What Causes the Virtual Agglomeration of Creative Industries?

Xu Chen 1, Chunhong Liu 1, Yao Jiang 2,* and Changchun Gao 1

Abstract: The agglomeration paradigm for creative industries has fundamentally changed under the digital economy, giving rise to a new form of virtual agglomeration within these industries. This study explores the causes of this virtual agglomeration. We collected online Chinese news texts related to the virtual agglomeration of the creative industry, used text mining to identify nine factors affecting its formation, and refined the internal and external factors for an analytical framework based on the PEST (political, economic, social, technological) and value-chain models. We then combined the relevant literature and the creative industry’s development practices, analyzed the mechanism of each driving factor, and constructed a driving-force model for the creative industry’s virtual agglomeration. The external driving factors were government policy planning, the digital economic environment, emerging consumer demand, and the application of innovative technology; the internal factors were the digitalization of cultural resources, flexible manufacturing, digital marketing and promotion, online interactive services, and virtual platform facilities. Each factor was found to contribute to virtual agglomeration through different internal mechanisms. This study’s findings have theoretical and practical value for cultivating the modes of virtual agglomeration within creative industries.

Keywords: virtual agglomeration; creative industries; driving factors; driving mechanisms; text mining

1. Introduction

With the development of digital technologies, virtual agglomeration has provided a new avenue for the geographic agglomeration of creative industries. Various creative industry enterprises, institutions, and users rely on digital technologies to conveniently share creative ideas, develop creative resources, and design products in virtual spaces. For example, Yipin Witkey, a creative service platform in China, gathers creative service providers in fields such as design, development, planning, and marketing, and it offers more than 300 kinds of creative services. The platform has more than 22 million registered users and has supported the creative needs of numerous enterprises [1]. Meanwhile, the US-based creative gaming platform Steam brings together thousands of game creators and publishers; it offers about 30,000 games played by over 90 million users around the world [2]. These types of virtual agglomeration practices extend beyond traditional concepts of geographical agglomeration; we call this phenomenon “the virtual agglomeration of creative industries” [3,4].

Research on virtual agglomeration began with the EU’s program for the Competitiveness of Enterprises and SMEs (COSME) [5,6]. Virtual agglomeration is also sometimes referred to as “e-clusters” or “virtual clusters” [7–9]. Brown and Lockett [10] noted that the emergence of the Internet provided firms with platforms on which they could form interconnected digital enterprise communities. Moreover, Mason et al. [11] noted that the development of information and communications technology (ICT) has connected organizations that were previously scattered across regions, thereby forming e-clusters in the information space. Based on the literature [12–14], we define virtual agglomeration...
as a phenomenon in which enterprises, institutions, and users within an industry are concentrated in virtual spaces via the connectivity afforded by digital technology. Unlike geographical agglomeration, which concerns geographical proximity in a specific region [15], virtual agglomeration is characterized by digital proximity among multiple subjects in virtual spaces [16,17]. Virtual agglomeration can overcome spatiotemporal constraints and expand the functions of sensing, linking, data, and computation in all aspects of a product, from production to consumption, and it supports the rapid transmission of information between enterprises and consumers [18]. Correspondingly, the virtual agglomeration of creative industries has changed creative design, production, marketing, and services. These changes significantly improve operational efficiency, thus promoting sustainable development in these creative industries [19]. Thus, a question arises: How did the phenomenon of the creative industry’s virtual agglomeration emerge?

Various studies have investigated the driving factors of virtual agglomeration in different creative industries, such as animation [20], information services [21], digital publishing [22], fashion [23], and online video [24]. These studies have also extracted relevant factors, including digital technology, government, demand, capital, and network platform factors. They have not, however, systematically explored the specific mechanisms of each factor nor have they built relevant theoretical models. This study therefore aimed to explain how the virtual agglomeration of creative industries formed. Our specific goals were to: (1) identify the driving internal and external factors of the creative industry’s virtual agglomeration and analyze them using relevant theoretical models, and (2) explore the mechanisms for each driving factor and build a unified driving-force model.

Our research proceeded as follows. First, we used Octoparse’s web crawler technology to collect Chinese online news reports related to the creative industry’s virtual agglomeration. Second, a Python-based text-mining algorithm was used to mine the resulting news corpus to identify the key driving factors of the creative industry’s virtual agglomeration. Third, the identified driving factors were refined into internal and external factors based on PEST (political, economic, social, technological) and value-chain models. Each driving factor and mechanism was then analyzed based on the relevant literature and industry practices, thus forming a driving-force model of the creative industry’s virtual agglomeration. The identified external driving factors were government policy planning, the digital economic environment, emerging consumer demand, and application of innovative technology; the internal factors were the digitalization of cultural resources, flexible manufacturing, digital marketing and promotion, online interactive services, and virtual platform facilities. Each factor was found to contribute to the creative industry’s virtual agglomeration through different internal mechanisms.

This study makes two main contributions. First, we focused on the causes of the creative industry’s virtual agglomeration through text mining and analysis. We clarified the internal and external driving factors and mechanisms and built a related driving-force model. Second, we used big data web crawler technology to collect Chinese online news content and, on that basis, identified the driving factors of the creative industry’s virtual agglomeration. Such methods served to ensure the scientific basis of this research, verify its accuracy, and broaden its application to related fields.

The rest of this paper is organized as follows. Section 2 reviews the literature on the virtual agglomeration of creative industries. Section 3 introduces the research design and methods. Section 4 describes the data collection process and analysis results. Section 5 analyzes the driving factors and mechanisms, while Section 6 presents the main conclusions and implications.

2. Literature Review

Geographic agglomeration has long been a hot topic in research on creative industries [25–32]. The geographic agglomeration of the creative industry is a spatial organizational system in which creative industry entities are codependent because of their commonalities and are therefore spatially close to each other [33]. Geographic agglomer-
tion based on a specific location reduces spatial distance in enterprise collaboration, thereby reducing production and transaction costs and increasing production efficiency [34]. Such agglomeration requires enterprises to move into the cluster, and the benefits they obtain can outweigh the costs [14]. Creative industries, however, have difficulty overcoming the boundary constraints of geographic space and providing benefits for those outside the cluster location [35]. Moreover, the geographic agglomeration of the creative industry exists in a relatively closed-loop geographical space [36], and the number and scale of enterprises it can accommodate are limited [14]. This limitation weakens the ability of industry clusters to form more efficient organizational systems.

Virtual agglomeration transcends geographic boundaries [37] and provides a new direction for the development of creative industry agglomeration. It is characterized by data capitalization, online information, demand fragmentation, transaction ubiquity, production flexibility, platform “mega-ization,” and full chain integration [38]. Such characteristics are consistent with basic creative industry attributes, such as production digitalization, communication networking, consumption informatization, and innovation globalization [35,39]. The COSME project initially characterized virtual agglomeration as a collection of companies that can mobilize their respective capabilities and operate a virtual enterprise, such that market opportunities are shared among all member companies [5]. Later, Passiante and Secundo [13] defined a “virtual cluster” as a community in which suppliers, distributors, service providers, and customers can cooperate or compete based on technological business networks. Wu and Li [40] suggested that virtual agglomeration can not only incorporate geographical agglomeration but also integrate resources across regions and industries, thus accelerating enterprise innovation. Virtual agglomeration has thus become an important mode of industrial development in the digital economy. Wei and Wu [41] proposed that the virtual clustering of creative industries is embedded in various relationships—such as between production and consumption, global and local, and online and offline—and that connectivity and flexibility are essential indicators for measuring its competitiveness.

As the main carriers of virtual agglomeration, network platforms are characterized by openness and inclusivity [42] and facilitate low-cost, high-efficiency connections between numerous subjects, such as creative product/service providers, creative consumers, and related institutions. Based on intelligent algorithms, perfect rules, and humanized service functions of creative network platforms, participants can also realize precise information matching, network collaboration, and value trading [3]. This emergence of virtual agglomeration in creative industries thus creates a need for systematic research on the causes of its formation.

Previous research has focused on specific creative industries, using qualitative methods to summarize the factors affecting the formation of virtual agglomeration. Studying China’s animation industry, Du et al. [20] classified the driving factors of virtual agglomeration into three levels: social environment (market demand, network infrastructure, industrial policy); industry (internal organization, institutional construction, cultural construction, network platform, brand effect); and enterprise (information system construction, scale, ability to attract capital, creative ability, learning-conversion ability, technological ability). Wang [21], meanwhile, found that the internal factors affecting virtual agglomeration in the information service industry were survival pressure, interest attraction, user demand, and economic scale; the external driving factors were technical, government, and credit guarantees. Studying the fashion industry, Crewe [23] found that digitally mediated communication technologies significantly influenced virtual agglomeration. Finally, Xiao [22] studied virtual agglomeration in digital publishing and found that it relies on the dual elements of technology and capital, which include capabilities such as digital technology R&D, transformation of technological achievements, collaborative innovation, and capital operation.

Given the differences in the attributes of different creative industries, no consistent findings have emerged regarding the causes of virtual agglomeration. In addition, few studies have systematically developed theoretical frameworks to study the mechanisms of these
driving factors. Chen et al. [3] only proposed the mechanisms of network coordination, freedom of participation and guarantee of trust based on interviews and questionnaires of creative industries, but they did not analyze these mechanisms in depth. Because the creative industry’s virtual agglomeration is a new type of relationship formed in a virtual space, existing research conclusions regarding geographic agglomeration cannot be used to explain its formation. As such, there is a need to establish a new theoretical framework to analyze its driving factors and mechanisms.

Unlike traditional agglomeration, virtual agglomeration is built on network platforms using big data [43], which has benefits but also introduces high complexity [44]. Qualitative methods are not entirely suitable for this research. With the ongoing advancements in big-data-related research, quantitative methods based on big data web crawling and text mining can provide excellent solutions [45,46], and researchers are increasingly obtaining results that better reflect reality using big data text mining methods [47–49]. Taking the Shenzhen Dafen Oil Painting Village as an example, Yuan et al. [50] studied the convergence between creative industry parks and tourism using ROST Content Mining 6 and performing an analysis based on the network text. Applying text analysis to 100,000 tweets, Casadei et al. [51] investigated the interlinked relationships between the fashion industry and locations in four major cities (London, New York, Milan, and Paris). In light of such research, the present study used big-data-related methods to identify the driving factors of the virtual agglomeration of creative industries.

The existing research on the creative industry’s virtual agglomeration is fragmented. The main driving factors and mechanisms have not been sufficiently explored and related theoretical frameworks have not been built. Therefore, this study used big data analysis to identify these external and internal driving factors and used PEST and value-chain models to build an analytical framework. The PEST model is a classic tool used for analyzing the external macroenvironment [52]. Lee and Rim [53], for example, used a PEST-SWOT model to analyze the external and internal factors affecting human resource development systems in Korea’s software industry. Xie et al. [54], meanwhile, used PEST to refine the external factors affecting the entrepreneurial ecosystem of the online cultural industry. Porter introduced the value-chain model in 1985 to analyze the internal industrial environment; this model assumes that an enterprise’s production and operation activities create value, and these different but interrelated activities form a dynamic value-creation process called a value chain [55]. The value chain for the creative industry refers to the organic whole constituted by each value-added link among creative industries [56]. Horng et al. [57] used a value-chain model to analyze the value-creation processes of Taiwan’s cultural and creative industries. Madudová [58], meanwhile, used a value-chain model to evaluate industry participants, value-creation processes, supporting environments, and stakeholder relationships in the creative industries of advertising, architecture, and design.

In light of the above, this study analyzed the causes of the creative industry’s virtual agglomeration using big data web-crawler technology and text mining to identify the key factors. These were refined into internal and external factors based on PEST and value-chain models, and the mechanism of each factor was analyzed. Finally, a unified driving-force model of the creative industry’s virtual agglomeration was built to provide a theoretical basis for guiding creative industry development.

3. Research Design and Methods

3.1. Research Design

News is a record of noteworthy events [59]. It includes not only events but also details about participants, time, place, causes, and processes. It is characterized by newness, importance, proximity, significance, and level of interest [60]. As an extension of traditional news, online news has quickly adopted information technology, featuring fast, multifaceted, multichannel, multimedia, and interactive formats [61]. This study therefore used Chinese online news materials to investigate the virtual agglomeration of creative industries. The specific process is described below.
First, Octoparse’s web-crawler technology was used to collect Chinese online news reports related to the creative industry’s virtual agglomeration to form a news corpus.

Second, Python-based text mining was performed on the collected data, including text preprocessing, feature-word extraction, text vectorization, and text clustering. This helped identify the key factors driving the creative industry’s virtual agglomeration.

Third, the identified factors were refined into internal and external factors based on PEST and value-chain models. Each mechanism was analyzed based on the related literature and industry practices. In this way, a driving-force model of the creative industry’s virtual agglomeration was built. Figure 1 shows the research framework.

Figure 1. Research framework.

3.2. Research Methods

Feldman and Dagan proposed the concept of text mining in 1995 [62]. It involves obtaining user interest or useful patterns from unstructured text information [63]. We used text mining to process and analyze the online news corpus; the specific steps are as follows.

The first step is text preprocessing. The complexity of the collected information made text processing and mining difficult. It was therefore necessary to preprocess the text to normalize the data. Preprocessing included text regularization, Chinese word segmentation, part-of-speech tagging, and stop-word removal [64].

The second step is feature-word extraction. Feature-word extraction involves extracting representative and informative phrases from preprocessed text. Using specific algorithms can make this process more efficient and accurate and avoid the disruptive effects of dimensionality in text clustering [65]. Common feature-word extraction methods include the Term Frequency–Inverse Document Frequency (TF-IDF) algorithm, the Text Rank algorithm, and the Topic Model algorithm. Among them, TF-IDF is the most widely used and effective method. The principle of TF-IDF is that the most meaningful feature items for distinguishing texts should be words that appear frequently in the text but infrequently in other texts. This approach extracts feature words based on word-frequency statistics combined with weighted calculation [66]. We therefore used TF-IDF to extract feature words from text information.

The third step is text vectorization. The text information in feature words cannot be directly understood by machine-learning models. Thus, the information needs to be converted into text representations that are easily processed by machine-learning algorithms. Text vectorization represents a text as a series of vectors that can express textual semantics. In 2013, Google released the word-embedding tool Word2vec [67]. Word2vec can be trained on millions of dictionaries and hundreds of millions of datasets, and the obtained word vectors can be a good measure of word-to-word textual similarity. Word2vec is a shallow neural network that depends on two crucial algorithm models, Continuous Bag of Words and Skip-gram, and two efficient training methods, Negative Sampling and Hierarchical Softmax [68]. We used the Word2vec tool to calculate word vectors.
The last step is text clustering. After text vectorization, the word vectors can be grouped with a clustering algorithm. A widely used clustering algorithm is the $k$-means algorithm [69], which divides the original samples into different clusters by calculating the similarity between samples. Samples in the same cluster are similar to each other, and those in different clusters are different [70]. The main goal of the algorithm is to find independent clusters with more close distances. The least quadratic error of the loss function of the algorithm needs to be satisfied to accomplish this. The equation for the loss function is as follows:

$$RSS_k = \sum_{|x-t(w_k)|^2} |x-t(w_k)|^2, \quad (1)$$

where $x$ is the position of the scattered points in the cluster, $w_k$ is the $k$th cluster, $t(w_k)$ is the center point of the $k$th cluster, and $RSS_k$ is the loss function of the $k$th cluster. The goal of algorithm optimization is to choose a reasonable clustering scheme to minimize $RSS_k$.

In $k$-means clustering, the $k$ number can be determined by the silhouette coefficient. In this case, the $k$ value corresponding to the larger silhouette coefficient was selected. The equation for the silhouette coefficient is as follows:

$$S(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \quad (2)$$

where $a(i)$ is the average distance from the $i$ vector to other points in the same cluster, and $b(i)$ is the average distance from the $i$ vector to all points in the nearest cluster. The value of the silhouette coefficient is in the range of $[-1,1]$. A value closer to 1 means that cohesion and separation are relatively good, and a value closer to $-1$ means cohesion and separation are relatively bad [71]. The total silhouette coefficient of the clustering result can be obtained by averaging the silhouette coefficients of all of the data points. We used the $k$-means algorithm for clustering analysis and the silhouette coefficient method to determine the $k$ number in the clustering process.

4. Data Collection and Analysis
4.1. Data Collection
4.1.1. Data Sources

People.com.cn and Chinanews.com are the two major national online news platforms in China. Their authority, timeliness, and depth and breadth of content have been widely recognized in the field of news communication. We therefore used web crawler technology to collect online news from these two platforms about the creative industry’s virtual agglomeration through keyword retrieval and formed an online news corpus.

4.1.2. Search Term Selection

To ensure accurate, comprehensive data collection, core search terms were selected based on the combination of three keyword categories: virtual network, creative industry, and agglomeration space [26,72]. This generated 891 search terms, such as “online game community”, “virtual design services platform”, “online publishing cluster”, “digital advertising platform”, and “online fashion space”. Relevant news texts were collected from People.com.cn and Chinanews.com based on these search terms. Table 1 shows the composition of the search terms.

| Categories       | Keywords                                                                 |
|------------------|---------------------------------------------------------------------------|
| Virtual Network  | E-Commerce, Online, Virtual, Cloud, Internet, Network, Digital Creativity, Culture, Advertising, Architecture, Art, Design Services, Fashion, Television, Film, Publishing, Media, News, Software, Information Services, Leisure & Entertainment, Games, Animation, Radio |
| Creative Industry| Community, Platform, Space, Forum, Agglomeration, Cluster                  |
| Agglomeration Space|                                                                 |

Table 1. Composition of search terms.
4.1.3. Web Crawler

Octoparse’s data collection system was used to collect news reports related to the creative industry’s virtual agglomeration from People.com.cn and Chinanews.com using the search terms identified above. Octoparse is based on a distributed cloud computing platform that can automatically search, filter by time and category, and collect news information. Online news texts were collected in the period 1–5 January 2021. The content included four parts: title, release time, media source, and body content. Initially, 955 news texts were collected, including 604 from People.com.cn and 391 from Chinanews.com. After eliminating duplicate or irrelevant items, 812 news texts were finally collected, spanning 1 January 2012, to 31 December 2020.

4.2. Data Analysis

4.2.1. Text Preprocessing

First, we used regular expressions to remove English, numbers, and Chinese and English symbols in the text; preprocessing was set to [a-zb-fhj-qs-uw-z0-36-9]. Other specialized terms such as 4G, 5G, AI, AR, and others were thus excluded from the text regularization process. Second, the Jieba toolkit in Python was used to segment the regularized text. Third, we expanded the related stop words based on the Harbin Institute of Technology’s stop word database, forming a database containing 2245 stop words. Based on this, stop words were removed during text segmentation.

4.2.2. Feature-Word Extraction

The news corpus was transformed from text data into a collection of words after preprocessing. The words, however, were still too complex to identify the key factors affecting the creative industry’s virtual agglomeration. Therefore, the words needed to be filtered by feature-word extraction. The TF-IDF algorithm was used to extract feature words, and the extracted feature words and their TF-IDF values were saved as a “feature-words.csv” file for subsequent analysis.

4.2.3. Text Vectorization

Selecting a suitable Chinese word vector corpus and training for natural language processing requires many resources and computing power. The Chinese Information Processing Institute of Beijing Normal University and the DBIIR Laboratory of Renmin University of China provide open-source word vectors based on dozens of Chinese corpora for training in different domains (e.g., Baidu Baike, Chinese Wikipedia, Zhihu Q&A, and Weibo), and they contain a variety of training settings [73]. Among them, the word vector of Zhihu Q&A is obtained from the training results of the Zhihu Q&A corpus, which has a total capacity of 2.1 GB and 1117 k-phrases. The Skip-gram model in the Word2vec toolkit was used to train the corpus (basic settings: window size = 5, dynamic window = yes, subsampling = $1 \times 10^{-5}$, low-frequency word = 10, iteration = 5, negative sampling = 5). We calculated the word vectors in the feature-words.csv file and constructed the word vector matrix by loading the trained Zhihu Q&A word vectors.

4.2.4. Text Clustering

We used the $k$-means algorithm to cluster the word vector matrix of the news text. It was necessary to set the number of clusters ($k$ value) in the results manually because the $k$-means clustering algorithm is an unsupervised learning method with no category attribute. Our study used the silhouette coefficient method to determine and improve the accuracy of the selection of the $k$ value. The variation of the silhouette coefficient for different $k$ values is shown in Figure 2.
Figure 2. Variations in the silhouette coefficient.

As shown in Figure 2, the silhouette coefficient fluctuates between 0.09 and 0.45 as the value of \( k \) increases. When the value of \( k \) is 9, the silhouette coefficient peaks at 0.45, which is the closest the value is to 1. Thus, the value of \( k \) is 9. Table 2 shows the clustering results when the \( k \) value is 9.

| Cluster | Representative Words | Title |
|---------|----------------------|-------|
| 1       | Media, Marketing, Promotion, Short Video, Weibo, Publicity, News, Video, Live Broadcast, Vlog, Film, Self-Media, TV, Radio, Advertising, SEM, SEO, Forum, Headlines, Circle of Friends | Digital marketing and promotion |
| 2       | Platform, E-Commerce, Transaction, Flow, Credit, Service Platform, Trading Platform, Online Platform, Market, Online Shopping, Order, Visits, Subsidies, Tencent, Alibaba, Taobao, JD, WeChat, Client-Side, Baidu | Virtual platform facilities |
| 3       | Consumer, Demand, New Generation, Personalization, Uniqueness, Emerging, Consumption, Habit, Differentiation, Diversification, Characteristic, Authenticity, Professionalism, Rationality, Initiative, Experiential, Entertainment, Emotion, Interest, Challenging | Emerging consumer demand |
| 4       | Cloud Computing, AI, 5G, Blockchain, AR, VR, MR, 4G, Big Data, Internet of Things, ICT, Distributed Computing, Mobile Internet, Internet, Data Center, Technology, Network, Information, Data Mining, R&D | Application of innovative technology |
| 5       | Digital Economy, Digitalization, Intelligence, Sharing, Transformation, Upgrading, Network Economy, Cloud, Green, Palm, Fingertip, Immersive, Precision, Innovation, Reform, New Retail, New Format, Online, Digital Facilities, New Economy | Digital economy environment |
| 6       | Central, Government Department, National, Strategy, Intellectual Property, Party Committees, Provincial, Municipal, Government Affairs, Public Management, Implementation, Planning, Think Tank, Command, Supervision, Measures, Subject, Laws, Regulations, Copyright | Government policy planning |
| 7       | User, Topic, Communication, Interaction, Service, Fans, Viewer, Word of Mouth, Popularity, Praise, Favorites, Retweets, Referrals, Hotspot, Comment, Content, Socialization, Community, Public Opinion, Entertainment | Online interactive service |

Table 2. Clustering results \((k = 9)\).
Table 2. Cont.

| Cluster | Representative Words                                                                 | Title                                |
|---------|--------------------------------------------------------------------------------------|--------------------------------------|
| 8       | Customization, Reservation, Division of Labor, Integration, Synergy, Order Placement, Small Batch, One-Stop, Pop-Up, Cooperation, Alliance, Industry Chain, Value Chain, Interconnection, Resource Integration, Complementary Advantage, Smart Manufacturing, Intelligent Manufacturing, Virtual Workshop, Virtual Production | Flexible manufacturing               |
| 9       | Digital Culture, Digital Creativity, Digital Museum, Digital Art Museum, Digital Music, Digital Artwork, Cultural Museum, Gallery, Art Base, Digital Town, Ancient Town, Forbidden City, Dunhuang, National Museum, Cultural City, Cultural Sites, Silk Road, Ethnic Minority, Folklore, World Heritage | Digitalization of cultural resources |

In Table 2, the feature words of news texts related to the creative industry’s virtual agglomeration are clustered into nine categories. Cluster 1 (digital marketing and promotion) includes words such as “media”, “marketing”, “promotion”, “short video”, “Weibo”, and other words that reflect the characteristics of digital marketing and promotion.

Cluster 2 (virtual platform facilities) includes “platform”, “e-commerce”, “transaction”, “flow”, “credit”, and other words that reflect the characteristics of online virtual platform facilities.

Cluster 3 (emerging consumer demand) includes “consumer”, “demand”, “new generation”, “personalization”, “uniqueness”, and other words that reflect the characteristics of emerging consumer demand.

Cluster 4 (application of innovative technology) includes “cloud computing”, “artificial intelligence”, “5G”, “blockchain”, “big data”, and other words that reflect the characteristics of digital technology applications.

Cluster 5 (digital economic environment) includes “digital economy”, “digitalization”, “intelligence”, “sharing”, “transformation”, “upgrading”, and other words that reflect the characteristics of digital economic environment.

Cluster 6 (government policy planning) includes “central”, “government department”, “national”, “strategy”, “measures”, “intellectual property”, and other words that reflect the characteristics of government policies and strategic planning.

Cluster 7 (online interactive service) includes “user”, “topic”, “communication”, “interaction”, “service”, and other words that reflect the characteristics of communication and interaction services with users.

Cluster 8 (flexible manufacturing) mainly “customization”, “reservation”, “division of labor”, “integration”, “synergy”, and other words that reflect the characteristics of flexible production based on customization.

Cluster 9 (digitalization of cultural resources) includes “digital culture”, “digital creativity”, “digital museum”, “digital art museum”, and other words that reflect the characteristics of the digital transformation of cultural resources.

5. Driving Factors and Mechanism Analysis

5.1. Internal and External Driving Factors

The results of the text clustering indicate that the creative industry’s virtual agglomeration is influenced by nine driving factors: digital marketing and promotion, virtual platform facilities, emerging consumer demand, application of innovative technology, digital economic environment, government policy planning, online interactive service, flexible manufacturing, and the digitalization of cultural resources. These factors are too dispersed, however, to be explained at a systemic level. In creative industries, virtual agglomeration cannot appear without the combined effects of internal and external factors. We therefore
used the PEST and value-chain models to extract the driving factors from the external and internal aspects of this virtual agglomeration.

First, according to the PEST framework, the creative industry’s virtual agglomeration is driven by four external factors: government policy planning (political factor), digital economic environment (economic factor), emerging consumer demand (social factor), and the application of innovative technology (technological factor).

Second, with the rapid development of the digital economy, creative enterprises and institutions in the value chain increasingly rely on digital technology to achieve value creation leaps in virtual space and spawn a new form of virtual agglomeration. We therefore analyzed the internal driving factors of this virtual agglomeration based on an analysis of primary (design, production, marketing, services) and auxiliary (basic guarantees) activities in the value chain. The creative industry’s value chain in virtual space is created based on the activities “digitalization of cultural resources → flexible manufacturing → digital marketing and promotion → online interactive services” and “virtual platform facilities.” These five aspects constitute the internal driving factors of the creative industry’s virtual agglomeration.

This virtual agglomeration is jointly driven by strong external and internal factors. The external driving factors are government policy planning, digital economic environment, emerging consumer demand, and application of innovative technology; the internal factors are the digitalization of cultural resources, flexible manufacturing, digital marketing and promotion, online interactive service, and virtual platform facilities (Figure 3).

**Figure 3.** Driving factors of the virtual agglomeration of creative industries.

5.2. Mechanism Analysis

We combined the relevant literature and industry practices with the analysis of the mechanisms of each factor, as described below.

First, the driving mechanism of government policy planning is primarily reflected in digital strategy [74] and network supervision [75]. Regarding digital strategy, governments have formulated various strategic plans for the digital development of creative industries [76–78], which include accelerating the construction of new infrastructure, encouraging digital innovation in creative enterprises, and stimulating online creative consumption,
all of which provide opportunities for the creative industry’s virtual agglomeration. Regarding network supervision, the government has improved laws and regulations related to online transactions for creative products, safeguarded subjects’ legitimate interests [79], and created a suitable environment for creative production and trading activity on virtual platforms [80].

Second, the driving mechanism of the digital economic environment is reflected in industrial digital transformation [81] and innovation [82]. Regarding industrial digital transformation, the digital economy promotes the transformation of traditional creative industries. It provides more formats, such as digital publishing [83], film [84], and music [85], thus promoting the creative industry’s virtual agglomeration. Regarding industrial digital innovation, innovation in the design, production, marketing, and overall service provision of creative industries has accelerated under the digital economy, producing intelligent, personalized digital creative products and services. This will increasingly attract creative enterprises and users who gather in virtual spaces.

Third, the driving mechanism of emerging consumer demand is predominately reflected in the pursuit of spiritual culture [86] and the rise of online consumption [87]. With improved living standards, more people are pursuing diversified and personalized spiritual and cultural consumption, and a rich diversity of digital cultural and creative products and services in the virtual space can meet these changing consumption demands [88] in a way that promotes the creative industry’s virtual agglomeration. Various new online consumption models have also emerged, such as live e-commerce, creative content e-commerce, and creative community e-commerce. Increasingly, consumers are pursuing green, healthy consumption through this cloud-life model, which promotes prosperity in the online creative consumption market and accelerates the virtual clustering of creative industries.

Fourth, the driving mechanisms for the application of innovative technology appear in innovation support [89] and service intelligence [90]. Regarding innovation support, technologies such as AI, blockchain, cloud computing, big data, and 5G accelerate the formation of cooperative networks among creative enterprises [91], promote the efficient flow of resources [92], accelerate profit acquisition, and attract more creative enterprises to participate in virtual agglomeration. Creative enterprises can also use big data technology to record users’ browsing and purchasing behaviors in real time, capture creative consumption preferences, and portray crowd characteristics, thus bringing more personalized intelligent services to users.

Fifth, the driving mechanism of the digitalization of cultural resources appears in digital cultural inheritance [93,94] and activation [95]. Regarding the former, by combining traditional cultural resources with digital technology, creative entities can create virtual forms of creative products that not only protect these heritage and cultural resources but also allow more users to experience traditional culture free from the constraints of physical space. Regarding the latter, creative entities can browse, learn about, and utilize digital cultural resources; be stimulated to create content [96]; and develop digital creative derivatives to enrich the content ecology of creative industries [97], thus promoting the creative industry’s virtual agglomeration.

Sixth, flexible manufacturing is reflected in personalized customization [98,99] and flexible supply [100]. A flexible production system can achieve personalized customization through user-demand analysis, creative design, product optimization, logistics distribution, and service feedback. It can also greatly enhance the market responsiveness and conversion ability of creative enterprises [101]. With advanced flexible production systems and efficient resource matching on virtual platforms, creative enterprises can flexibly produce and supply according to market demand. This can help to avert inventory backlogs, reduce operating costs [102], and compel more creative subjects to cluster on virtual platforms.

Seventh, the driving mechanism of digital marketing and promotion is reflected in precise cross-media communication [103] and immersive virtual experience [104]. Creative companies use diversified digital media channels, such as Weibo, short videos, and WeChat, to form links with users. They can therefore more accurately explore online markets,
tap into user needs, deliver product ideas, and shape brand culture in virtual space. New communication channels can greatly improve product sales and brand stickiness for creative enterprises and attract more creative institutions and users to gather in virtual spaces [19]. Via digital technologies such as virtual reality, artificial reality, mixed reality, 5G, 4K, and 8K, users can fully immerse themselves in creative experiences, better sense the fit between products and their needs, and enhance their satisfaction [105].

Eighth, the driving mechanism of online interactive services is reflected in two-way service feedback [98,106] and shorter service distance [107]. The linear one-way service model between creative enterprises and users has been gradually replaced by an online two-way interactive service model in the social media environment [108]. Creative enterprises are increasingly inviting users into live rooms, where they join fan groups to share new product expectations, product usage tips, and service satisfaction. This can help improve products and services. Mobile Internet has removed the distance between creative enterprises and users. Users have more opportunities to participate in digital product creation, service development, and event hosting [102], thus promoting the virtual agglomeration of creative industries.

Ninth, the driving mechanism of virtual platform facilities is reflected in network collaboration [109], freedom of participation [3], and the guarantee of trust [110]. Diversified service functions, efficient information sharing [111], and convenient transaction processes of virtual platforms [112] help establish close network collaborations between creative enterprises, thereby reducing production costs, improving service efficiency, and enhancing value creation [113]. Open and inclusive platform features also lower the participation threshold for creative entities [114], causing a more diverse range of subjects to gather on platforms. Meanwhile, reviews on virtual platforms, reputation ratings, violation penalties, and other rules can constrain unethical or illegal online behaviors and protect the rights and interests of all participants [115].

Based on this analysis, Figure 4 shows our proposed driving-force model of the creative industry’s virtual agglomeration.
6. Conclusions and Implications

This study analyzed the causes of the virtual agglomeration of creative industries. We used web-crawler technology to collect relevant online news reports and used text mining to extract nine key factors. We also constructed an analysis framework for the internal and external factors based on the PEST and value-chain models. On that basis, we combined the relevant literature and industry practices and analyzed the mechanism of each factor. Finally, we built a driving-force model of the creative industry’s virtual agglomeration.

The main findings are as follows. First, the creative industry’s virtual agglomeration was found to result from a combination of internal and external factors. The external driving factors are government policy planning, the digital economic environment, emerging consumer demand, and application of innovative technology; the internal factors are the digitalization of cultural resources, flexible manufacturing, digital marketing and promotion, online interactive service, and virtual platform facilities. Second, each driving factor was found to promote the creative industry’s virtual agglomeration through different mechanisms. Among them, (1) government policy planning is based on digital strategy and network supervision [3]; (2) the digital economic environment is based on industrial digital transformation and innovation; (3) emerging consumer demand is based on the pursuit of spiritual culture and the rise of online consumption; (4) application of innovative technology is based on innovation support and service intelligence [84]; (5) the digitalization of cultural resources is based on digital cultural inheritance and activation [97]; (6) flexible manufacturing is based on personalized customization and flexible production and supply [92]; (7) digital marketing and promotion are based on cross-media communication and immersive virtual experiences; (8) online interactive services are based on two-way service feedback and shortened service distance [116]; and (9) virtual platform facilities are based on network coordination, freedom of participation, and guarantee of trust [117]. The creative industry’s virtual agglomeration arises from the influence of these factors and mechanisms.

Among the driving factors found in this study, government policy planning, the digital economic environment, and emerging consumer demand are key factors that also affect the geographic agglomeration of creative industries [118–121]. However, factors such as office rent, transportation facilities, and urban cultural facilities that affect this geographic agglomeration [122–124] are no longer significant in virtual agglomeration because virtual agglomeration is formed based on virtual space, which eliminates the restriction of geographical boundaries [37]. It will therefore be difficult for region-related factors to have a critical impact. This study also identified some new factors, such as the application of innovative technology, the digitization of cultural resources, flexible manufacturing, digital marketing and promotion, online interactive services, and virtual platform facilities. These emerging factors are worth exploring in future research.

This study enriches the research on the creative industry’s virtual agglomeration. First, we addressed its causes, discovered the internal and external driving factors and their mechanisms through big data text mining and theoretical analysis, and constructed a driving-force model. The proposed model can lay a theoretical foundation for future research on the creative industry’s virtual agglomeration. Second, we used web-crawler technology and text-mining methods (text preprocessing, feature-word extraction, text vectorization, and text clustering) to identify the key factors affecting virtual agglomeration, thus ensuring the reliability of the results. The use of big-data technology in this study broadens the choice of methods for related future research.

Our findings also have practical implications for guiding the creative industry’s virtual agglomeration. First, governments should focus on constructing digital infrastructure for creative industries and encourage creative enterprises to explore the commercial applications of digital facilities, such as 5G, cloud computing, big data, and the Internet of Things. In addition, to create a fair and mutually trusting digital economic environment, governments should innovate in terms of digital governance methods by guiding the digital innovation of creative industries, maintaining order in the digital economy, and
cracking down on illegal digital activities [125]. Second, creative entities should strengthen R&D on generic technologies, algorithms, and software and establish a network cooperation system for interconnectivity between the different parts of the creative-industry chain [126]. Creative entities should explore new technologies to creatively transform traditional cultural resources, develop creative digital products, and optimize the virtual ecology of creative industries [127]. Third, virtual platforms should continue to improve their intelligent services through more advanced algorithms, improved rules [128], and humanized functions. This will help to achieve more accurate information matching, network collaboration, and value trading among creative subjects [129]. Virtual platforms should also innovate in terms of their business models, promote the efficient use of creative resources, and integrate overall creative industry processes. Fourth, creative enterprises can change the creation mode of digital products and meet consumers’ social and creative needs by inviting them to participate in content creation, as well as developing digital products, such as interactive videos, movies/TV, and books. Creative enterprises should aim to customize digital creative products with precision based on user preferences, thus promoting positive interactions between supply and demand on virtual platforms.

This study has some limitations. First, we did not empirically test the effects of the factors and their mechanisms. Second, this study was conducted based on Chinese online news texts related to the creative industry’s virtual agglomeration. Whether our conclusions are equally applicable to Western creative industries needs to be verified. These limitations suggest directions for future research.

Author Contributions: All of the authors contributed to the study. Conceptualization: X.C. and C.L.; software: Y.J.; supervision: C.G.; writing (original draft): X.C.; writing (review and editing): Y.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Shanghai Science and Technology Innovation Action Plan: Shanghai Sailing Program (No. 21YF1415900); the Shanghai Science and Technology Innovation Action Plan: Soft Science Key Program (No. 21692192200); the National Natural Science Foundation of China (Nos. 71874027 and 72104137); the Major Science and Technology Innovation Project of Shanghai Municipal Education Commission (No. 2017-01-07-00-03-E00044).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors have no conflict of interest to declare.

References
1. Yipin Witkey’s Nine Years of Accumulation to Build the Benchmark of Enterprise Service Industry. Available online: https: //www.sohu.com/a/325019163_115470 (accessed on 5 July 2019).
2. Steamworks. Available online: https://partner.steamgames.com/ (accessed on 8 June 2021).
3. Chen, X.; Liu, C.; Gao, C.; Jiang, Y. Mechanism Underlying the Formation of Virtual Agglomeration of Creative Industries: Theoretical Analysis and Empirical Research. Sustainability 2021, 13, 1637. [CrossRef]
4. Aslesen, H.W.; Martin, R.; Sardo, S. The virtual is reality! On physical and virtual space in software firms’ knowledge formation. Entrep. Reg. Dev. 2019, 31, 669–682. [CrossRef]
5. Molina, A.; Flores, M.; Caballero, D. Virtual Enterprises: A Mexican Case Study. In Proceedings of the International Conference on Information Technology for Balanced Automation Systems, Prague, Czech Republic, 4 August 1998; Springer: Boston, MA, USA, 1998; pp. 159–170.
6. Molina, A.; Flores, M. A virtual enterprise in Mexico: From concepts to practice. J. Intell. Robot. Syst. 1999, 26, 289–302. [CrossRef]
7. Mason, C.; Castleman, T.; Parker, C. Communities of enterprise: Developing regional SMEs in the knowledge economy. J. Enterp. Inf. Manag. 2008, 21, 571–584. [CrossRef]
8. Salvador, E.; Mariotti, I.; Conicella, F. Science park or innovation cluster? Int. J. Entrep. Behav. Res. 2013, 19, 656–674. [CrossRef]
9. Fu, H.; Yeh, H.; Ma, R. A study of the CSFs of an e-cluster platform adoption for microenterprises. Inf. Technol. Manag. 2018, 19, 231–243. [CrossRef]
10. Brown, D.H.; Lockett, N.J. Engaging SMEs in e-commerce: The role of intermediaries within eClusters. Electronic Markets 2001, 11, 52–58. [CrossRef]
11. Mason, C.; Castleman, T.; Parker, C. Knowledge Management for SME-Based Regional Clusters. In Proceedings of the 2004 Collector E-Commerce Conference, Adelaide, Australia, 7–8 May 2004; pp. 1–17.

12. Caballero, D.; Molina, A.; Bauernhansl, T. A Methodology to Evaluate Enterprises to Become Members of Virtual Industry Clusters. In Proceedings of the Working Conference on Virtual Enterprises, Florianópolis, Brazil, 4–6 December 2000; pp. 443–454.

13. Passiante, G.; Secundo, G. From geographical innovation clusters towards virtual innovation clusters: The Innovation Virtual System. In Proceedings of the 42nd Congress of the European Regional Science Association: “From Industry to Advanced Services-Perspectives of European Metropolitan Regions”, Dortmund, Germany, 27–31 August 2002; pp. 1–22.

14. Chen, X.Y. Virtual Transformation of Industrial Cluster. China Ind. Econ. 2017, 357, 78–94.

15. Carrie, A. Integrated clusters—the future basis of competition. Int. J. Agil. Manag. Syst. 1999, 1, 45–50. [CrossRef]

16. Adebanjo, D.; Michaelides, R. Analysis of Web 2.0 enabled e-clusters: A case study. Technovation 2010, 30, 238–248. [CrossRef]

17. Kawa, A.; Rafajczak-Mrozek, M. Cloud Community in Logistics e-Cluster. In Asian Conference on Intelligent Information and Database Systems; Springer: Cham, Switzerland, 2014; pp. 495–503.

18. Davidovic, M. Building e-Clusters. Bus. Logist. Mod. Manag. 2013, 13, 211–223.

19. Han, R.B. Ecological Development and Dimensions of the Cultural and Creative Industries Integral Age. J. Shandong Univ. Bus. Manag. 2020, 49–59.

20. Du, Y.; Xu, C.; Wu, J. Exploring the construction framework of my country’s animation industry virtual cluster. Sci. Technol. Manag. Res. 2009, 29, 168–170.

21. Wang, Y. Research on Virtual Information Service Industry Cluster. Master J. Central China Norm. Univ. 2009.

22. Xiao, Y. The Development Motivation, Dilemma and Countermeasures of Digital Publishing Industry Cluster. Mod. Publ. 2014, 16–19.

23. Crewe, L. When virtual and material worlds collide: Democratic fashion in the digital age. Environ. Plan. A 2013, 45, 760–780. [CrossRef]

24. Zhang, X.; Sun, C.; Mei, L. Agglomerative patterns and cooperative networks of the online video industry in China. Reg. Stud. 2021, 55, 1–13. [CrossRef]

25. Zukin, S. Loft Living: Culture and Capital in Urban Change; Rutgers University Press: New Brunswick, NJ, USA, 1989.

26. Lazzaretto, L.; Boix, R.; Capone, F. Do Creative Industries Cluster? Mapping Creative Local Production Systems in Italy and Spain. Ind. Innov. 2008, 15, 549–567. [CrossRef]

27. Hong, J.; Yu, W.T.; Guo, X.M.; Zhao, D.T. Creative industries agglomeration, regional innovation and productivity growth in China. Chin Geogr. Sci. 2014, 24, 258–268. [CrossRef]

28. Gong, H.W.; Hassink, R. Exploring the clustering of creative industries. Eur. Plan. Stud. 2017, 25, 583–600. [CrossRef]

29. Granpayehvaghei, T.; Bonakdar, A.; Zandiatashbar, A.; Hamidi, S. The Quest for Creative Industries: A Multilevel National Study of the Impacts of Urban Form on the Geography of Creative Industries. Transp. Res. Rec. 2019, 2673, 157–168. [CrossRef]

30. Coll-Martinez, E.; Méndez-Ortega, C. Agglomeration and coagglomeration of co-working spaces and creative industries in the city. Eur. Plan. Stud. 2020, 1–22. [CrossRef]

31. Coll-Martinez, E.; Moreno-Monroy, A.I.; Arauzo-Carod, J.M. Agglomeration of creative industries: An intra-metropolitan analysis for Barcelona. Pap. Reg. Sci. 2019, 98, 409–431. [CrossRef]

32. Scott, A.J. The US recorded music industry: On the relations between organization, location, and creativity in the cultural economy. Environ. Plan. A 1999, 31, 1965–1984. [CrossRef]

33. Tao, J.; Ho, C.Y.; Luo, S.; Sheng, Y. Agglomeration economies in creative industries. Reg. Sci. Urban Econ. 2019, 77, 141–154. [CrossRef]

34. Branzanti, C. Creative clusters and district economies: Towards a taxonomy to interpret the phenomenon. Eur. Plan. Stud. 2015, 23, 1401–1418. [CrossRef]

35. Hracs, B.J. A Creative Industry in Transition: The Rise of Digitally Driven Independent Music Production. Growth Chang. 2012, 43, 442–461. [CrossRef]

36. Xiong, C.; Zhang, X. Composition and integration of creative industry elements under the perspective of cluster leap. Jiang-Huai Tribune 2020, 12, 135–140.

37. Picard, R.G. Media Clusters: Local Agglomeration in an Industry Developing Networked Virtual Clusters; Jönköping International Business School: Jönköping, Sweden, 2008.

38. Wang, R.Y.; Liang, Q.; Li, G.Q. Virtual Agglomeration: A New Form of Spatial Organization with Deep Integration of New Generation Information Technology and Real Economy. Manag. World 2018, 34, 13–21.

39. Zang, Z.P. GVC restructuring of the digital creative industries: Strategic position and China’s path. Stud. Sci. Sci. 2018, 36, 825–830.

40. Wu, Q.; Li, Y. Management Innovation of Virtual Industry Cluster. Econ. Manag. 2008, 30, 11–15.

41. Wei, J.; Wu, Q. “Internet +” and the transformation and upgrading of cultural and creative industry clusters: From the perspective of networked relations. J. Southw. Minzu Univ. 2019, 40, 148–154.

42. Lin, J.; de Kloet, J. Platformization of the unlikely creative class: Kuaishou and Chinese digital cultural production. Soc. Media+ Soc. 2019, 5, 1–12. [CrossRef]

43. Arsenault, A.H. The datafication of media: Big data and the media industries. Int. J. Media Cult. Polit. 2017, 13, 7–24. [CrossRef]

44. Li, H.; Quan, H. Research on Tourism Virtual Industry Cluster Based on Big Data Platform. Bus. Manag. J. 2018, 40, 21–38.
45. Mete, M.; Yuruk, N.; Xu, X.; Berleant, D. Knowledge discovery in textual databases: A concept-association mining approach. In *Data Engineering*; Springer: Boston, MA, USA, 2009; pp. 225–243.

46. Shah, S.M.; Lütjen, M.; Freitag, M. Text Mining for Supply Chain Risk Management in the Apparel Industry. *Appl. Sci.* 2021, 11, 2323. [CrossRef]

47. Vicente, M.R.; López-Menéndez, A.J.; Pérez, R. Forecasting unemployment with internet search data: Does it help to improve predictions when job destruction is skyrocketing? *Technol. Forecast. Soc. Chan.* 2015, 92, 132–139. [CrossRef]

48. Cavallaro, A.; Rigobon, R. The Billion Prices Project: Using Online Prices for Measurement and Research. *J. Econ. Perspect.* 2016, 30, 151–178. [CrossRef]

49. Lin, S.J.; Huang, L.M.; Shih, M.L.; You, H.J. Online Word-of-mouth Analysis of Customer Satisfaction with the Cultural and Creative Industries in Taiwan-an Example of Creativity Living Industry. *Int. J. Adv. Comput. Technol.* 2011, 3, 242–251. [CrossRef]

50. Yuan, J.; Liu, Y. Research on Tourism Development of Cultural and Creative Industry Park Based on the Network Text Analysis: A Case Study of Dafen Oil Painting Village in Shenzhen. In Proceedings of the 2017 3rd International Conference on Information Management (ICIM), Chengdu, China, 21–23 April 2017; pp. 37–41.

51. Casadei, P.; Lee, N. Global cities, creative industries and their representation on social media: A micro-data analysis of Twitter data on the fashion industry. *Environ. Plan. Econ. Space* 2020, 52, 1195–1220. [CrossRef]

52. Ho, J.K.-K. Formulation of a systemic PEST analysis for strategic analysis. *Eur. Acad. Res.* 2014, 2, 6478–6492.

53. Lee, J.M.; Rim, M.H. Derivation of Creative SW HRD Policy Using Analytic Hierarchy Process. *J. Digit. Converg.* 2013, 11, 95–102.

54. Xie, X.; Xie, X.; Martinez-Climent, C. Identifying the factors determining the entrepreneurial ecosystem of cultural and creative industries in emerging economies. *Int. Entrep. Manag. J.* 2019, 15, 503–522. [CrossRef]

55. Porter, M.E. *Competitive Advantage: Creating and Sustaining Superior Performance*; The Free Press: New York, NY, USA, 1985.

56. Xing, H. Research on the Value-chain Merge and the Development Path of the Cultural and Creative Industry. *Bus. Manag. J.* 2009, 31, 37–41.

57. Horng, S.-C.; Chang, A.-H.; Chen, K.-Y. The business model and value chain of cultural and creative industry. In *Thriving in a New World Economy*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 198–203.

58. Madudová, E. Creative industries value chain: The value chain logic in supply chain relationships. *Market. Brand. Res.* 2017, 4, 227–235. [CrossRef]

59. Wang, X.D.; Jian, G. Research on the Informal Scientific Communication of the Humanities and Social Sciences Based on the Extraction of News Information. *Libr. Inf.* 2018, 2, 43–53. [CrossRef]

60. Zihua, H. *News Theory Course*; Higher Education Press: Beijing, China, 2010.

61. Zhang, X.B. Self-evolution: A multi-dimensional interpretation of the evolution of online news. *Media* 2019, 85–87.

62. Feldman, R.; Dagan, I. Knowledge Discovery in Textual Databases (KDT). In Proceedings of the First ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Montreal, PQ, Canada, 26–29 August 1995; pp. 112–117.

63. Tan, A.-H. Text mining: The State of the Art and the Challenges. In Proceedings of the Pakdd 1999 Workshop on Knowledge Discovery from Advanced Databases, Beijing, China, 10 October 1999; pp. 65–70.

64. Zhang, T.; Ma, H.Q. Comparative Study on Artificial Intelligence Policies in China Based on Text Similarity Computation. *J. Intell.* 2021, 40, 39–47.

65. Zhang, M.Y.; Zhu, G.L.; Zhang, S.X.; Zhang, B. Grouping Microblog Users of Trending Topics Based on Sentiment Analysis. *Data Anal. Knowl. Discov.* 2021, 5, 43–49.

66. Ramos, J. Using tf-idf to Determine Word Relevance in Document Queries. In Proceedings of the The First Instructional Conference on Machine Learning, 3–8 December 2003; pp. 29–48.

67. Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G.; Dean, J. Distributed representations of words and phrases and their compositionality. *arXiv 2013*, arXiv:1310.4546.

68. Ma, L.; Zhang, Y. Using Word2Vec to Process Big Text Data. In Proceedings of the 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, USA, 29 October–1 November 2015; pp. 2895–2897.

69. MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 1 January 1967; pp. 281–297.

70. Zhang, L.J.; Gao, C.X. Indicators Screening Method Based on k-means Clustering and Rough Set Algorithm. *Oper. Res. Manag. Sci.* 2020, 29, 8–12.

71. Liu, R.L.; Ye, W.H.; Gao, R.Q.; Tang, M.J.; Wang, D.B. Research on Text Clustering Based on Requirements of Big Data Jobs. *Data Anal. Knowl. Discov.* 2017, 1, 32–40.

72. Wilkinson, A. An Assessment of productivity indicators for the creative industries. *DCMS London 2007*.

73. Li, S.; Zhao, Z.; Hu, R.; Li, W.; Liu, T.; Du, X. Analogical reasoning on chinese morphological and semantic relations. *arXiv 2018*, arXiv:1805.06504.

74. Snowball, J.; Tarentaal, D.; Sapsed, J. Innovation and diversity in the digital cultural and creative industries. *J. Cult. Econ.* 2021, 1–29. [CrossRef]

75. Liu, S.; Song, Y. Research on the Legalization Path of China’s Network Cultural Market Supervision. *J. Hunan Univ.* 2020, 34, 1–9.

76. Zang, Z. Research on the Restructuring Strategy of the Global Value Chain of the Digital Creative Industry: Based on the Three-dimensional Collaborative Innovation of Content, Technology and System. *Soc. Sci. Res.* 2018, 45–54.

77. Sun, S.; Min, X.; Tang, Y. Development status and prospects of digital creative industry. *Packaging Eng.* 2019, 40, 65–74.
Sustainability 2021, 13, 9232

78. Jin, X.; Li, K. Integrated Development and Transformation of UK’s Innovation Industry in the Context of Digital Economy Strategy. J. Shenzhen Univ. 2020, 37, 65–73.
79. Wei, W.; He, Z.; Rayman-Bacchus, L.; Xiang, G. Do Industrial Clusters Still Matter to Trust-Building in the Internet Era? A Network Embeddedness Perspective. SAGE Open 2019, 9, 1–12. [CrossRef]
80. Peng, B.; Zhang, Q. The formation and evolution of China’s internet governance model from 1994 to 2019. J. Commun. 2019, 27, 44–65.
81. Li, F. The digital transformation of business models in the creative industries: A holistic framework and emerging trends. Technovation 2020, 92–93, 102012. [CrossRef]
82. Bertola, P.; Teunissen, J. Fashion 4.0. Innovating fashion industry through digital transformation. Res. J. Text. Appar. 2018, 22, 352–369. [CrossRef]
83. Li, X.; Liao, C.; Xie, Y. Digital piracy, creative productivity, and customer care effort: Evidence from the digital publishing industry. Mark. Sci. 2021, 40, 685–707. [CrossRef]
84. Ma, S.M. The Application of Interactive Animation in the Film and Television Industry Under the Background of Digital Media. Agro Food Ind. Hi Tech. 2017, 28, 1578–1581.
85. Aguair, L.; Martens, B. Digital music consumption on the internet: Evidence from clickstream data. Inf. Econ. Policy 2016, 34, 27–43. [CrossRef]
86. Xiang, Y. Creative Fusion: The Development Trend and New Normal of China’s Cultural Industry. Arts Crit. 2015, 20, 33–38.
87. Fan, Z. Innovation and Development of Cultural Industry in the Revolution of Digital Economy. J. Shenzhen Univ. 2020, 37, 50–56.
88. Fan, Z. The Revolution in Creativity in the Digital Economy; The Commerical Press: Beijing, China, 2019.
89. Xie, X.; Zhang, J. Technology Empowerment: Digitalization and Intelligent Transformation of New Cultural and Creative Industries. View Publ. 2019, 9–13.
90. Jain, G.; Rakesh, S.; Nabi, M.K.; Chaturvedi, K. Hyper-personalization–fashion sustainability through digital clienteling. Res. J. Text. Appar. 2018, 22, 320–334. [CrossRef]
91. Xiong, L.; Ji, J.; Chen, P. Dynamic mechanism design: The collaborative innovation in digital content industry based on platform economics. Sci. Technol. Manag. Res. 2016, 36, 21–25.
92. Yan, W.; Chiou, S. Dimensions of Customer Value for the Development of Digital Customization in the Clothing Industry. Sustainability 2020, 12, 4639. [CrossRef]
93. Borissova, V. Cultural heritage digitization and related intellectual property issues. J. Cult. Herit. 2018, 34, 145–150. [CrossRef]
94. Psomadaki, O.I.; Dimoulas, C.A.; Kalliris, G.M.; Paschalidis, G. Digital storytelling and audience engagement in cultural heritage management: A collaborative model based on the Digital City of Thessaloniki. J. Cult. Herit. 2019, 36, 12–22. [CrossRef]
95. Shi, B.; Guangyun, Z. Analysis on Advantages and Influencing Factors of Design and Development of Museum Cultural & Creative Products. In Proceedings of the 2019 7th International Forum on Industrial Design, Luoyang, China, 17–19 May 2019; pp. 1–6.
96. Gong, H.; Xin, X. Buzz and tranquility, what matters for creativity? A case study of the online games industry in Shanghai. Geoforum 2019, 106, 105–114. [CrossRef]
97. Guan, W.; Qiao, L.; Chen, K. Dissemination of Marine History and Culture Based on Virtual Museum Technology. J. Coast. Res. 2020, 110, 150–153. [CrossRef]
98. Liao, C.-S. Self-Construal, Personalization, User Experience, and Willingness to Use Codesign for Online Games. Inform. Commun. Soc. 2012, 15, 1298–1322. [CrossRef]
99. Shang, X.; Liu, X.; Xiong, G.; Cheng, C.; Ma, Y.; Nyberg, T.R. Social Manufacturing Cloud Service Platform for the Mass Customization in Apparel Industry. In Proceedings of the 2013 IEEE International Conference on Service Operations and Logistics, and Informatics, Dongguan, China, 28–30 July 2013; pp. 220–224.
100. Zawadzki, P.; Żywicki, K. Smart Product Design and Production Control for Effective Mass Customization in the Industry 4.0 Concept. Manag. Prod. Eng. Rev. 2016, 7, 105–112. [CrossRef]
101. Li, G.; Ding, X.H.; Hu, B.L. The impact of business flexibility on urban production development driven by “internet + “ technologies. Stud. Sci. Sci. 2019, 37, 1797–1804.
102. Qi, J.D.; Xiao, X. Transformation of Enterprise Management in the Era of Digital Economy. Manag. World 2020, 36, 135–152+250.
103. Kim, S.M. How can we make a socially optimal large-scale media platform? Analysis of a monopolistic Internet media platform using two-sided market theory. Telecommun. Policy 2016, 40, 899–918. [CrossRef]
104. Marchand, A.; Hennig-Thurau, T. Value Creation in the Video Game Industry: Industry Economics, Consumer Benefits, and Research Opportunities. J. Interact. Mark. 2013, 27, 141–157. [CrossRef]
105. Yao, X.; Qin, X.B. Technology and Survival: The Essence of Digital Marketing. Journalism Res. 2013, 122, 58–63.
106. Arakji, R.Y.; Lang, K.R. Digital consumer networks and producer-consumer collaboration: Innovation and product development in the video game industry. J. Manag. Inform. Syst. 2007, 24, 195–219. [CrossRef]
107. Brogi, S.; Calabrèse, A.; Campisi, D.; Capece, G.; Costa, R.; Di Pillo, F. The effects of online brand communities on brand equity in the luxury fashion industry. Int. J. Eng. Bus. Manag. 2013, 5, 5–32. [CrossRef]
108. Yang, W.; Wang, H. Digital Marketing Strategy of Higher Education Press Based on SIVA Paradigm. Publ. J. 2018, 26, 81–84.
109. Gu, N.; Singh, V.; London, K.; Brankovic, L.; Taylor, C. Adopting Building Information Modeling (BIM) as Collaboration Platform in the Design Industry. In Proceedings of the CAADRIA 2008: Beyond Computer-Aided Design: Proceedings of the 13th Conference on Computer Aided Architectural Design Research in Asia, Chiang Mai, Thailand, 9–12 April, 2008; pp. 1–9.
110. Wang, Z.; Yang, L.; Wang, Q.; Liu, D.; Xu, Z.; Liu, S. ArtChain: Blockchain-Enabled Platform for Art Marketplace. In Proceedings of the 2019 IEEE International Conference on Blockchain (Blockchain), Atlanta, GA, USA, 14–17 July 2019; pp. 447–454.

111. Chen, C.; Hung, S. To give or to receive? Factors influencing members’ knowledge sharing and community promotion in professional virtual communities. Inf. Manag. 2010, 47, 226–236. [CrossRef]

112. Wu, B. Research on operation mechanism of business ecosystem of creative design platform enterprises. J. Natl. Acad. Art 2017, 38, 110–116.

113. Komninos, N.; Kakderi, C.; Collado, A.; Papadaki, I.; Panori, A. Digital Transformation of City Ecosystems: Platforms Shaping Engagement and Externalities across Vertical Markets. J. Urban Technol. 2020, 28, 1–22.

114. Rai, A.; Constantinides, P.; Sarker, S. Next-generation digital platforms: Toward human-AI hybrids. MIS Q. 2019, 43, iii–ix.

115. Gareev, T.R. Platform Markets: Their Place in the Theory of Mesoeconomic System: Development and a Challenge to Spatial Studies. Baltic Reg. 2018, 10, 26–38. [CrossRef]

116. Suseno, Y.; Laurell, C.; Sick, N. Assessing value creation in digital innovation ecosystems: A Social Media Analytics approach. J. Strateg. Inf. Syst. 2018, 27, 335–349. [CrossRef]

117. Wei, C.; Wang, Q.N.; Liu, C.Y. Research on Construction of a Cloud Platform for Tourism Information Intelligent Service Based on Blockchain Technology. Wirel. Commun. Mob. Comput. 2020, 2020, 8877625. [CrossRef]

118. Chen, J.; Ge, B. Analysis on the Agglomeration Effect and Influence Factors of Cultural and Creative Industries. Contemp. Econ. Manag. 2008, 30, 71–75.

119. Yuan, J. Research on the Generation and Optimization of the Culture Creative Industry Clusters. Chongqing Soc. Sci. 2015, pp. 55–61. Available online: https://en.cnki.com.cn/Article_en/CJFDTotal-CQSK201506009.htm (accessed on 8 June 2021).

120. Yu, J. Agglomeration of cultural and creative industries: Taking Jinan as an example. Shandong Soc. Sci. 2017, 156–161.

121. Jiang, Y.; Gao, C.; Chen, X. Formation of creative industrial spatial agglomeration: Knowledge spillover and mutualism. Sci. Res. Manag. 2020, 41, 119–129.

122. Newbigin, J. Creative Economy and Cultural Industries. Part 1: Creative Economy: A Guide for Beginners; British Council: London, UK, 2010.

123. Lawton, P.; Murphy, E.; Redmond, D. Residential preferences of the ‘creative class’? Cities 2013, 31, 47–56. [CrossRef]

124. Qian, H.; Liu, S. Cultural entrepreneurship in US cities. J. Urban Affairs 2018, 40, 1043–1065. [CrossRef]

125. Tsai, H.; Lee, H.; Yu, H. Developing the Digital Content Industry in Taiwan. Rev. Policy Res. 2008, 25, 169–188. [CrossRef]

126. Olsson, H.H.; Bosch, J. Going digital: Disruption and transformation in software-intensive embedded systems ecosystems. J. Softw. Evol. Proc. 2020, 32, e2249. [CrossRef]

127. Parc, J.; Kim, S.D. The Digital Transformation of the Korean Music Industry and the Global Emergence of K-Pop. Sustainability 2020, 12, 7790. [CrossRef]

128. Ren, H.; Wang, R.; Zhang, S.; Zhang, A. How Do Internet Enterprises Obtain Sustainable Development of Organizational Ecology? A Case Study of LeEco Using Institutional Logic Theory. Sustainability 2017, 9, 1375. [CrossRef]

129. Hänninen, M.; Smedlund, A.; Mitronen, L. Digitalization in retailing: Multi-sided platforms as drivers of industry transformation. Balt. J Manag. 2018, 13, 152–168. [CrossRef]