

The first theme consists of Topic 1 and Topic 2, named "policy". In this section, we will examine the factors that led to the emergence of the theme and analyze how the COVID-19 pandemic has affected the development of the game industry and the impact of government policies on the industry. Our research focuses on the role of the government in shaping the development of the game industry and includes topics such as finance, research and development, infrastructure development, and overall industry direction.

In this theme, it is shown that the COVID-19 pandemic has led to a surge in video game engagement. The impact of the pandemic on the game industry has resulted in game publishers acquiring a substantial player base, as gaming became a popular stay-at-home activity during the crisis. This is supported by a constant flow of news reporting the increased prevalence of gaming during the pandemic.

Before the COVID-19 pandemic, the Chinese government had already placed emphasis on the boosting of the digital economy through measures such as the construction of infrastructure and industrial parks [22]. The government supported industries such as home entertainment, artificial intelligence, short videos, video games, and e-sports, advocating for the digital economy [24]. These high-tech industries, offering services such as remote and online options, proved to be the least affected and most valuable during the COVID-19 pandemic [25]. During the outbreak, China, like other countries, implemented economic control measures and made policy changes to combat the pandemic's effects [23].

As a provider of digital content, the gaming industry has certainly benefited from the rise of the digital economy [26]. For those in quarantine,

online entertainment and socialization opportunities made possible by technological advancements have made staying at home less monotonous [27]. The gaming industry has therefore taken on the responsibility of maintaining social stability to a certain extent. The game industry offers entertainment and socialization options during a pandemic, and the promotion of digital economics by the government furthers the development of the industry. The positive external [28] and all the necessary conditions for growth are present, making the expansion of the industry a logical outcome.

A large number of policy keywords appear in the theme because the government has a massive impact on the industry in formulating policies. It is normal for the game industry to experience development setbacks due to lack of or excessive government intervention. The Korean government's hostile policy to introducing domestic games has led to a sluggish market, making it more difficult for developing small enterprises to obtain substantial government policy support [29]. China has also restricted game distribution due to game review issues, and investors have lost confidence in the industry's future prospects; investment in future projects has decreased [30]. Due to the influence of policy factors, China's e-sports industry has been battered, and the policy on the protection of minors has had a specific impact on the selection of excellent e-sports players [31]. Note that this will also affect the future layout of China's game industry.

By contrast, Japan companies have benefited from Japan's loose policy management and unique corporate management mechanism. Japanese game companies once surpassed the United States in the game development market [32]. The Chinese government supports domestic game development and proposes a strategy to enter overseas markets [33] to promote the globalization of China's game industry. In 2021, the actual sales revenue of China's self-developed games overseas market will reached US$18.013 billion, an increase of US$2.563 billion or 16.59% year-on-year [34]. The successful development of Chinese companies in the overseas game market in recent years is inseparable from the government's policy promotion. The government's subdivision management of the industry can enable enterprises to find suitable corporate positioning in the game market with high competition pressure. To sum up, COVID-19 has played a catalytic role in the development of games, and the social background and policy support have promoted the development of the industry. However, in the game industry's development direction, China must carefully consider policy limits [57].

## 5.2 Industrial Transformation

The theme 2 is called "industrial transformation." This section consists of two topics. They are Topic 3 and Topic 4. The theme reveals the future development direction and prospects of Chinese games.

From the perspective of game technology, the game ecosystem refers to a technology cluster that realizes large-scale applications in video games and aims to enrich and improve interactive experience [35]. It is mainly virtual reality, game engine, rendering, etc. Those game companies with high-tech technologies began to innovate in response to social needs. For example, many smart city platforms based on game engines are used [36]. Coupled with the government's support for high-tech industries, China's Internet companies have begun to develop products such as metaverse and XR. In 2022, Tencent expects to widely apply game technology development content in cultural tourism, AI, industry, and other fields in the future, and policy support will benefit the company's long-term development [37]. In terms of digital healthcare, combined with the current situation of the large-scale isolation policy, many chronic patients do not have adequate medical protection at home. Hence the use of telemedicine, remote patient monitoring, and artificial intelligence in healthcare continues to grow in the age of COVID-19. And this cooperative and innovative operating model still has prospects. In the post-epidemic era, patients will also choose to continue to receive services because of their convenience. Moreover, the need for adjunctive treatment of chronic and psychiatric diseases for individual patients in hospitals and at home has enabled the development of digital health [38].

The gaming industry has also fostered the development of telemedicine. For example, in the medical field, there are experiments in training ADHD patients to strengthen their attention through game content [39]. Not only that, there are already related examples of SaMD (software as a medical device) in China. An eye protection game has been developed by an ophthalmology institution in China and a Chinese game company. The project introduced experts from the external medical industry, and the rehabilitation treatment content for children with strabismus and amblyopia has effectively improved the degree of cooperation of patients with treatment. Currently, the therapeutic content of SaMD games is developed by medical institutions cooperating with game companies. Participating in developing SaMD game content is an attempt by Chinese game developers to transform their businesses in the era of COVID-19. However, game companies should pay attention to the development direction compared with game content. SaMD does not focus

on entertainment factors but requires corresponding scientific arguments, research results, and clinical applications [41].

In this section, according to the analysis of the theme chart, it is concluded that the current COVID-19 era has gradually become a new industrial transformation opportunity for Chinese game companies. While adapting to the environment and policies of the times, the Chinese game industry has begun to provide high-tech technical support for traditional industries to help enterprises improve production efficiency. Enterprise transformation is one of the critical goals of the current industry.

## 5.3 Community and Social Support

The third theme is "community and social support." It consists of topic 5. In this theme, we reflect that Chinese society has great opinions on the game industry. This research points out the problems exposed in the current Chinese game development.

In the era of COVID-19, due to the continued economic downturn and the impact of home isolation, most people began to choose to use the Internet as a new way of entertainment. With the epidemic, the Chinese game industry ushered in the second explosive growth (the first explosive growth was in 2018 when mobile games led to the update and explosive growth of the industry [42]). Therefore, minors have been exposed to online games for a longer time. It caused more social problems: children addicted to games leads to academic regression, teenagers addicted to online games often run away from home, prolonged exposure to screens leads to the deepening of myopia, etc. [43]. In China, parents, academics, and educators have expressed concern about the rapid development of the e-sports industry. The conservative opinion generally holds that video games are harmful to teenagers. As a result, society has gradually formed a negative public opinion of the games [44]. Internet game addiction gradually becomes the reason for the failure of adolescents' education [45]. In this case, the pressure of public opinion means games are often brought up as a bad influence factor in the family and education. The positive effects of games are challenging to publicize [46]. Due to COVID-19, some students' outings and entertainment activities have been affected. When students are in isolation, they often choose to play games to relieve their inner pressure. Some students are addicted to online games and waste their studies. Due to the epidemic's impact, students risk developing Internet addiction due to the long-term use of computers when receiving online education [47]. Needless to say, public opinion and the education circle in China have a strong negative attitude toward the game industry.

From theme 3, we can see that the Chinese government's policy on the management of the game industry has affected the development of China's e-sports industry. In 2022, the State Press and Publication Administration of China issued a "Notice on Further Strict Management to Effectively Prevent Minors from Indulging in Online Games." [48]. The main content is that the time played and type of online games used by minors are strictly restricted. In addition, the Chinese government has issued policies to prevent minors from becoming addicted to the Internet, e.g minors play games for 3 hours a week, and online games are subject to a real-name system. Such measures have affected the selection of young e-sports players.

The growth of the game industry due to the new crown epidemic has not been affected by these policy measures in the short term, and minors are not the main consumer groups in the eyes of game manufacturers. However, from a long-term perspective, restrictions on the game industry will affect the absorption of potential game users. Although the minor user group cannot generate much revenue, it has contributed a lot of online users to major games. After the policy was introduced, online game data inevitably declined to a certain extent. Moreover, the industrial layout of the capital market and game companies will also be affected. From the perspective of the industry's sustainable development, this is undoubtedly a blow to the game industry.

## 6  Suggestions and Limitations

### 6.1  Suggestions for the Future Development of the Game Industry

Based on the analysis, the rapid growth of the gaming industry during the COVID-19 pandemic can be attributed to the shift in entertainment habits caused by the pandemic and supportive policies in China. The industry has evolved, now prioritizing technology integration with traditional industries. However, the impact of COVID-19 has also magnified existing challenges in the industry. Thus, the following three recommendations are proposed based on these findings.

- Game companies should strive for diversification and enhance the technical diversity in their products and services, and collaborate with other industries to offer game development technology to aid traditional industries in their digital transformation. One example could be the use of digital twin technology [49], which acts as a virtual-based digital production aid. Finally game companies should proactively fulfill their

social responsibilities and provide technical support to other industries in the post-pandemic era, to meet the digital modernization needs of traditional industries. For instance, online AR shopping can increase the sales of low-value-added products, which could lead to technological reform of the manufacturing industry and enhance the competitiveness of businesses [50].

- The government should offer more specific policy support for the growth of the game industry. The success of businesses is closely tied to government policies, and as such, the government should provide ample legal protection for market players, inspire game companies to overcome challenges, and use regulations to enhance the development of game companies and businesses, enhancing the digital economy market and generating job opportunities [51]. Currently, setting up a comprehensive game rating and management system is a pressing requirement for the advancement of China's game industry.

- At the social level, positive public perception can greatly benefit the growth of the game industry [52]. Positive media coverage showcasing the positive impact of the industry on society can enhance its image and increase public trust, leading to more support. Developing e-sports players requires collaboration between society and families [53]. Game companies should also adopt better practices in managing minors' game time and promoting healthy habits. Improving the company's social credibility by enhancing social recognition can also help improve the industry's overall image [54].

## 6.2 Limitations and Future Research

However, we acknowledge the limitations of this study. The LDA analysis may be influenced by personal bias factors, which could affect the choice of topics and keywords. Subjective evaluation of the analysis may also impact the resulting topics. Moreover, the diversity in the amount of text in the data may lead to an uneven distribution of results, constrained by the analytical model.

Additionally, if related articles contain words in the same context that change due to temporal relationships, they may not be included in the same dataset, potentially leading to data inconsistency or bias. Thus, in the following research stage, we plan to enhance the model by utilizing the DTM model [55] to analyze news texts. We aim to improve the accuracy of synonym analysis results over time and analyze the text at a temporal level.

We also intend to examine the public opinion of news media and analyze the public's views on news topics according to the media's comments.

This study suggests that when analyzing success metrics in the gaming industry, analytical methods should not solely focus on technical aspects but should also take into consideration the industry's complexity and breadth. Examining the industry's evolution through data analysis can provide a fresh perspective and help improve its public image. Future research in the gaming industry based on LDA analysis can support the formulation of more objective and rational policies and development strategies.

## Acknowledgments

## References

[1] China Game Publishers Association Publications Committee's (GPC), "China Gaming Industry Report (2021)," *White Paper*, 2022.

[2] D. Vitale, P. Ferragina, U. Scaiella. "Classification of short texts by deploying topical annotations," *European Conference on Information Retrieval*, Berlin, 2012, pp. 376–387.

[3] H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, L. Zhao, "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimedia Tools and Applications*, vol. 78, pp. 15169–15211, Nov 2019

[4] B. Wang, S. B. Liu, K. Ding, Z. Y. Liu, "Patent content analysis method based on LDA topic model," *Research Management*, vol. 36, pp. 111–117, Mar 2015.

[5] L. Wang, L. X. Zou, X. W. Liu, "Visualizing document correlation based on LDA model," *Data Analysis and Knowledge Discovery*, vol. 2, pp. 98–106, Mar 2018.

[6] Y. P. Wang, "Research progress of scientific and technical literature topic detection and evolution based on topic model in China," *Library and Information Service*, vol. 60, pp. 130–137, 2016

[7] L. Guo, C. J. Vargo, Z. Pan, W. Ding, P. Ishwar, et al., "Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling,"

*Journalism & Mass Communication Quarterly*, vol. 93, pp. 332–359, Feb 2021.

[8] Q. Liu, Z. Zheng, J. Zheng, "Health communication through news media during the early stage of the COVID-19 outbreak in China: digital topic modeling approach," *Journal of Medical Internet Research*, vol. 22, e19118, Apr 2020.

[9] S. Han and T. Kim, "News Big Data analysis of 'Metaverse' using topic modeling analysis," *Journal of Digital Contents Society*, vol. 22, pp. 1091–1099, Jul 2021.

[10] M. Kumar and J. Ng, "Using text mining and topic modelling to understand success and growth factors in global renewable energy projects," *Renewable Energy Focus*, vol. 42, pp. 211–220, 2022.

[11] V. G. Nair, *Getting Started with Beautiful Soup.* Packt Publishing Ltd, 2014.

[12] J. H. Liu and L. L. Yu, "Survey on topic focused Web crawler," *Application Research of Computers*, pp. 26–29 and 47, Oct 2007.

[13] R. Mihalcea and P. Tarau, "Textrank: Bringing order into text," In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*, Jul 2004, pp. 404–411.

[14] M. Y. Day and M. Y. Lee, "Deep learning for financial sentiment analysis on finance news providers," In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Aug 2016, pp. 1127–1134.

[15] M. D. Blei, Y. A. Ng, Y. A. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, pp. 993–1022, Jan 2003.

[16] J. Han, J. Pei, H. Tong, *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2022.

[17] D. M. Blei and M. I. Jordan, "Variational inference for Dirichlet process mixtures," *Bayesian Analysis*, vol. 1, pp. 121–143, Jan 2006.

[18] S. Graham, S. Weingart, I. Milligan, *Getting Started with Topic Modeling and MALLET*. Programming Historian, 2012.

[19] Y. S. Chen, L. H. Chen, Y. Takama, "Proposal of LDA-based sentiment visualization of hotel reviews," In *2015 IEEE International Conference on Data Mining Workshop (ICDMW)*, Nov 2015, pp. 687–693.

[20] S. F. Chen and J. Goodman, "An empirical study of smoothing techniques for language modeling," *Computer Speech & Language*, vol. 13, pp. 359–394, Apr 1999.

[21] D. Mimno, H. Wallach, E. Talley, M. Leenders, A. McCallum, "Optimizing semantic coherence in topic models," In *Proceedings of the 2011*

*Conference on Empirical Methods in Natural Language Processing*, July 2011, pp. 262–272.

[22] China Academy of Information and Communications Technology (CAICT), "2022 China Digital Economy Development Report," *White Paper*, Jul 2022.

[23] W. J. Chen, L. Y. Yang, "Global digital economy governance: Elements, mechanisms and breakthroughs," *International Economic Review*, vol. 2, pp. 60–87, 2022.

[24] R. Bukht and R. Heeks, "Defining, conceptualising and measuring the digital economy," *Development Informatics Working Paper*, no. 68, Aug 2017.

[25] S. A. Zahra, "International entrepreneurship in the post Covid world," *Journal of World Business*, vol. 56, 101143, Jan 2021.

[26] F. E. Alvarez, D. Argente, F. Lippi, "A simple planning problem for covid-19 lockdown (no. w26981)," *National Bureau of Economic Research*, Apr 2020.

[27] A. Khalil, P. von Dadelszen, E. Kalafat, M. Sebghati, S. Ladhani, A. Ugwumadu, L. Magee, "Change in obstetric attendance and activities during the COVID-19 pandemic," *The Lancet Infectious Diseases*, vol. 21, p. e115, 2021.

[28] G. Nimrod and D. A. Kleiber, "Reconsidering change and continuity in later life: Toward an innovation theory of successful aging," *The International Journal of Aging and Human Development*, vol. 65, pp. 1–22, 2007.

[29] J. H. Wi, "Korea government policy evaluation on game industry-focused on industrial policies on game regulation and promotion by the ministry of culture, sports and tourism, Korea," *Journal of Korea Game Society*, vol. 19, pp. 15–24, 2019.

[30] W. J. Li, "Online game marketing opportunity survey," *Economic Outlook the Bohai Sea*, vol. 3, pp. 107–109, 2022.

[31] S. Mao, "The power of idol? The impact of age regulation on internet addiction," In *2022 8th International Conference on Humanities and Social Science Research (ICHSSR 2022)*, Jun 2022, Atlantis Press, pp. 2066–2071.

[32] C. Storz, F. Riboldazzi, M. John, "Mobility and innovation: A cross-country comparison in the video games industry," *Research Policy,* vol. 44, pp. 121–137, 2015.

[33] P. L. Zeng, Y. X. Deng, "From 'communicative medium' to 'innovative subject': A new way to bring Chinese video games to the world," *Journalism Research*, vol. 5, pp. 94–104, 2022.

[34] China Game Publishers Association Publications Committee's (GPC), "China Gaming Industry Report (2018)," *White Paper*, 2019.

[35] P. Zackariasson and T. L. Wilson (eds.), *The Video Game Industry: Formation, Present State, and Future*. Routledge, 2012.

[36] S. Beckett, "Smart city digital twins, 3D modeling and visualization tools, and spatial cognition algorithms in artificial intelligence-based urban design and planning," *Geopolitics, History, and International Relations*, vol. 14, pp. 123–138, Jan 2022.

[37] Tencent, "Tencent Announces 2022 First Quarter Results," *White Paper*, 2022.

[38] Z. Haider, B. Aweid, P. Subramanian, F. Iranpour, "Telemedicine in orthopaedics during COVID-19 and beyond: a systematic review," *Journal of Telemedicine and Telecare*, vol. 28, pp. 391–403, 2022.

[39] S. H. Kollins, D. J. DeLoss, E. Cañadas, et al, "A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): a randomised controlled trial," *Lancet Digital Health.* vol. 2, pp. 168–178, Apr 2020.

[40] M. Zhang, H. Su, J. Wen, "Analysis and mining of internet public opinion based on LDA subject classification," *Journal of Web Engineering*, vol. 20, pp. 2457–2472. Nov 2021.

[41] J. Shuren and P. M. Doraiswamy, "Digital therapeutics for MCI and Alzheimer's disease: A regulatory perspective—highlights from the clinical trials on Alzheimer's disease conference (CTAD)," *The Journal of Prevention of Alzheimer's Disease*, pp. 1–5. 2022.

[42] China Game Publishers Association Publications Committee's (GPC), "2018 China Game Industry report," *White Paper*, 2018.

[43] J. X. Zhang and X. Y. Hu, "Psychodynamic interpretation of online game addiction and educational implications," *Chinese Journal of Special Education*, pp. 64–69, Jan 2016.

[44] G. L. Fan, "Research on the factors affecting the industrialization of e-sports in my country and its system model," *Market Modernization*, vol. 13, pp. 353–354, 2007.

[45] L. Cheng, "The psychological mechanism and treatment of internet addiction in adolescents," *Contemporary Education Sciences*, vol. 23, pp. 42–44, 2003.

[46] Y. Cao and H. M. Wang, "Negative effects of role-playing online games on adolescent participants and normative suggestions," *Theory Research*, vol.18, pp. 70–73, 2008.

[47] H. Y. Wan, J. Q. Yu, N. Q. Yan, "Relationships between learning burnout and Internet addiction among undergraduates during the COVID-19 pneumonia: Mediating effect of career adaptability," *China Journal of Health Psychology*, vol. 5, pp. 695–701, 2021.

[48] National Press and Publication Administration, "Notice of the National Press and Publication Administration on Preventing Minors from Indulging in Online Games (in Chinese)," NPPA.gov.cn modified Nov 2019. accessed Sep. 19. 2021. http://www.nppa.gov.cn/nppa/contents/279/74483.shtml

[49] R. Z. Lin, X. J. Wang, Y. P. Feng, J. Q. Zhao, "Experimental system of intelligent assembly manipulator based on digital twin," *Research and Exploration in Laboratory*, vol. 12, pp. 83–88, 2019.

[50] Y. Z. Yang, F. X. Wen, S. Y. Yang, Y. Zheng, S. Y. Tang, "Research on new online shopping experience—AR+VR shopping innovation development research," *Henan Science and Technology*. vol. 14, pp. 24–26, 2020.

[51] Y. J. Huang, "Comparative Study: How Metaverse Connect with China Laws," Available at SSRN 3955900, Oct 2022.

[52] Y. Jiang and J. J. Shang, "Should children be allowed to play games?," *People's Education*, vol. 8, pp. 24–28, 2019.

[53] Y. Meng-Lewis, D. Wong, Y. Zhao, G. Lewis, "Understanding complexity and dynamics in the career development of eSports athletes," *Sport Management Review*, vol. 25, pp. 106–133, 2022.

[54] L. Xiao, L. Henderson, Y. Yang, P. Newall, "Gaming the system: Suboptimal compliance with loot box probability disclosure regulations in China," *Behavioural Public Policy*, pp. 1–27, 2021.

[55] D. M. Blei and J. D. Lafferty, "Dynamic topic models," In *Proceedings of the 23rd International Conference on Machine Learning*, pp. 113–120, Jun 2006.

[56] C. Sievent and K. Shirley, "LDAvis: A method for visualizing and interpreting topics," In Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, Jun 2014, pp. 63–70.

[57] M. Colder Carras, V. Stavropoulos, F. Motti-Stefanidi, A. Labrique, M. D. Griffiths, "Draconian policy measures are unlikely to prevent disordered gaming," *Journal of Behavioral Addictions*, pp. 849–853, 2021.

[58] P. Tijare and P. J. Rani, "Exploring popular topic models," *Journal of Physics: Conference Series,* vol. 1706, p. 012171, Dec 2020.

[59] S. Syed and M. Spruit, "Full-text or abstract? Examining topic coherence scores using latent Dirichlet allocation," In *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pp. 165–174, Oct 2017.

## Biographies



**Yi-Qian Han** is currently a graduate student in the IT Convergence Engineering at Gachon University, Republic of Korea. His research interests include 3D content design, text mining, and evaluation methods of game industry.



**Won-Jun Jeong** received his Masters in Game Engineering from College of IT Convergence Engineering of Gachon University in 2020. He is currently a graduate student in the IT Convergence Engineering at Gachon University, Republic of Korea. His research interests include VR motion sickness, evaluation methods of XR content and digital therapeutics.

**Gi-Sung Oh** received his Masters in Game Engineering from College of IT Convergence Engineering of Gachon University in 2020. He is currently a graduate student in the IT Convergence Engineering at Gachon University, Republic of Korea. His research interests include VR motion sickness, digital therapeutics, game design and evaluation methods of XR content.

**Seok Hee Oh** is a Program Director at R&D Division, Korea Creative Contents Agency (KOCCA) in Korea. He was as an associate professor in the Department of Computer Engineering at Gachon University, Seongnam from 2016 to 2021. Oh received his Ph.D. degree in 2016 from the Department of IT Convergence Engineering, Gachon University Graduate School. Oh's current and previous research interests include virtual reality, HCI, UX, and game design. He is a member of the Korea Game Society and the Gamification Forum.

**Taeg-Keun Whangbo (Member, IEEE)** received his M.Sc. degree in computer science from the City University of New York in 1988, and his Ph.D. degree in computer science from the Stevens Institute of Technology in 1995. He is currently a Professor with the Department of Computer Science, Gachon University, South Korea. Before he joined Gachon University, he was a Software Developer for Q-Systems, which is located in New Jersey, from 1988 to 1993. He was also a Researcher with Samsung Electronics, from 2005 to 2007. His research interests include artificial intelligence, computer vision, healthcare and AR/VR.