



Research on the Characteristics and Evolution of Digital Economic Policy Topics Based on BERTopic Model¹

Ye Li^{a*}, Cunyang Zhang^b, Linfang Zhao^c

Institute of Science and Technology for Development of Shandong, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China

^a*liye@sdas.org; ^bzbcunyal@163.com; ^c2697856726@qq.com;

Abstract. The digital economy is an important engine for building a digital China and a modern economic system, and digital economic policies are an important guarantee for the high-quality development of the digital economy. This article collects 794 digital economy policies at the national and provincial levels from 2003 to 2023. The BERTopic algorithm is used to mine the topics of digital economy policy texts. Based on landmark events, the development of the digital economy is vertically divided into three stages, and horizontally divided into eight directions. The thematic characteristics of policies at each stage are analyzed. Research has found that in the early days, the government focused on the digital transformation of public services, and then on industrial development. Currently, attention distribution tends to be balanced. The policy coverage is becoming wider and deeper, and a complete policy system has now been formed. In the future, policies will continue to focus on data development and utilization, as well as data market construction. In addition, the heat of policies related to improving the security system of the digital economy and enhancing the governance capacity of the digital economy is continuously increasing.

Keywords: digital economy; policy topic mining; BERTopic; topic model;

1 Introduction

The digital economy is an important engine for shaping new drivers and advantages of high-quality economic development. The development of the digital economy cannot be separated from the guidance and support of policies. In order to seize development opportunities, since the 18th National Congress of the Communist Party of China, the Party Central Committee has elevated the development of the digital economy to a national strategy and formulated a series of major policies and guidelines. Currently, China's digital economy is flourishing and has achieved a series of innovative breakthroughs. With the deepening of development, the development of the digital economy has become increasingly complex, bringing many challenges to the formulation and

¹ This work was supported by a grant from Social Sciences Planning Research Program of Shandong [grant number 23CSDJ13].

implementation of policies. Reviewing and organizing existing digital economy policies, exploring the development context and trends of policies, is of great significance for further enhancing the scientific, precise, and effective nature of policies.

Based on this, this article conducts a quantitative study of policy texts on digital economy policies at the national and provincial levels in China. The main research approach is to use a total of 794 policy documents related to the digital economy at the national and provincial levels from January 2003 to December 2023 as research objects and samples, divide them into three development stages, sort out the basic characteristics and thematic evolution laws of digital economy policies, study the evolution process of China's digital economy policies, and provide practical basis for subsequent formulation.

2 Research Status

Most research on policy texts built analytical frameworks based on policy tool theory to study policy effectiveness such as the "tool-goal" framework[1] by Yang Qiaoyun et al., the "tool-preference" framework[2] by Tan Haibo et al., and the "subject-tool-cycle" framework[3] by Lei Hongzhu et al. Scholars coded and classified policy texts according to theoretical frameworks to discuss issues such as whether policy tools were unbalanced, whether policy priorities were biased, and how well the policy goals had been achieved. Recently, Chen Mei et al. have attempted to use text processing techniques to extract text topic words, draw a co word network based on social network analysis, and combine or assist other methods to carry out policy research[4]. However, the level of this research tends to be micro level, and there is insufficient exploration of the overall policy context. In addition, exploring the evolution of policy by using the PMC (Policy Modeling Consistency) index model is also the mainstream of current quantitative research on policy texts. By constructing a PMC index model, Bu Lingtong et al. compared regional differences for referencing[5]. Wang Weiguang et al. evaluated 62 policies from 31 provinces, then suggested that the resource advantages and the characteristics of technological development in different regions should be fully considered[6]. Cai Dongsong et al. quantitatively evaluated the digital economy policies in Jilin Province and proposed optimization suggestions from four aspects: policy nature, policy areas, policy effectiveness, and policy priorities[7]. Overall, current research on digital economy policies has formed a certain theoretical foundation. However, related research is mostly limited to analyzing policy tool frameworks or constructing PCM index models for quantitative evaluation. There is still little research on policy topic analysis and the application of topic mining models.

At present, text topic mining mostly relies on the Latent Dirichlet Allocation (LDA) model [8], but the problem of missing semantics in the LDA calculation process leads to a lot of meaningless vocabulary in the results, and it is difficult to make targeted improvements to the study of text characteristics in this drama. The new achievement in the field of topic mining, the BERTopic model [9], can effectively solve the problem of lexical semantic loss and obtain topic results with moderate granularity by adjusting parameters, which is better than LDA [10]. In addition, BERTopic is suitable for short

text topic extraction and is suitable for policy texts with clear semantic boundaries between paragraphs. Therefore, this article chooses the BERTopic model as a policy text processing tool.

3 Data Processing

3.1 Collecting Samples and Dividing Stages

The samples studied in this article mainly come from the portal websites of the Central People's Government and the people's governments of 31 provinces (autonomous regions, municipalities directly under the central government), as well as the government document database of China National Knowledge Infrastructure. Set the deadline as December 2023, select keywords such as "digital", "data", and "intelligence" to search for policy titles, filter the search results, and ultimately retain 794 policy documents, including 67 national policies and 727 provincial policies.

Policy evolution analysis is usually based on landmark events as nodes to divide policy development stages for research. The first policy collected in this article dates back to 2003, so it is used as the starting point. In 2015, the State Council issued the Guiding Opinions on Actively Promoting the "Internet plus" Action, which led the central departments and local governments to launch a large number of "Internet+" policies in industries, agriculture, medical care and other fields. Since then, China has entered a new stage where "Internet + Industry" is the core radiation to drive the rapid development of other digital technologies. In 2021, the State Council released the first national level special plan for digital economy, the "14th Five Year Plan for Digital Economy Development," which clearly pointed out that China's digital economy has entered a new stage of deep application, standardized development, and inclusive sharing. Therefore, this article divides digital economy policies into preliminary exploration stage (2003-2014), rapid improvement stage (2015-2020), and comprehensive development stage (2021-2023) to explore the development process of digital economy policies.

3.2 BERTopic Modeling

BERTopic is a topic mining model that generates multiple topics by clustering text. The modeling process can be divided into text embedding, text vector dimensionality reduction, topic clustering, calculating keyword scores, and generating topic representations. The calculation method is shown as formula (1).

$$W_{t,c} = tf_{t,c} \cdot \log\left(1 + \frac{A}{tf_t}\right) \quad (1)$$

Among them, $W_{t,c}$ represents the importance score of keyword t in topic cluster c , $tf_{t,c}$ represents the frequency of keyword t in topic cluster c , tf_t represents the frequency of keyword t appearing in all clusters, and A represents the average vocabulary of each cluster. The detailed modeling process is shown in Figure 1.

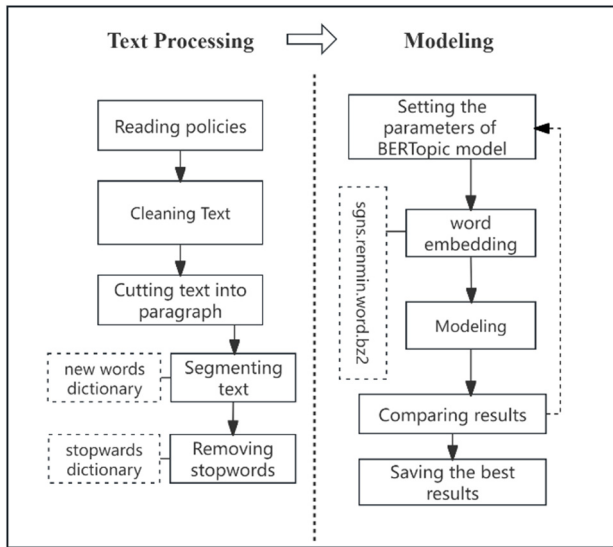


Fig. 1. Modeling process diagram

Text Preprocessing

Firstly, reading and cleaning the text, and removing invisible characters from the text; secondly, dividing the text into paragraphs and assigning a paragraph to a document object; then, the paragraph is segmented by using a new vocabulary dictionary developed by our research group; finally, removing stop words.

Modeling

Firstly, the "sgns.renmin.word.bz2" is used to embed policy paragraphs into text, converting the text into semantic vectors; secondly, reducing the dimensionality of the text vector by using the UMAP algorithm to solving the problem of dimensionality curse, which algorithm can preserve the high-dimensional information of the text vector as much as possible; then, the HDBSCAN algorithm is used for topic clustering, which clusters all document paragraphs into multiple topic clusters; finally, the c-TF-IDF algorithm is used to extract keywords from each topic cluster. The c-TF-IDF assigns weights to each keyword based on global information, and selects keywords with higher weights as the topic keywords to form a topic representation.

After extracting the topics, in order to clarify, this article divides policy implementation into eight directions based on the "14th Five Year Plan for Digital Economy Development", and based on this, summarizes the extracted topics into eight categories. At the initial stage, the topic only covered some categories, and gradually covered all categories as the stages progressed.

3.3 Calculating Topic Intensity

In the process of generating topics, the topic model will calculate the probability that each document belongs to each topic. By defining the time interval of the document, the sum of the probabilities that all documents within a certain time interval belong to that topic is defined as the topic intensity[11]. The calculation formula is as follows.

$$S_z^t = \frac{\sum_{d=1}^{D_t} \theta_z^d}{D_t} \quad (2)$$

Among them, S_z^t represents the intensity of topic z in the time interval t , θ_z^d represents the probability that the document d belongs to the topic z , and D_t is the total number of documents in the time interval t .

Topic intensity reflects the importance or influence of the topic[12]. The greater the intensity, the more important the topic is in the policy, and the more likely the policy is related to this topic.

3.4 Calculating Topic Heat

The heat of a topic reflects the degree of policy preference for it, and the higher the heat, the more attention the government allocates to it. This article calculates the heat of each topic, and then calculates the heat of each topic category, in order to study the policy path of the government to promote the development of the digital economy.

The heat of a topic is the proportion of documents belonging to that topic among all documents within a time interval t [13]. The specific calculation formula is as follows.

$$I_z^t = \frac{D_z^t}{D_t} \quad (3)$$

Among them, I_z^t is the heat of topic z in the time interval t , D_z^t is the number of documents on topic z in the time interval t , and D_t is the number of all documents in the time interval t .

Topic category heat refers to the sum of topic heat within a category, calculated using the following formula.

$$I_{t,c} = \sum_z I_{t,c}^z \quad (4)$$

$I_{t,c}$ refers to the heat of topic category c in time interval t , $I_{t,c}^z$ and refers to the heat of topic z within topic category c in time interval t .

4 Analysis of Policy Topics

In the topic analysis of each stage, first, combined with the historical background, analyze the extracted topic words of the topic and explain the characteristics of policy stages. Secondly, analyze the hotspots at each stage based on the heat and intensity of the topic.

4.1 Analysis of Policy Topics in the Preliminary Exploration Stage (2003-2014)

Classifying and organizing the topics in this stage, and the relevant data is shown in Table 1.

Table 1. Topics in the Preliminary Exploration Stage

Topic category	Topic representation	Topic intensity	category heat
Industrial digitalization	Industrialization/Modernization/Innovation/Digitalization;	0.0639	8.07%
	Digital culture/Cultural industry/Digital publishing/Publishing;	0.0448	
	Information resources/Integration/Sharing/Personnel/Equipment;	0.0632	
Digital industrialization	Internet of things/Cloud computing /Enterprise/Technology/Products;	0.199	28.86%
	Investment and financing/Big data industry/Enterprise/Guarantee;	0.052	
Digital infrastructure	Network/Communication/Broadband access/Cities/Townships;	0.0695	7.91%
Digitalization of public services	Platform/e-Government/System/Data;	0.0757	46.94%
	Education/Teachers/Digital campus/Information Technology;	0.0767	
	Intelligence/Monitoring/System/Environmental Protection/Monitoring/Traffic;	0.0742	
	Urban management/Digitalization/Coordination/Digital city;	0.097	
	Agriculture/Rural/Information service/Network system;	0.0587	
	Urban and Rural Planning/Digitalization/Municipal Management;	0.0815	
	Civil affairs/Digital/Civil affairs work/Social organization;	0.0494	
Digital security	Information security/Protection/Network/Regulations/Internet;	0.0696	8.22%

Analysis of Policy Topics

In 1997, China elevated informatization to a national strategy and focused on promoting industrial modernization through informatization. The industrial digitization emerged in the gap between the integration of informatization and industrialization. The development needs of informatization and digitization have prompted policies to focus on guiding enterprises to equip professional personnel and hardware equipment,

and to use information technology to break down barriers between various departments and achieve resource integration and optimization. In 2004, Google brought big data technology to people's attention, and China's government seized the opportunity to formulate financial preferential policies to provide financial support for the big data industry. Around 2009, the United States sparked a wave of IoT and cloud computing technologies internationally, while China's policies promoted the development of digital industries related to technology application and product production. At this stage, China's infrastructure is not perfect. The government vigorously promotes the construction of broadband and is committed to achieving universal Internet access. The rise of emerging things is often accompanied by new challenges and problems. In view of the information leakage, network viruses and other thorny problems caused by the popularization of the Internet, the government ensures digital security by standardizing Internet applications and guiding citizens to use the Internet in a standardized manner. The digital category of public services contains the most topics. Public services are related to the livelihoods of people. Therefore, the government regards them as the primary object of digital transformation. Through the application of intelligent monitoring systems and the popularization of information technology education courses, it takes the lead in promoting the digital process in key areas such as urban management, education and teaching, and civil affairs services. At this stage, the digitization of the social economy is still in the exploration period, and the policy mainly focuses on the application of emerging digital technology and the construction of Internet infrastructure and other basic levels, laying the foundation for the development of the digital economy.

Analysis Based on Category Heat and Topic Intensity

Compared to the heat of various thematic categories, the digital heat of public services is as high as 46.94%, followed by the digital industrialization category, reaching 28.86%, while the heat of other categories is relatively low, around 8%. These data indicate that the government attaches special importance to the development of public services and digital industrialization. Observing the intensity of various topics within each category, balancing the intensity of various topics in the field of public services, mainly focusing on social management, government services, and school education, and emphasizing the enhancement of the government's own management and service capabilities through digitization. The digital industrialization includes two topics. The Internet of Things, cloud computing and other technologies have the highest intensity of application and are also the most intense topics in the entire stage. However, the intensity of financing and investment in the big data industry is relatively low, indicating that the development of the digital industry in this stage is in the early stage of technology application, and the industrial scale and industrial chain have not yet formed.

4.2 Analysis of Policy Topics in the Rapid Improvement Stage (2015-2020)

The policy topics in the rapid improvement stage are shown in Table 2.

Table 2. Topics in the rapid improvement stage

Topic category	Topic representation	Topic intensity	Category heat
Industrial digitalization	Industry/Manufacturing/Industrial Internet/Key links/Intelligent manufacturing;	0.0432	20.01%
	Internet+/Overall coordination/Joint force/Mechanism/Implementation;	0.0362	
	Agriculture/Agricultural products/Agricultural machinery/Internet of things/Smart agriculture/Experiment;	0.0308	
	Logistics and distribution/Public information/Smart logistics/Intelligent/Port;	0.0234	
	E-commerce/Cross-border e-commerce/Platform/Demonstration county/Demonstration base;	0.0259	
	Energy/Green/Smart grid/Distributed energy;	0.0256	
	Financial services/Internet+/Securities/Inclusive finance;	0.0253	
	Digital empowerment/Action plan/Healthy development/Compliance;	0.0118	
Digital industrialization	Innovation/Entrepreneurship/Big data/Digital economy/Innovation driven;	0.0427	22.03%
	Artificial Intelligence/Innovation/Technology/R&D/Enterprise/New Generation;	0.0366	
	Talent training/Introduction/Professional/Leading talents/Scientists;	0.0296	
	Financing/Enterprise/Social capital/Special funds/Key projects;	0.0345	
	Leading enterprises/Big data/Software/Investment Attraction/Model innovation;	0.0309	
	Theory/Knowledge/Reasoning/Perception/R&D/Key common technology;	0.0312	
	Smart sensors/R&D/intelligent terminals/chips/integrated circuits;	0.0384	
	Industrial park/Base/Big data/Industry-university-research platform;	0.0352	
	Culture/ Animation /Digital film and television/Creativity/Games/Music;	0.0271	
Network broadband/IPv6/5G/Optical fiber/Interconnection;	0.0356	5.51%	

Digital infrastructure	Data center/Computer room/Standard rack/Green;	0.0314	
Digitalization of public services	Government affairs/Data/Platform/Government services/Sharing;	0.112	40.93%
	Medical health/Elderly care/Internet+/Public health/Platform;	0.0313	
	Transportation/Roads/Integrated law enforcement/Infrastructure/Emergency;	0.0316	
	Education/Intelligence/Educational resources/Smart education/Demonstration project;	0.0249	
	Environment/Monitoring/Natural resources/Pollution sources/Ecology;	0.0295	
	Smart tourism/Information platform/Integration/Marketing/Big data;	0.0264	
	Media/Broadcasting and television/Upgrading/Smart broadcasting and television;	0.0297	
	Emergency/Public safety/Command/Natural disaster/Early warning;	0.0335	
	Public services/People's livelihood/Universal benefits/Equalization/Information benefits the people;	0.0321	
	Smart city/City management/Architecture/Brain/City operation;	0.032	
Digital security	Network and information security/Protection/Level/Data security/System/Risk assessment;	0.0367	4.82%
Data value conversion	Scientific data/Data management/Aggregation/Management system;	0.0236	3.28%
Digital governance	Regulation/System/Market/Negative list/Breaking/Bottom line;	0.0346	3.43%
	Social credit/Information platform/Corporate credit/Credit reporting/Supervision;	0.0242	
	Market supervision/Food and drugs/List/Regulatory matters/Standardization;	0.0333	

Analysis of Policy Topics

The period from 2015 to 2020 is a stage of rapid growth in China's digital economy, with the number of policy topics rapidly increasing as the scope of development expands, reaching out to more fields. Compared to the previous stage, new content related to releasing data value and digital governance has emerged in the policy.

In terms of industrial digitalization, on the basis of the comprehensive popularization of broadband Internet in the previous stage, e-commerce, cross-border e-commerce, rural e-commerce and other platform economies have developed rapidly, driving the

digital transformation of logistics, finance and other related industries. Due to the weakening of traditional economic momentum in China, technological innovation has become a new driving force for economic development. From the topic words of digital industrialization, it can be seen that policies attach great importance to strengthening the ecological support of technological and industrial innovation, ensuring the comprehensive supply of resources such as talents, platforms, and projects, and increasing support for theoretical research and technological research and development, especially basic key technologies such as electronic manufacturing. In terms of digital infrastructure, in addition to continuing to upgrade and optimize network infrastructure, the government has begun to vigorously build data centers to support the storage and calculation of massive data, and enhance the basic support capacity of the digital economy foundation. The digitalization policy of public services has been extended to areas such as healthcare, cultural media, and tourism that are more closely related to people's lives. The digital security policy no longer simply emphasizes protection, but begins to strengthen risk assessment work and build institutional systems to ensure the security of data and information. In the newly emerging thematic categories, the policy related to data value conversion targets the massive amount of data derived from digital technology applications, releases a series of management systems and measures to sort out and store data, and paves the way for subsequent development and utilization of data value. Digital governance mainly utilizes digital technology to enhance market supervision capabilities, upgrade and enrich existing governance methods.

The policies at this stage have further expanded the application scope of digital technology and continued to strengthen the soft and hard support of the digital economy, such as providing resources to assist technology research and development, establishing institutional systems to regulate data management, upgrading and constructing digital infrastructure to improve carrying capacity.

Analysis Based on Category Heat and Topic Intensity

Compared to the heat of various topic categories, the digital heat value of public services in this stage is 40.93%, still ranking first. Next are digital industrialization and industrial digitization, with heat values of 22.03% and 20.01%. The heat of other categories ranges from 3% to 5%. The policy continues to promote the digitalization of public services to a deeper level. However, the significant increase in policy attention to the digital economy industry indicates that the policy always regards economic development as the overall goal. After cultivating a good social digital foundation, it quickly shifts to cultivating emerging industries and promoting the development and growth of the digital economy industry. Observing the intensity of intra category topics, in the digitalization of public services, the topic intensity of building a government data platform is significantly higher at 0.112, indicating that government services are of utmost importance. The intensity of other topics is around 0.03, and the number of topics has increased, indicating that the digitalization of public services is developing comprehensively and evenly. In digital industrialization and industrial digitization, the intensity of topics related to industrial manufacturing digitalization and digital economy innovation and entrepreneurship is relatively high, with intensity values of 0.0432 and 0.0427 respectively, while the intensity of other topics is basically around 0.03. The

policy focuses on promoting the digital transformation of industries, which are the backbone of the national economy, and encouraging mass entrepreneurship, enterprise innovation, and cultivating new forces in the digital economy. This indicates that the government has shifted to promoting the development of the digital economy from the entire industry level, and the "points" that were first tested in the previous stage of policy choices have gradually become "lines".

4.3 Analysis of Policy Topics in the Comprehensive Development Stage (2021-2023)

The policy topics in the rapid improvement stage are shown in Table 3.

Table 3. Topics in the Comprehensive Development Stage

Topic category	Topic representation	Topic intensity	Category heat
Industrial digitalization	Industry/Industrial Internet/Network collaboration/Benchmark project;	0.0434	20.71%
	Agriculture/Agricultural machinery/Precision monitoring/Early warning/Refinement/Intelligent;	0.0284	
	Financing/Financial institutions/Financial services/Enterprises/Investments/Banks;	0.0346	
	Energy/Oil & Gas/Smart/Grid/Consumption/Regulation;	0.0289	
	5G/Commercial/Rural/Rural commercial/Agricultural supermarket/Internet;	0.0208	
	Waterway/Ship/Smart Waterway/Shipping/Port/Navigation;	0.0221	
Digital industrialization	Chip/Material/Base/Manufacturing/Sensor/Semiconductor;	0.0329	11.17%
	Artificial Intelligence/Algorithm/Model/Technology/Blockchain;	0.0351	
	Talent/Training/Cultivation/Digital skills/Digital economy talent;	0.0303	
	Digital economy/R&D/Core technology/High-end talents/Team/Enterprise;	0.0355	
	Digital/Ecology/Digital economy/Future industry;	0.0334	
Digital infrastructure	Data center/Computing power /New type/Layout/Green/Computing/Integration;	0.0336	11.00%
	Transportation/Infrastructure/Vehicle-Road Collaboration/Perception/Intelligence;	0.0283	
	Innovation/Technological Innovation/Laboratory/Innovation Center/R&D Platform;	0.0391	

	Network/5G/IPv6/Dual gigabit/Backbone network/Capacity expansion and upgrade;	0.0339	
Digitalization of public services	Government data/Sharing/Integration/Platform/System;	0.0889	28.53%
	Cultural tourism /Resource advantage/Platform/Live broadcast/Quality/Digital cultural center;	0.0314	
	Education/School/Students/Digital campus/Network/Smart learning/Courses;	0.0268	
	Government services/Processing/Electronic certificates/One thing/Inter-provincial processing;	0.0440	
	Public services/Elderly care/Disabled people/People's livelihood;	0.0360	
	Natural Resources/Ecology/Monitoring/Ocean/Forestry/Soil;	0.0240	
	City brain/Digital twin/City operation/Digitalization;	0.0352	
	Emergency/Civil Air Defense/Command/Natural Disasters/Risk/Monitoring and Warning/Meteorology;	0.0284	
	Medical/Health/Insurance/Diagnosis and Treatment/Internet Hospital/Smart Equipment;	0.0260	
	Rural/Digital rural/Administrative village/Rural governance/Public safety;	0.0297	
	Digital security	Network security/Public data/Personal information/Law/Classification;	
Data value conversion	Data trading/Trading venue/Pricing/Standards/Registration/Compliance;	0.0375	1.82%
Digital governance	Supervision/Credit/Special equipment/Drugs/Accurate portraits/Image perception;	0.0350	10%
	Standardization/National standards/Industry standards/Digital countryside/Data development;	0.0379	
	Economic operation/Analysis/Research and judgment /Economic adjustment/Economic governance;	0.0376	
	Digital government/Operation/Mechanism/Coordination/Sustainable and healthy development;	0.0391	
	ASEAN/Cross-border trade/Information Port/Digital Silk Road;	0.0316	3.48%

International cooperation and exchange	Publicity/Digital Economy/Conference/International/Exchange/Forum/Summit/Platform;	0.0314	
--	--	--------	--

Analysis of Policy Topics

At this stage, the number of topics has not increased, but the keywords are sufficient to demonstrate policy changes and new topics for international exchange and cooperation have emerged. Since entering the 14th Five Year Plan period in 2021, policies have clearly tended towards refinement, and the targeting of topic words has become stronger. In terms of industrial digitalization, the policy focuses on the construction of industrial Internet to improve network synergy. In agriculture, monitoring and calculation are used to accurately complete each step of the work based on soil, weather, and other conditions. In addition, the topic words indicate that the digital economy has penetrated into complex high-end fields such as waterway construction, shipbuilding, and port logistics. In terms of hardware in the digital industry, the government encourages the research and production of high-end chips, new display materials, and third-generation semiconductors; In terms of software, accelerating the innovation of algorithms and models related to artificial intelligence, blockchain and other technologies to improve independent innovation capabilities and achieve technological self-reliance and self-improvement. The topic of digitalization of public services is still the most prevalent and continues to develop towards deeper levels. The digital security guarantee system has entered the track of legalization. In terms of data value conversion, the policy focus has shifted from data management to data trading, with a focus on addressing issues such as transaction rights confirmation, pricing, and regulations. The number of topics in digital governance has increased, and the focus of the content has shifted to addressing the problems that arise in the digital economy. Through measures such as monitoring economic operations and setting development standards, the risks and negative social impacts of the digital economy have been reduced. The emerging international cooperation and exchange focus on cross-border digital trade with ASEAN, and through international conferences, exchange forums and other platforms, keep up with the forefront of the digital economy, promote the high-end and high-quality development of China's digital economy.

Analysis Based on Category Heat and Topic Intensity

Comparing the heat of topic categories, it was found that the heat of digitalization of public services decreased to 28.53%, while the heat of digitalization of industries remained at 20.71%. The heat of digital infrastructure, digital security, and digital governance all increased to 10% or more. This indicates that as the digital economy enters a stage of comprehensive development, policy attention is gradually balanced, no longer focusing solely on social services and economic industries, but increasingly on digital security and governance capabilities to ensure the sustainable and healthy development of the digital economy. The heat of data value conversion is only 1.82%. The reason for this is that the three-year action plan for "Data Factor ×" (2024-2026) jointly released by the National Data Bureau and 17 other departments in 2024 was not

included as a sample after the experiment, resulting in a low heat of data factor value. It is foreseeable that in the future, policies related to the value of data factors will further promote the development and utilization of data resources and the release of data factor value in China, providing a strong driving force for the high-quality development of the digital economy. Observation shows that there is not much difference in intensity values between topics during this stage, and overall, they are relatively low. According to the logic of calculating topic intensity, the reason is that the policy in the comprehensive development stage has a high degree of integration and coordination in task arrangements in various fields, and the content of each policy paragraph is comprehensive. However, a topic only focuses on one aspect, resulting in a low probability of the policy belonging to a certain topic. It can be seen from this that the development situation of China's digital economy is becoming increasingly complex, with deeper levels of development, and policies paying more attention to the coordinated development of various fields of the digital economy.

5 Conclusions

This article explores the evolution of policy topics in stages, and finds that policies take the digitization of public services and the application of digital technology as entry points, radiating and driving the digital transformation of industries such as industry and agriculture. With the deepening application of digital technology, emphasis is placed on cultivating emerging industries and continuously improving digital infrastructure construction. At present, policies focus on the key and difficult points in the development of the digital economy, as well as the emerging new problems and demands in the development process, to build a digital ecosystem support, and to strengthen security and governance guarantees. Meanwhile, as the development stage leaps, the scope of policy topics continues to expand, gradually covering from five categories to eight categories. The policy content is developing towards refinement and intensification, emphasizing the overall coordination of the development pattern of the digital economy. At present, China's digital economy policies have formed a policy development pattern of overall promotion, precise efforts, and individual breakthroughs in parallel.

According to the analysis, data utilization development and data market construction will be the current and future core policy concerns. At the same time, the heat of digital security and digital governance continues to rise, and the topic indicates that policies are increasingly emphasizing the problems inherent in the development of the digital economy. The increasingly complex situation of digital economy development will promote policies to continuously strengthen the digital economy security system and governance capabilities.

Acknowledgment

Thanks for the support of Taishan Scholars Program [NO. tsqn202306253].

References

1. Yang Qiaoyun, Qiao Yingying, Liang Shilu. Research on Digital Economic Policy of Provincial Government Based on Compatibility of Policy Goals and Policy Tools [J]. *Reform of Economic System*, 2021(03):193-200.
2. Tan Haibo, Guo Yitong. Balance and Adaptability: A Quantitative Research on Digital Economic Policies of Local Governments—An Analytical Perspective Based on "Policy Tool-Policy Preference" [J]. *Journal of Guangxi Normal University (Philosophy and Social Sciences Edition)*, 2022, 58(05):100-116.
3. Lei Hongzhu, Wang Qian. A Quantitative Study on Chinese Local Government Digital Economy Policy Texts [J]. *Journal of Technical Economics & Management*, 2022(05):91-94.
4. Chen Mei, Sun Ruiqian. Quantitative Analysis of China's Provincial Digital Economy Policies from the Perspective of Policy Tools: Based on LDA Topic Social Network Analysis [J]. *Journal of Intelligence*, 2023, 42(11): 174-182.
5. Bu Lingtong, Zhang Jiawei. Quantitative Evaluation of Digital Economy Policy Based on PMC Index Model [J]. *Statistics & Decision*, 2023, 39(07): 22-27.
6. Wang Weiguang, Song Hongling. Quantitative Evaluation of Policy Tools for Digital Economy Support: Text Analysis Based on Inter-Provincial Level Policy [J]. *Forum on Science and Technology in China*, 2023(06):97-107.
7. Cai Dongsong, Chai Yilin, Tian Zhixiong. Quantitative Evaluation of Jilin Province Digital Economy Policy Based on PMC Index Model [J]. *Information Science*, 2021, 39(12): 139-145.
8. Blei DM, Ng AY, Jordan M I. Latent dirichlet allocation[J]. *Journal of machine learning research*, 2003, 3(Jan): 993-1022.
9. Grootendorst M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure[J]. *arXiv preprint arXiv:2203.05794*, 2022.
10. Abuzayed A, Al-Khalifa H. BERT for Arabic topic modeling: An experimental study on BERTopic technique[J]. *Procedia computer science*, 2021, 189: 191-194.
11. Sun Ruiying, Chen Yihong. Research on the Hot Topics and Future Research Orientation of Domestic Smart Reading Based on LDA Model[J]. *Library Development*, 2023, (03):82-93+103.
12. Gao Nan, Gao Jiaqi, Chen Hongpu. Emerging Technology Identification Method and Evolution Path—Take the Field of Integrated Circuits as an Example [J]. *Information Science*, 2023, 41(03):127-135+172.
13. Wu Ruipeng, Li Yongnan, Liu Shuai, etc. DTM-Based Analysis of Hot Topics and Evolution of American Artificial Intelligence Strategy [J]. *Journal of Intelligence*, 2023, 42(12):134-143.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

