



ARTICLE



<https://doi.org/10.1057/s41599-025-06046-6>

OPEN

Trends and perceptions of youth entrepreneurship in China: a mixed-text mining analysis

Xin Feng¹, Yiqiao Liu¹, Linan Zhang²✉ & Shuting Zhang²

Youth entrepreneurship is a vital force for China's economic growth, educational reform, and rural revitalization. This study investigates the trends and perceptions of youth entrepreneurship in China using a mixed-text mining approach. We first conduct a bibliometric analysis of 260 academic papers from the CNKI&WOS database to identify key research themes and trends. We then apply Latent Dirichlet Allocation (LDA) topic modeling to analyze 7493 entries from Zhihu, a popular Chinese Q&A platform, to understand public perceptions of college student entrepreneurship. Our findings reveal a dynamic interplay between academic research and societal discourse. Key themes in academic literature include entrepreneurial education, policy support, and rural entrepreneurship. The LDA analysis of Zhihu data highlights concerns about entrepreneurial capabilities, business operations, and industry selection. We integrate these findings within an innovation ecosystem framework, identifying three core influences: subjective entrepreneurial capability, the macro-policy environment, and objective social resources. Notably, subjective capability is the most influential factor, accounting for 81.38% of the total topic proportion in the Zhihu data, emphasizing its critical role in entrepreneurial success. This research underscores the need for universities to enhance entrepreneurial skills training, for governments to provide robust policy support and improve policy awareness, and for industries to foster innovation and resource accessibility.

¹School of Management, Shijiazhuang Tiedao University, Shijiazhuang, China. ² School of Marxism, Shijiazhuang Tiedao University, Shijiazhuang, China.
✉email: 534894264@qq.com

Introduction

Youth entrepreneurship is increasingly recognized globally as a pivotal engine for fostering economic growth, driving innovation, and facilitating social progress. In China, against the backdrop of national strategies emphasizing innovation-driven development and rural revitalization, youth entrepreneurship plays an exceptionally vital role in shaping the economic landscape and promoting sustainable development (Guobin, 2023; Yan, 2024). It is seen as a crucial pathway for youth employment, talent utilization, and activating grassroots economic vitality. However, despite significant policy support and a burgeoning entrepreneurial ecosystem, young entrepreneurs in China continue to face considerable challenges. Recent reports highlight issues such as limited access to funding, intense market competition, lack of practical experience, and the pressures arising from high graduate employment rates, which paradoxically also push more young people towards entrepreneurial ventures (Augustino et al., 2023). Understanding the dynamic interplay of factors influencing youth entrepreneurship in this complex environment is therefore essential. This study investigates the trends in and perceptions of youth entrepreneurship in China through a mixed-method text mining analysis, examining both academic literature and online public discourse.

Prior academic research has extensively explored various facets of youth entrepreneurship. Studies have delved into individual entrepreneurial traits and capabilities (Svotwa et al., 2022; Campo et al., 2022), the influence of educational programs and policy support (Augustino et al., 2023), and the impact of social networks and industry ecosystems (Chuan et al., 2021). Chinese scholarship has specifically examined the entrepreneurial environment within China, focusing on different youth demographics (e.g., rural youth, college students) and the role of government initiatives and social platforms (Zeping et al., 2022; Jiaqing and Deshui 2022). While this body of work provides a foundational understanding of the theoretical and structural aspects of youth entrepreneurship, there appears to be a gap in systematically connecting these academic perspectives with the ground-level concerns, questions, and real-world challenges articulated by young people engaged in or contemplating entrepreneurship. The divergence between the themes prominent in scholarly analysis and the issues discussed within public online forums represents a critical area where insights are needed to better align support mechanisms with actual needs.

The motivation for this study is multi-faceted, arising from both significant societal relevance and crucial academic imperatives. Firstly, acknowledging the critical role of youth entrepreneurship for China's sustainable development and youth empowerment, a primary driver is to address the substantial real-world challenges faced by young entrepreneurs, such as resource limitations and intense competition (as highlighted by recent reports), which hinder their success and impact. Secondly, from a methodological perspective, this research is motivated by the need to employ rigorous analytical techniques, specifically quantitative text mining integrated within the innovation ecosystem framework, to comprehensively understand the complex factors influencing youth entrepreneurship. This approach aims to provide a deeper, empirically grounded understanding of the entrepreneurial landscape. Thirdly, a distinct academic motivation arises from the observed disparity between the prevalent themes in existing academic literature and the authentic concerns and perceptions expressed by young people in public online discourse. Addressing this gap is essential for ensuring that both scholarly insights and practical support initiatives are effectively aligned with the actual needs and experiences of aspiring entrepreneurs.

Aligned with this motivation, this study aims to achieve two main objectives: (1) To map the key research themes and evolving

trends in academic literature on Chinese youth entrepreneurship using bibliometric analysis. This objective seeks to systematically outline the intellectual structure and historical trajectory of the field, revealing what topics have been prioritized by scholars. (2) To identify and analyze the dominant perceptions, interests, and concerns of college students regarding entrepreneurship, as expressed in online public discussions, through LDA topic modeling. This objective is designed to tap into the authentic, ground-level discourse to uncover the practical challenges and real-world needs articulated by aspiring young entrepreneurs. Our research is novel in its mixed-method approach, which uniquely integrates these two distinct data sources—formal academic publications and informal online discourse—to provide a comprehensive and comparative view.

The contributions of this study are multifaceted. To the literature, it offers a novel framework for analyzing youth entrepreneurship by explicitly bridging academic research trends and public perceptions, enriching theoretical models like the innovation ecosystem by highlighting the dynamic interplay of influencing factors from a dual perspective. For practitioners and educators, the findings provide data-driven insights into the actual challenges and concerns faced by young entrepreneurs, which can inform the design of more relevant training programs and support services. Policymakers can leverage this research to refine existing policies, ensuring they are more responsive to the real needs identified in both academic study and public discourse. Ultimately, by shedding light on the complex dynamics of youth entrepreneurship from multiple angles, this study contributes to fostering a more effective and supportive entrepreneurial ecosystem in China, thereby having positive social implications.

The remainder of this paper is structured as follows. Section “Literature review” provides a brief overview of the relevant literature. Section “Data sources and research design” details the data collection and research methodologies. Section “Bibliometric analysis” presents the results of the bibliometric analysis and LDA topic modeling. Section “Empirical analysis based on LDA theme analysis” discusses the key findings within the context of the innovation ecosystem framework. Finally, Section “Influencing factors in entrepreneurial ecosystem” summarizes the conclusions, offers practical recommendations, acknowledges limitations, and suggests directions for future research.

Literature review

Current state of youth entrepreneurship research in China from the perspective of innovation ecosystem. The concept of the innovation ecosystem originates from the theory of system evolution in ecology, with its core tenet being that innovation is no longer an isolated act but rather the product of collaborative competition among diverse entities within a complex network (Moore, 1993). Adner (2006) further emphasized the importance of value co-creation and interdependent relationships among participants within an ecosystem. Chinese scholars have interpreted and applied the innovation ecosystem theory within the local context. The innovation ecosystem theory emphasizes synergistic cooperation among various actors within the system, achieving sustainable development through resource sharing and complementary capabilities (Mingfeng and Yafang, 2021; Tie and Linling, 2022). It also highlights the dynamic evolutionary process of the system, positing that interactions among internal actors and changes in the external environment jointly drive the system's evolution (Ruzhong et al., 2011).

In China's societal development, the policy environment plays a crucial supporting role within the innovation ecosystem. By establishing robust innovation policy systems, enhancing

intellectual property protection, and fostering a fair competitive environment, the government can create fertile institutional ground for entrepreneurship among youth from Hong Kong, Macao, and those returning to their hometowns (Aiping and Xiaoyun, 2022). However, research indicates that the application of policy tools within China's youth entrepreneurship policy system exhibits diverse yet unbalanced characteristics. Persistent issues include the insufficient use of demand-driven policy tools, significant internal biases in tool structure, and limited variety in the categories of tools employed (Guobin, 2023). Yuhua and Yongzhi (2019) also pointed out that the imbalanced development of local innovation ecosystems leads to significant disparities in entrepreneurial opportunity structures across different regions, necessitating coordinated development through regional innovation policies.

Resources are pivotal to the success of youth entrepreneurship. Entrepreneurs' human capital, social networks, financial resources, and knowledge and cultural resources constitute a multi-dimensional framework for entrepreneurship research. Tommaso et al. (2014) highlighted that young entrepreneurs often face resource disadvantages, relying on incubators, university science parks, and government entrepreneurial platforms for resource integration. Zhenxing (2016), studying China's high-tech zones, found that a sound entrepreneurial finance system (encompassing technology finance, industrial finance, and intellectual property finance) can significantly enhance innovation capabilities. Concurrently, Weidong and Junxia (2021) proposed that cross-border resource integration capability is a crucial quality for young entrepreneurs to distinguish themselves within the innovation ecosystem.

The innovation ecosystem emphasizes interaction and cooperation among its actors. Young entrepreneurs can compensate for their lack of resources and experience by establishing cooperative relationships with university research teams, large enterprises, and government departments, thereby achieving technological innovation and market expansion. Zeping et al. (2021), considering China's large number of startups, active venture capital, and progressively improving entrepreneurial quality and ecosystem, proposed targeted recommendations from the perspectives of the government, the Communist Youth League, and universities. Yinghui (2020) further analyzed the mediating effect of social networks on the success of youth entrepreneurship, suggesting that strong ties facilitate access to key resources, while weak ties are beneficial for discovering innovation opportunities. Among the multiple actors, the subjective capability of the entrepreneur or entrepreneurial team holds a prominent position. The need for control and achievement are the most significant factors influencing young people's decision to bear employment risks (Ana and Nelu, 2019), while entrepreneurial self-efficacy and attitude positively impact entrepreneurial capability and help stimulate entrepreneurial thinking (Svotwa et al. 2022).

Within the innovation ecosystem, young entrepreneurs must not only acquire resources but also constantly adapt to ecological changes. Pillai and Ahamat (2018) discloses a variety of inter-related factors within the social network ecosystem that stimulate and sometimes stifle youth entrepreneurship, primarily through the active agency of social-cultural capital. Young entrepreneurs need to possess ecological adaptability—the ability to rapidly adjust their strategies and organizational structures in response to technological iterations, market fluctuations, and policy adjustments. Different categories of entrepreneurs have varying requirements for entrepreneurial conditions. Youth returning to their hometowns for entrepreneurship are most concerned about infrastructure (Longfei, 2019; Xiaohong, 2017) and the lack of entrepreneurial experience and funding (Yuan et al., 2022). In

contrast, Taiwanese youth venturing into mainland China place greater importance on marketing and business development, venture capital support, investment and financing, interpersonal resources, and convenient administrative services (Chang and Lai, 2021).

Application status of bibliometrics and text mining in entrepreneurship research. In recent years, bibliometric and text analysis methods have been widely applied in the field of entrepreneurship research. These approaches not only offer new perspectives for revealing the evolution of research themes, hotspot trends, and disciplinary structures but also hold significant importance for understanding the gap between theory and practice. Scientometrics can serve as a systematic tool for understanding the knowledge base in management research (Timothy et al., 2019; Elisa et al., 2021; Toschi et al., 2023; Emily et al., 2025). Among these, numerous scholars have utilized bibliometric analysis to map the knowledge landscape of social entrepreneurship, clearly demonstrating the evolutionary process of different research themes (Meidong et al., 2022; Fang and Zhu, 2020). Furthermore, Francisco et al. (2023) combined social network methods with bibliometric approaches in innovation research, effectively identifying core issues within the research domain. Text mining techniques have attracted increasing attention due to their ability to extract meaningful information from large-scale unstructured data (Gosztonyi et al. 2022; Haiyang et al. 2024). Text analysis has also contributed to refining developmental issues and offering breakthroughs for various entrepreneurial problems such as policy entrepreneurship (Kim et al., 2025), sustainable entrepreneurship (Zhou et al., 2022), and institutional entrepreneurship (Maria and Stefano, 2017). These studies provide a solid foundation for our methodological choices but also indicate that relying solely on traditional literature reviews and single-perspective insights may be insufficient to grasp the rapidly changing entrepreneurial practice environment.

However, while scholars have employed bibliometric or text mining methods to review entrepreneurship research, studies that combine both methods to deeply characterize the perceptions and needs regarding youth entrepreneurship from different domains remain relatively scarce. Moral-Muñoz et al. (2020), in assessing the feasibility of integrating bibliometrics and text mining, pointed out that combining the two methods can more effectively identify potential research gaps and future development directions. Building upon this, the present study adopts academic and public perspectives as entry points, conducting bibliometric and text mining analyses under the theoretical guidance of the innovation ecosystem framework. Our aim is to systematically analyze the current status of youth entrepreneurship in both domains from a systemic perspective, seeking to expand new analytical dimensions beyond existing work. The objective is to reveal the intersections, differences, or potential disconnects between academic research trends and actual public perceptions.

Data sources and research design

Data sources. This study investigates the trends and patterns in academic research on youth entrepreneurship in China through a bibliometric analysis of scholarly publications. We collected data from two key databases to ensure comprehensive coverage: China National Knowledge Infrastructure (CNKI) and the Web of Science (WOS) Core Collection.

To ensure the high quality and academic rigor of the included literature, our search was specifically limited to journals indexed in the Chinese Social Sciences Citation Index (CSSCI). We searched the CNKI database for papers published between January 1, 1994, and May 1, 2025. This comprehensive timeframe

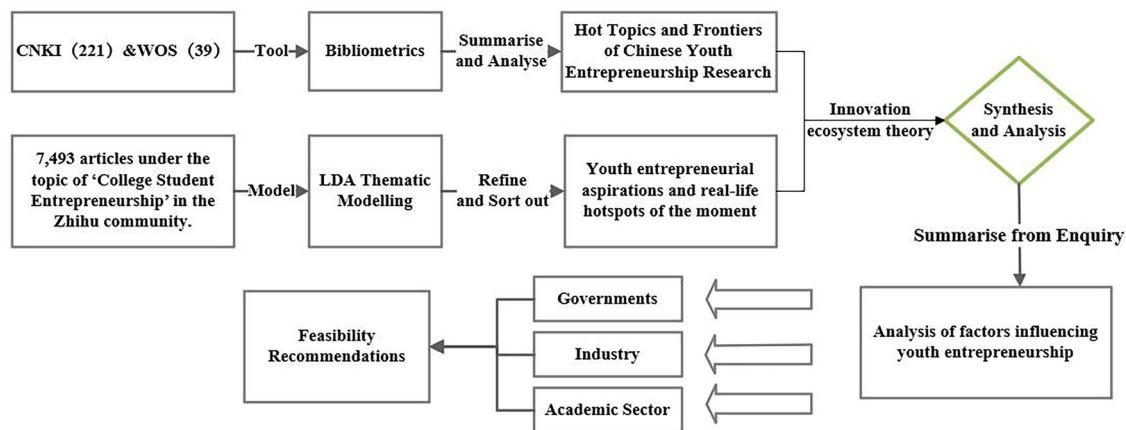


Fig. 1 Research design diagram.

was chosen to capture the historical evolution of research on youth entrepreneurship in China from its early stages through various policy initiatives and societal developments up to the present. The search focused on the theme “youth entrepreneurship” in the title or keywords, a focused strategy designed to capture highly relevant academic discourse within these key journal collections. Following the initial retrieval, a preliminary screening based on relevance of title and abstract was conducted, followed by a subsequent rigorous review to exclude non-academic publications (such as news reports, conference abstracts, and book reviews) and ensure the included studies were empirical or theoretical research directly focused on youth entrepreneurship. This multi-stage filtering process, combined with the journal indexing criteria, was implemented to ensure the high quality and direct relevance of the dataset to our bibliometric analysis. Ultimately, 221 academic papers from CNKI were included in the dataset after this rigorous filtering process.

To complement the CNKI data and include relevant research by Chinese scholars published internationally, we also conducted a supplementary search on the Web of Science (WOS) Core Collection. WOS was chosen for its extensive coverage of high-impact international journals across various disciplines, providing a global perspective and capturing research potentially missed in a purely domestic database search. We used the keywords “Youth” AND “Entrepreneurship” AND “China” in the topic search (TS). The inclusion of “China” was specifically added to narrow the focus to studies relevant to the Chinese context, as broader searches yielded a high volume of globally-focused papers not directly pertinent to our study. The search was limited to Article and Review Article types to focus on peer-reviewed, full-length research findings. The timeframe for the WOS search was the same as CNKI, from January 1, 1994, to May 1, 2025, to provide a parallel historical scope of internationally published research related to the Chinese context. This search yielded 39 relevant articles after relevance screening based on title, abstract, and keywords, using the same relevance criteria as the CNKI filtering.

In total, our academic bibliometric analysis dataset comprises 260 academic papers (221 from CNKI and 39 from WOS), providing a robust and more comprehensive foundation for mapping the research landscape.

Additionally, to understand public perceptions and real-world concerns regarding youth entrepreneurship, we collected data from Zhihu, a popular Chinese online Q&A platform. Zhihu was selected for its large user base, particularly among young intellectuals, and its format which facilitates detailed questions and answers, providing rich, authentic user-generated content

that can capture diverse perspectives and experiences regarding entrepreneurship. Using a web crawler tool, data from 7493 entries from the “College Student Entrepreneurship” topic on the Zhihu platform was collected. This specific topic was chosen because college students represent a significant and active demographic in youth entrepreneurship discourse and practice in China, making them a relevant and substantial source for exploring prevalent public perceptions. While acknowledging this focus limits the data to primarily the college student perspective, it provides deep insight into this crucial group. This dataset consists of user questions (capturing entrepreneurial needs, challenges, and points of confusion) and answers (indicating key areas of interest, shared experiences, and practical advice). The use of a web crawler tool was employed for its efficiency in collecting a large volume of publicly available data from the platform.

Research design. The study employs a mixed-method text mining approach, combining bibliometric analysis of academic literature and LDA topic modeling of social media data, to investigate youth entrepreneurship from both academic and public perspectives. This dual analysis is conducted and interpreted within the conceptual lens of the innovation ecosystem framework, providing a structured basis to understand the interplay of various influencing factors identified from both data sources. Using frequency distribution charts, keyword co-occurrence maps, and clustering diagrams, the bibliometric analysis provides an intuitive understanding of the keywords and themes in the academic articles, identifying research hotspots and trends in Chinese youth entrepreneurship.

The Zhihu data was processed using the LDA model to uncover public perceptions and concerns around youth entrepreneurship. By comparing academic literature with real-world youth entrepreneurship needs, this analysis reveals discrepancies between theoretical research and practical challenges. This comparison highlights gaps in current research and helps identify critical factors influencing youth entrepreneurship, leading to practical recommendations for further development. The research design is illustrated in Fig. 1.

Bibliometric analysis

Keyword analysis. This section presents the results of the bibliometric analysis of 260 academic papers on youth entrepreneurship retrieved from the CNKI(221) & WOS(39) database. We focus on identifying key research trends, hotspots, and thematic clusters.

Table 1 Hot Topics in Chinese Youth Entrepreneurship Research (1994–2025).

Frequency	Centrality	Year	Keyword	Frequency	Centrality	Year	Keyword
23	0.42	2003	Youth	5	0.02	2012	Entrepreneurial Education
21	0.27	2007	Entrepreneurship	5	0.09	2014	Entrepreneurial Environment
21	0.29	2008	Youth Entrepreneurship	4	0.04	2014	Social Entrepreneurship
10	0.10	2017	Taiwanese Youth	4	0.03	2014	Innovation
9	0.12	2019	Rural Revitalization	4	0.04	2011	College Students
8	0.14	2009	Innovation and Entrepreneurship	4	0.08	2022	Returning Youth
8	0.05	2010	Employment and Entrepreneurship	3	0.01	2016	Entrepreneurial Policy
6	0.13	2005	Rural Youth	3	0.08	2005	Entrepreneurial Spirit
6	0.15	2013	Influencing Factors	3	0.00	2016	Entrepreneurial Capability
5	0.07	2016	Entrepreneurial Willingness	3	0.00	2019	Urban-Rural Integration

The most representative 20 keywords are shown in Table 1. This research primarily revolves around the two main themes, “Youth” and “Entrepreneurship.” The concept of “Youth Entrepreneurship” essentially took shape around 2008. In July 2003, the 15th National Congress of the Communist Youth League proposed the initiative to “promote youth employment and entrepreneurship.” In November, the Central Committee of the Communist Youth League, the All-China Youth Federation, and the All-China Federation of Industry and Commerce jointly launched the “China Youth Entrepreneurship International Program.” Against this backdrop, youth entrepreneurship garnered significant public attention, leading to a large number of related research papers. Relevant research and concepts emerged extensively between 2010 and 2020, with diversified development directions including entrepreneurial groups (Taiwanese youth, rural youth, returning youth), entrepreneurial capabilities (entrepreneurial willingness, innovation), entrepreneurial pathways (social entrepreneurship, urban-rural integration), and the entrepreneurial environment (entrepreneurial policy, innovation and entrepreneurship), among others.

Cluster analysis. To identify major research themes, we performed a cluster analysis of the keywords using CiteSpace. The clustering results show that Modularity (Q) is 0.8016, which is greater than the critical value of 0.3, indicating that the network community structure is significant and the clustering effect is good; Weighted Mean Weight Mean Silhouette (S) is 0.9371, which is not only much greater than the critical value of 0.5, but even greater than 0.7, which indicates that the clustering results have a high degree of confidence. The keyword clustering results have a significant structure and good clustering effect, which helps to analyse and grasp the general characteristics and evolutionary context of youth entrepreneurship research. Figure 2 visualizes the keyword clustering, and Table 2 summarizes the nine major themes identified:

In Table 2, ‘Mean (Year)’ represents the average publication year of the nodes within a cluster, reflecting the approximate temporal characteristics of that cluster. The ‘log-likelihood ratio (LLR)’ indicates the difference between the frequency of a word appearing within the ‘current cluster’ and its frequency elsewhere in the ‘entire dataset.’ It is used to measure the importance of a specific term to a particular cluster. The ‘p-level’ signifies whether the phenomenon of ‘this word being particularly high-frequency in this cluster’ could be due to chance. The p-level is used to test the statistical significance of the aforementioned LLR results, with a value less than 0.05 generally considered statistically significant. The keyword clustering results exhibit a significant structure and good clustering effect, which aids in analyzing and understanding the overall characteristics and evolutionary context of youth entrepreneurship research. The following is a specific analysis of

four types derived from the clusters in Chinese youth entrepreneurship research:

- (1) **Entrepreneurial Groups**
This category primarily focuses on distinct groups within the youth population. It includes two major clusters (#3 Returning Youth, #5 Youth Groups) and characteristic nodes from other clusters (including “Taiwanese Youth” “Hong Kong and Macao Youth” “International Students” “College Students”). Youth are the most vibrant force in the entire society, and cultivating their entrepreneurship is highly beneficial for the full utilization of talent resources. In the context of rural revitalization, the policy system for youth returning to their hometowns to start businesses currently has clear objectives, rich content, and diverse methods (Guobin, 2023). Among youth, the most distinct groups are those from Hong Kong, Macao, and Taiwan. As a new generation leading the future direction of cross-strait relations, their willingness to start businesses in the mainland contributes significantly to deepening cross-strait integration (Cheyuan and Zeguang, 2018; Cheyuan et al. 2019). Issues of identity perception and misalignment with the mainland environment are factors influencing their employment and entrepreneurship decisions in the mainland. Therefore, strengthening youth exchange and cooperation and building an interactive cycle mechanism for Taiwanese youth to seek employment and entrepreneurship in the mainland and reconstruct their “Chinese identity” will help achieve positive symbiotic development among young people.
- (2) **Entrepreneurial Capabilities**
This category primarily focuses on the comprehensive capabilities of entrepreneurial groups. It includes two cluster clusters (#6 Entrepreneurial Willingness, #8 Soft Skills) and characteristic terms such as “organization” “intelligence” and “collective action”. The willingness of a partner may bring more social rewards than the ability to do so (Qiang et al., 2024). Therefore, firm entrepreneurial willingness is a prerequisite for cultivating entrepreneurial capability. Entrepreneurial self-efficacy and entrepreneurial attitude have a positive impact on entrepreneurial capability and help stimulate entrepreneurial thinking. Subjective capability is one of the core elements of entrepreneurship. The development of the times places higher demands on entrepreneurs; enterprise scale requires organizational ability from entrepreneurs, and information transmission requires intelligent processing ability from entrepreneurs.
- (3) **Entrepreneurial Pathways**
Entrepreneurial pathways include “#4 Innovation and Entrepreneurship” “Rural Factory” and “Policy Implementation”. These characteristic terms focus on the success and opportunities for entrepreneurs in their choice of

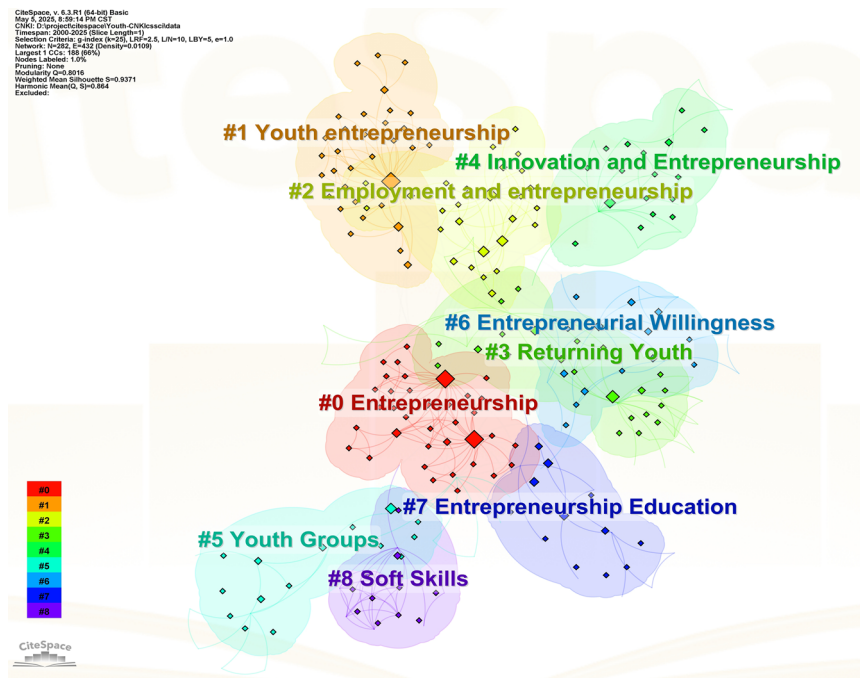


Fig. 2 Keyword clustering of youth entrepreneurship research in China.

Table 2 Youth entrepreneurship research clusters (1994–2025).

Cluster	Mean (Year)	Top Terms (log-likelihood ratio, p-level)
#0 Entrepreneurship	2013	Entrepreneurship (17.6, 1.0E-4); Innovation (8.62, 0.005); Youth (7.63, 0.01); Youth Entrepreneurship (4.01, 0.05); Employment (2.84, 0.1)
#1 Youth Entrepreneurship	2014	Youth Entrepreneurship (17.5, 1.0E-4); Entrepreneurial Environment (7.38, 0.01); Organization (3.66, 0.1); Intelligence (3.66, 0.1); Collective Action (3.66, 0.1)
#2 Employment and Entrepreneurship	2017	Employment and Entrepreneurship (11.64, 0.001); Taiwanese Youth (11.64, 0.001); Mainland Entrepreneurship (7.69, 0.01); Identity Perception (3.81, 0.1); COVID-19 Pandemic (3.81, 0.1)
#3 Returning Youth	2021	Returning Youth (13.66, 0.001); Rural Revitalization (9.01, 0.005); Living Arrangement (4.45, 0.05); Entrepreneurial Risk (4.45, 0.05); Rural Factory (4.45, 0.05)
#4 Innovation and Entrepreneurship	2018	Innovation and Entrepreneurship (16.19, 1.0E-4); Hong Kong and Macao Youth (5.24, 0.05); Social Ecology (5.24, 0.05); Policy Implementation (5.24, 0.05); Incubator (5.24, 0.05)
#5 Youth Groups	2012	Youth Groups (10.98, 0.001); Entrepreneurial Culture (10.98, 0.001); Apprenticeship and Students (5.4, 0.05); Countermeasures (5.4, 0.05); Problems (5.4, 0.05)
#6 Entrepreneurial Willingness	2020	Entrepreneurial Willingness (12.79, 0.001); Digital Literacy (6.26, 0.05); Youth Individual (6.26, 0.05); Living Arrangement (6.26, 0.05); Digital Era (6.26, 0.05)
#7 Entrepreneurship Education	2014	Entrepreneurship Education (10.31, 0.005); Youth Development (10.31, 0.005); College Students (10.31, 0.005); Entrepreneurial Policy (5.08, 0.05); International Students (5.08, 0.05)
#8 Soft Skills	2014	Soft Skills (5.79, 0.05); Chinese Youth (5.79, 0.05); Chinese Education (5.79, 0.05); Entrepreneurial Spirit (5.79, 0.05); China (5.79, 0.05)

entrepreneurial pathways. Optimizing the youth innovation and entrepreneurship environment has always been a key focus of government work. Since the 18th National Congress of the Communist Party of China, the comprehensive implementation of the rural revitalization strategy, coupled with “mass entrepreneurship and innovation” policy measures, has further activated youth innovation and creativity.

(4) Entrepreneurial Environment

This theme includes two clusters (#2 Employment and Entrepreneurship, #7 Entrepreneurship Education) and characteristic nodes such as “Rural Revitalization” “Digital Era” and “COVID-19 Pandemic”. This cluster focuses on the multi-dimensional entrepreneurial environment for youth, encompassing time, region, and historical stages. The popularization of internet technology and e-commerce platforms has broadened market channels, and government

guidance and support have created a favorable rural entrepreneurial environment. Under the dual influence of policy dividends and market potential, various rural areas continue to develop, and rural entrepreneurship presents diversified models, successively driving rural economic development. In such a macro environment, education becomes the main source for entrepreneurs to acquire knowledge. Universities, as the main educational bodies, have expanded channels for entrepreneurial practice, including entrepreneurial competitions and diverse teaching formats, enabling students to learn specific and rich entrepreneurial knowledge through practice. The establishment of youth employment and entrepreneurship internship bases in many places has also become an incubator promoting youth success in employment and entrepreneurship (Campo et al., 2022).

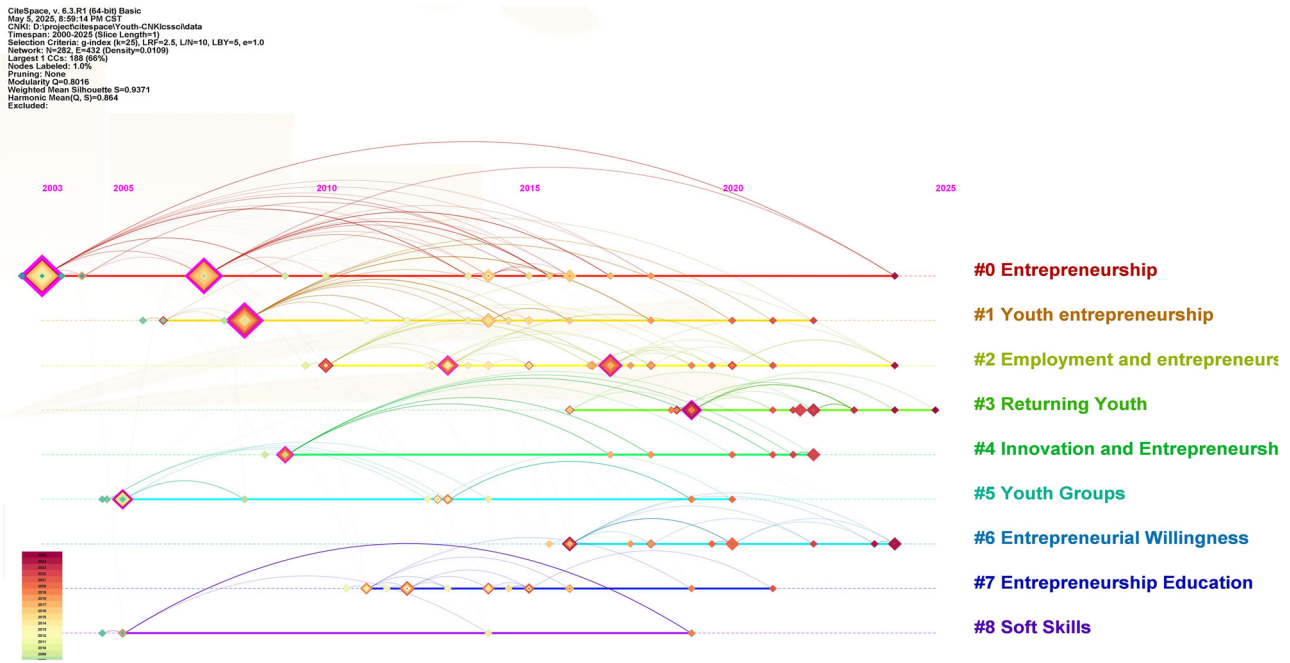


Fig. 3 Timeline of key words in domestic youth entrepreneurship.

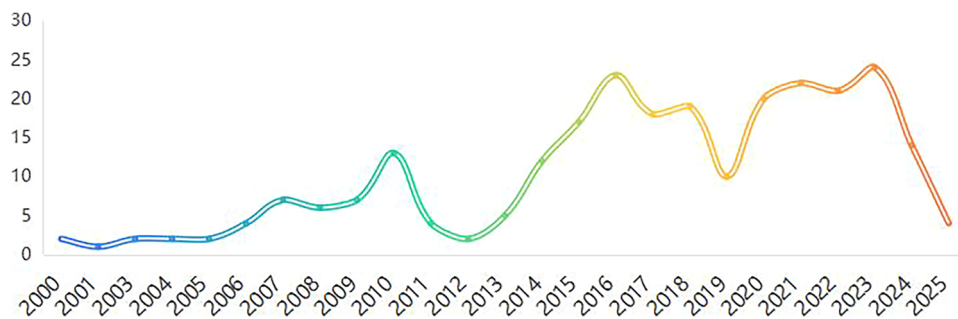


Fig. 4 Trends in the number of articles.

Temporal analysis. Figure 3 presents a timeline visualization of the keyword clusters, illustrating the historical trajectory and development of research themes over time. This analysis reveals the dynamic evolution of youth entrepreneurship research in China. With “Entrepreneurship” at its core, the theme is focused and has a long duration, laying the central axis for the field of youth entrepreneurship. From 2003 to 2005, subdivided topics such as “Youth Entrepreneurship” and “Employment and Entrepreneurship” rose rapidly, demonstrating the significant impact of policy orientation on the layout of research agendas. Subsequently, the research field gradually expanded to directions like “Returning Youth” and “Innovation and Entrepreneurship”, reflecting the regionality and diversity of entrepreneurial phenomena against the backdrop of social structural changes. The timeline analysis also shows the peak publication periods and the shifting research focus over the years.

Combining the information from Fig. 3, the timeline of key words, and Fig. 4, the trend chart of publication volume, the research process on domestic youth entrepreneurship from 1994 to 2025 can be roughly divided into the following four stages:

The Initial Stage of Youth Entrepreneurship Research (1994–2000). This stage saw a limited number of high-quality publications in the entrepreneurial domain, indicating that the focus was primarily on the awareness and preparation for youth entrepreneurship, centering on the entrepreneurial culture of

youth groups and the awakening of consciousness among young farmers.

The Development Stage of Youth Entrepreneurship Research (2000–2011) During this phase, a significant volume of high-quality research on youth entrepreneurship emerged. Relevant literature revolved around three aspects: local entrepreneurial status, entrepreneurial environment, and pathway construction. Terms such as “entrepreneurial environment” “innovation awareness” “national governance” “youth groups” and “path construction” became research hotspots. Many scholars integrated the processes of economic globalization and the macroeconomic background, comprehensively analyzing the roles of governments, schools, enterprises, and youth organizations in developing and allocating youth entrepreneurial resources, attempting to build a multi-party collaborative entrepreneurial environment. In 2002, the Ministry of Education held the “National College Entrepreneurship Education Seminar” and launched “Entrepreneurship Education Projects” which were implemented by various provinces and cities according to local conditions. With policy support, the momentum of youth entrepreneurship in China was strong, but issues such as insufficient venture capital, lack of knowledge, and social experience persisted (Research Group from Guangzhou University, 2021).

The Intensification Stage of Youth Entrepreneurship Research (2012–Present) The volume of research on youth

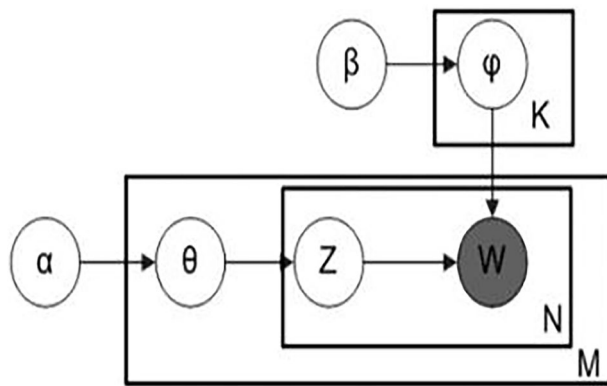


Fig. 5 LDA model workflow diagram.

entrepreneurship peaked in 2016 during this stage, followed by a yearly decrease, which aligns with the general development cycle of phenomena. The evolutionary process of the entrepreneurial ecosystem includes emergence, growth, stability, and decline phases. After experiencing the entrepreneurial boom driven by China's rapid economic development, the initial cohort of entrepreneurs has gone through a cycle of development, and related research has also entered a temporary buffer period. However, this does not signify a regression in entrepreneurship research but rather a deepening. Under the strategic deployment of "mass entrepreneurship and innovation" by the central government in 2014, policy support propelled entrepreneurship-related research to its current peak. Driven by the deep integration of industry, academia, and research, and the innovation-driven development strategy, the social entrepreneurial environment improved, meeting the development needs of youth. The connection between youth entrepreneurship research and the application of innovation practices was further strengthened, playing a significant leading and diffusing role in the subsequent development of youth entrepreneurship research. Concurrently, many new elements and hotspots have entered the field of youth entrepreneurship, with a large number of new research topics emerging, such as "Internet" "Maker Spaces" "Rural Revitalization" "Hong Kong and Macao Youth" and "Chinese Dream". The new situation poses new demands on youth entrepreneurship and has spurred scholars to research topics related to youth entrepreneurship and economic transformation.

This stage focuses on rural revitalization, grasping youth entrepreneurship as a "new engine" that promotes sustained economic growth, employment, and consumption, and stimulates social vitality, aiming to promote the modernization of agriculture and rural areas and the sustainable development of innovation and entrepreneurship. The academic research perspective and needs have also conducted in-depth discussions on the guidance mechanisms and development paths of youth entrepreneurship in line with the industrial changes under the new situation and the new era.

Empirical analysis based on LDA theme analysis

LDA model overview. LDA is a generative probabilistic model that assumes each document is a mixture of topics, and each topic is a distribution over words. The model works by iteratively assigning words to topics and topics to documents, ultimately revealing the underlying thematic structure of the corpus. Mathematically, the core idea of LDA can be simplified as a process of finding the probability distribution of topics in documents and words in topics, given a corpus of text. Consequently, a document can be seen as a blend of various topics. The core idea

of LDA is mathematically represented as follows:

$$(\theta|\alpha) = \text{Dir}(\theta|\rho\alpha)$$

LDA functions as a three-layer Bayesian probability model organized into document–topic–word hierarchies. As shown in Fig. 5, LDA selects topics probabilistically and then chooses words based on those topics, continuously repeating this process to generate all words in the document. This method allows for a clustering of terms to reveal latent semantic relationships across documents, while also reducing dimensionality and alleviating data sparsity. For this study, we applied LDA to Zhihu's "College Student Entrepreneurship" dataset, preprocessing the questions and answers before training the model to categorize each response under relevant topics. This enabled a systematic analysis of recurring themes within youth entrepreneurship discussions.

Pre analyze data from a comprehensive perspective. Analyzing 7493 Zhihu topic entries, Fig. 6 displays the trend in "College Student Entrepreneurship" discussions from 2011 to 2024. Notably, in 2012, government efforts to promote innovation and entrepreneurship sparked an initial surge in topic growth. This trend continued, peaking in 2014 as entrepreneurship became a mainstream concern. Although topic frequency slightly declined during the COVID-19 pandemic in 2020, it rebounded in 2022, as challenging employment conditions intensified public interest in entrepreneurship.

These trends underscore a rising public interest in youth entrepreneurship, driven by market constraints and heightened job competition. Additional analysis, as shown in Fig. 7, highlights word co-occurrences within the dataset. Larger nodes signify higher word frequency and greater connectivity with other terms, with 19,881 connections between 200 primary terms. The dense interconnectedness illustrates a wide-ranging and inter-related public interest in youth entrepreneurship.

Data preprocessing and model training. Before applying the LDA model, we performed several crucial preprocessing steps on the 7,493 Zhihu entries to prepare the text data for analysis. These steps were systematically applied to clean the data, reduce noise, and transform the text into a suitable format for the subsequent topic modeling process. First, text segmentation was performed, which is fundamental for analyzing Chinese text. To enhance the accuracy of segmentation, particularly for specialized terms or specific vocabulary, we loaded a custom dictionary ('dict.txt'). Simultaneously, part-of-speech (POS) tagging was applied to the segmented words. Next, the segmented words were filtered based on several criteria to remove noise and focus on terms with significant semantic meaning. These criteria included: (1) Removing non-Chinese characters. (2) Excluding words present in a comprehensive stopword list and words with a length less than 2 characters. The stopword list was constructed by manually downloading the 'Chinese.txt' file from the 'punkt' package, reading the stopwords, and removing irrelevant characters such as newline characters. (3) Retaining only words with specified POS tags (nouns 'n', other proper nouns 'nz', verbal nouns 'vn'). These three filtering criteria collectively ensured that the analysis focused on words carrying substantial semantic weight relevant to youth entrepreneurship discourse. Finally, text vectorization was performed using the CountVectorizer method. This step converted the preprocessed text data into a term-document matrix, which is the numerical format required for LDA modeling. After completing these preprocessing and vectorization steps, we trained the LDA model on the resulting numerical data.

We then trained the LDA model on the preprocessed data. A critical step in applying the LDA model is determining the

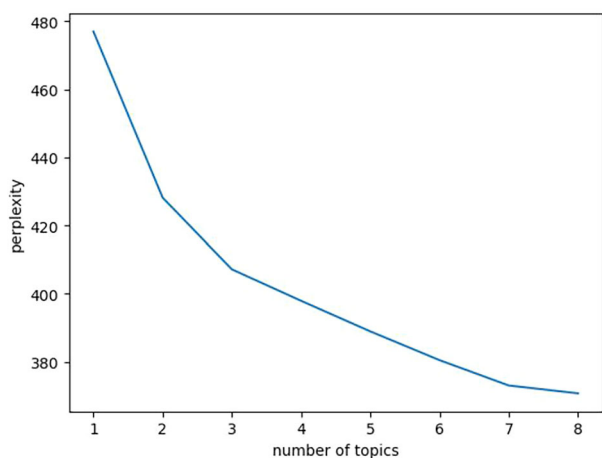


Fig. 8 Topic Perplexity.

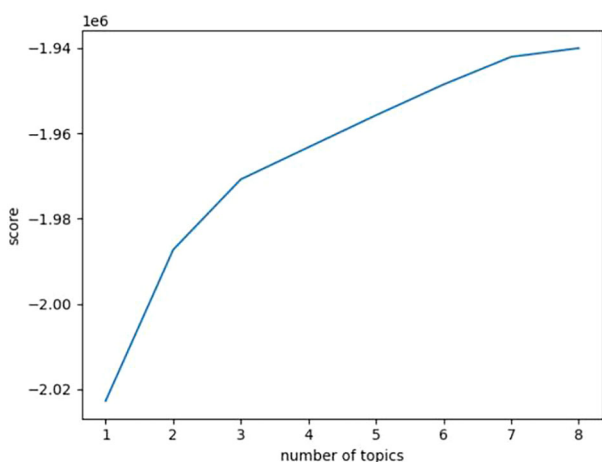


Fig. 9 Topic coherence analysis.

thereafter. Based on the optimal balance clearly indicated by these quantitative metrics, and considering the potential for meaningful interpretation of the resulting topics, we selected $K=7$ as the number of topics for our LDA analysis.

Entrepreneurship theme analysis and identification of core influencing factors.

(1) LDA Topic Keyword Distribution and Visualization Analysis

Applying the LDA model with $K=7$ yielded seven distinct topics within the “College Student Entrepreneurship” responses. Each topic’s top 15 keywords were extracted and classified, resulting in the themes shown in Table 3. To provide a more integrated and theoretically grounded understanding, Table 3 also maps these themes to relevant literature and theoretical bases. The analysis reveals that youth entrepreneurship interests revolve around team management, entrepreneurial preparation, online/offline integration, entrepreneurial pathways, business operations, market trends, and foundational experience.

To gain a deeper understanding of the discourse structure of youth entrepreneurship revealed by the LDA model, we mapped and dialogued the seven themes with existing theories and literature in the field of entrepreneurship research. This not only allows us to validate the effectiveness of our findings but also to uncover their theoretical implications within specific contexts.

Topic 1: foundational elements of entrepreneurship and risk management. This theme reflects the most core issues in entrepreneurship research: the process of new venture creation. Its keywords—“company” “product” “market” “team” and “resources”—form the basic framework of entrepreneurial activity, which highly aligns with the “opportunity, resources, team” three-element model emphasized by the classic Timmons Entrepreneurship Model (Timmons and Spinelli, 2008). Furthermore, terms like “investment” “funding” and “risk” highlight the critical role of resource acquisition and risk management in the entrepreneurial process, confirming the core tenet of the Resource-Based View (Barney, 1991), which states that the source of a firm’s competitive advantage lies in its unique ability to possess and integrate valuable, rare resources. Therefore, this theme can be regarded as a “textbook” discussion of universal principles of entrepreneurship within the discourse of youth entrepreneurship.

Topic 2: personal growth and career preparation. This theme focuses on the “pre-stage” of entrepreneurial behavior, namely, individual preparation and accumulation within the campus environment. Keywords such as “university” “school” “skills” and “specialization” point to Human Capital Theory, which posits that an individual’s knowledge, skills, and educational background are important determinants of their future productivity (Becker, 1964). Simultaneously, “friends” “peers” and “teachers” reflect the role of Social Capital, i.e., an individual’s ability to acquire information, support, and resources through social networks. These two theories collectively constitute important antecedent variables for the Theory of Entrepreneurial Intention, explaining why university students represent a significant group within youth entrepreneurship (Ajzen, 1991; Sid et al., 2025).

Topic 3: internet platforms and digital entrepreneurship. This theme distinctly reflects a new paradigm of entrepreneurship in the digital age. Keywords such as “platform” “e-commerce” “traffic” and “user” are core elements of Platform Economics theory, which involves creating value by connecting multi-sided groups (e.g., producers and consumers) through technological platforms (Parker et al., 2016). Specifically, terms like “video” “content” and “streaming” point to the rise of Digital Entrepreneurship and content creation models in recent years, characterized by asset-light operations, high iteration, and data-driven approaches. This finding indicates that, within the youth entrepreneurship discourse system, business model innovation based on internet platforms has become an extremely important and active topic.

Topics 4 and 7: entrepreneurship competitions and project practice and scientific research innovation and technology application. These two themes collectively reveal a typical path of “academic-style” entrepreneurship: the incubation process from idea to project. “Competition” “event” and “campus” embody an important practical form of Entrepreneurship Education—business plan competitions. Extensive research shows that such competitions can effectively enhance students’ entrepreneurial skills and self-efficacy. Meanwhile, “research” “scientific” and “analysis” in Theme 7 further narrow the focus to the early stages of Technology Commercialization, specifically how to transform scientific research findings into viable business ideas. This aligns with the “build-measure-learn” loop advocated by the Lean Startup methodology (Ries, 2011), emphasizing thorough analysis and validation before committing significant resources.

Topic 5: brick-and-mortar entrepreneurship and business management. Unlike high-tech or platform-based entrepreneurship, this theme depicts a more traditional and grounded

Table 3 Potential themes and top keywords for seven topics.

Topic ID	Top 15 Keywords	Topic Name	Corresponding Literature
Topic 1	Company, product, market, entrepreneur, team, project, investment, issues, enterprise, industry, risk, funding, resources, experience, capability	Foundational Elements of Entrepreneurship and Risk Management	(Timmons and Spinelli, 2008; Barney, 1991)
Topic 2	Work, university, time, school, life, graduation, skills, students, specialization, issues, teachers, friends, peers, society	Personal Growth and Career Preparation	(Becker, 1964; Ajzen, 1991; Sid et al., 2025)
Topic 3	Platform, video, product, e-commerce, traffic, user, Taobao, content, games, projects, operations, goods, internet, streaming, media	Internet Platforms and Digital Entrepreneurship	(Parker et al., 2016)
Topic 4	Project, business plan, competition, internet, team, business, event, product, tech, campus, analysis, participation, company, platform	Entrepreneurship Competitions and Project Practice	(Ries, 2011)
Topic 5	Boss, customer, cost, location, friends, price, issues, school, profit, fruits, management, students, cash, consumption, store	Brick-and-Mortar Entrepreneurship and Business Management	(Szivas, 2001)
Topic 6	Development, e-commerce, industry, society, enterprise, technology, market, economy, capability, national, issues, capital, direction, specialization, education	Macro Environment and Industry Development	(Isenberg, 2011; Stam, 2015; North, 1990)
Topic 7	Competition, project, resources, event, experience, research, business plan, analysis, specialization, interface, applications, time, section, scientific, ideas	Scientific Research Innovation and Technology Application	(Ries, 2011)

entrepreneurial landscape. Words like “store” “cost” “price” “profit” and “location” are fundamental concepts found in Small Business Management textbooks. These types of entrepreneurial activities are often categorized as Necessity Entrepreneurship or Lifestyle Entrepreneurship, where the primary goal might be to secure employment or maintain a specific way of life, rather than pursuing rapid growth (Szivas, 2001). This reminds us that youth entrepreneurship is not solely about an elite facet; a significant amount of practice exists within the microeconomic activities of daily life.

Topic 6: macro environment and industry development. This theme examines youth entrepreneurship within a broader context. Grand terms such as “society” “economy” “national” and “industry” directly point to the Entrepreneurial Ecosystems theory (Isenberg, 2011; Stam, 2015). This theory emphasizes that the flourishing of entrepreneurial activity relies on the support of systemic environmental factors such as policy, finance, culture, and market. Our study found that in the discourse of Chinese youth entrepreneurship, particular attention is paid to national development (“national”), socio-economics (“society” “economy”), and technological trends (“technology”). This may reflect the strong influence of top-down policy guidance and national development strategies in shaping the wave of youth entrepreneurship, which is also consistent with the tenets of Institutional Theory (North, 1990).

Following the identification of the seven distinct topics from the Zhihu data using the LDA model (visualized in Fig. 10), it can be seen from the clustering classification diagram that most of the topics do not overlap with each other, indicating that the model works better. Topic Feature Word Distributions represent the top 30 feature words within a topic, with the light blue colour denoting the frequency of their occurrence and the dark red colour denoting their topic weights. Topic Feature Word Distributions Topic 2 and Topic 3 have some feature words that are related, and Topic 4 and Topic 7 are farther away from the JSDs of the other four topics, indicating that it is more different from the other topics.

(2) Valuation of Influencing Factors

After applying the LDA model with $K = 7$ and identifying the top keywords for each generated topic (as detailed in Table 3), we

calculated each topic’s proportion in the corpus, which represents its relative influence in the Zhihu discussions. We sought to group these themes into broader categories (as detailed in Table 4) to provide a more integrated and theoretically grounded understanding of the influencing factors on youth entrepreneurship from the public’s perspective. Drawing upon the innovation ecosystem framework, which differentiates between internal factors related to the entrepreneur or venture and external environmental conditions, we classified these seven topics into two overarching categories: Subjective Entrepreneurial Capabilities and Objective Social Factors. Subjective Entrepreneurial Capabilities encompass themes primarily focused on the skills, knowledge, mindset, experiences, operational aspects, and processes directly related to the entrepreneur and their venture’s internal functioning. Objective Social Factors include themes related to the external environment, market dynamics, industry context, and available resources within the ecosystem.

This two-category classification is guided by the fundamental components of the innovation ecosystem framework and reflects a natural division between internally-driven entrepreneurial elements and external ecosystem influences as discussed in the Zhihu data. This framework helps to conceptualize the interplay of individual agency and environmental factors in entrepreneurship. We assigned each of the seven LDA topics to one of these two categories based on the dominant semantic meaning of its top keywords:

- Subjective Entrepreneurial Capabilities: Topic 1 (Foundational Elements of Entrepreneurship and Risk Management), Topic 2 (Personal Growth and Career Preparation), Topic 4 (Entrepreneurship Competitions and Project Practice), Topic 5 (Brick-and-Mortar Entrepreneurship and Business Management), and Topic 7 (Scientific Research Innovation and Technology Application) were grouped here. This category aligns with theories emphasizing the individual’s role and internal attributes in entrepreneurial success, such as human capital theory, self-efficacy, and entrepreneurial intentions. For instance, Topic 4, ‘Entrepreneurial Pathways’ while influenced by external factors like market competition or internet platforms (reflected in keywords), predominantly features terms indicating the entrepreneur’s actions, planning, team involvement, and analytical efforts (project, business plan,

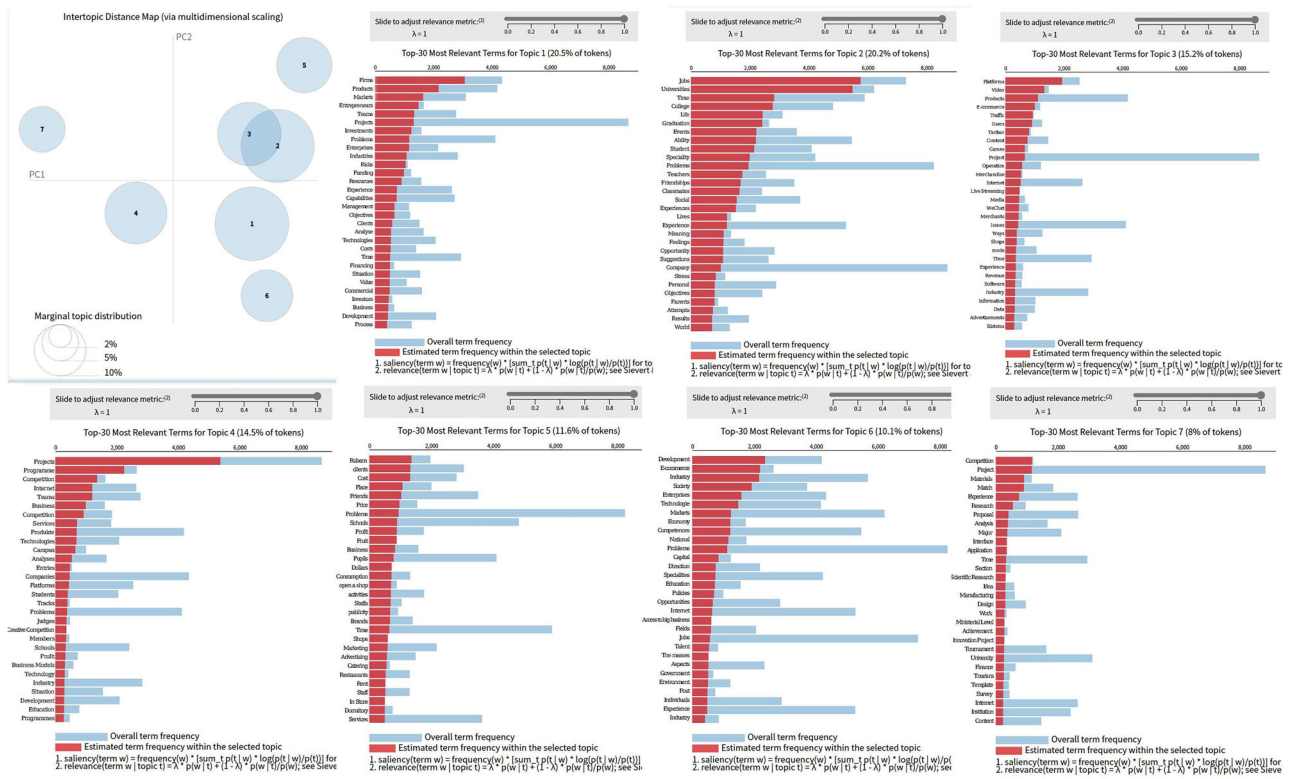


Fig. 10 LDA thematic clustering map and topic feature word distributions.

Table 4 Classification of LDA topics based on innovation ecosystem framework and influence proportion.

Topic ID	Topic Name	Proportion	Factor Type
Topic 2	Personal Growth and Career Preparation	15.48%	Subjective Capability
Topic 4	Entrepreneurship Competitions and Project Practice	3.42%	
Topic 7	Scientific Research Innovation and Technology Application	8.21%	
Topic 1	Foundational Elements of Entrepreneurship and Risk Management	26.27%	Objective Social Factors
Topic 5	Brick-and-Mortar Entrepreneurship and Business Management	28.00%	
Topic 3	Internet Platforms and Digital Entrepreneurship	12.77%	
Topic 6	Macro Environment and Industry Development	5.85%	

competition, team, analysis, participation). These elements are strongly aligned with the subjective aspects of building and navigating a venture within the ecosystem. Similarly, the other topics in this category focus on internal skills, preparation, operations, and foundational experience.

- Objective Social Factors: Topic 3 (Internet Platforms and Digital Entrepreneurship) and Topic 6 (Macro Environment and Industry Development) were grouped here, as their core keywords directly pertain to external market conditions, industry characteristics, and broader developmental trends within the entrepreneurial ecosystem. These factors are key components of the entrepreneurial ecosystem framework, highlighting the influence of external support structures, market dynamics, and institutional environment on entrepreneurial activities.

Table 4 below presents this classification of the seven topics into these two overarching categories, along with their respective proportions in the Zhihu data, highlighting the relative perceived importance of these factor types in public discourse from this specific dataset.

Topics related to entrepreneurial skills (such as Team Management and Business Operations) had the greatest

influence, indicating the importance of capabilities within the entrepreneurial cycle. In contrast, external factors like social environment and industry development were more supplementary. Notably, there was limited discussion around macro policy topics, possibly due to the nature of Zhihu as a user-driven platform where policy discussions might be less frequent compared to practical concerns. This limited discourse on macro-policy topics could stem from two potential factors. One possibility is a gap in youth awareness or understanding of available entrepreneurial policies and their relevance. Alternatively, this phenomenon may reflect the nature of the Zhihu platform itself, which is primarily user-driven for practical problem-solving rather than for policy-focused discussions. Regardless of the primary cause, this finding highlights a potential disconnect that warrants attention from policymakers and educators regarding effective channels for communicating policy information and fostering policy engagement among young entrepreneurs.

Influencing factors in entrepreneurial ecosystem

Based on a review of scholarly research and an analysis of current discussions surrounding youth entrepreneurship, three primary

factors influencing youth entrepreneurship emerge when viewed through the lens of innovation ecosystem theory.

(1) External Macro Environment: Foundational Support for Youth Entrepreneurship

The macro environment, particularly national policy support, serves as a foundational assurance for fostering youth entrepreneurial traits and capabilities. Policies such as the “Three Supports and One Assistance” (public service initiatives) and the “integration, collaboration, and sharing” framework, alongside the comprehensive implementation of rural revitalization strategies, play significant roles in enabling youth to engage in entrepreneurship and employment opportunities that are both feasible and sustainable. Such policies provide grounding for grassroots development while helping young people realize their self-worth. For example, under the “Mass Entrepreneurship and Innovation” initiative, the government has actively promoted entrepreneurial agricultural economies, resulting in notable rural economic growth driven by well-prepared young village leaders. As policies evolve, more youth are opting for careers that prioritize rural engagement, with programs for university graduates serving as a launchpad for many. By enabling young people to realize their personal and professional goals, these macro policies foster the alignment of individual value with national development objectives.

(2) Objective Social Resources: Enabling Platforms for Youth Entrepreneurship

Social resources provide critical platforms that empower youth entrepreneurship. With limited resources and a dynamic entrepreneurial landscape, competitive pressures are intensified. Industries serve as primary resource providers, offering opportunities and platforms that align with youth entrepreneurial aspirations. Innovation hubs and entrepreneurial platforms function as central gathering points, where talent and creativity converge. In the digital era, emerging opportunities in technology and data integration across traditional industries create new pathways and help establish a favorable ecosystem that nurtures youth entrepreneurship, supports industrial development, drives entrepreneurship, and creates employment opportunities.

(3) Academic Guidance: Navigational Framework for Youth Entrepreneurship

Academic institutions are essential in equipping young entrepreneurs with knowledge, fostering experiential learning, and imparting awareness of broader entrepreneurial trends and policies. From a developmental perspective, young entrepreneurs possess a blend of advanced academic knowledge and limited practical experience. According to John Dewey, early experiences are crucial in shaping one’s future. Therefore, enhancing overall competencies, understanding market competition, and cultivating practical skills are essential for future entrepreneurial success. Data from People’s Daily indicate that only 3% of college entrepreneurs succeed, with 70.2% citing a lack of experience and networks as major challenges. Given these realities, fostering experiential learning is beneficial for young entrepreneurs, as accumulated experience enhances their ability to tackle real-world challenges. Consequently, deepening entrepreneurial education should be prioritized to cultivate entrepreneurial qualities and capabilities.

Conclusions

This study, utilizing a mixed-method text mining approach, aimed to map academic research trends and public perceptions of

youth entrepreneurship in China, identifying key influencing factors through the innovation ecosystem framework. Our analysis explicitly links thematic findings from public discourse to established theoretical concepts, thereby enriching the understanding of entrepreneurial phenomena.

Our bibliometric analysis revealed that academic literature predominantly focuses on entrepreneurial education, policy support, and rural entrepreneurship, with “Returning Youth” and “Innovation and Entrepreneurship” emerging as significant themes. Conversely, the LDA analysis of Zhihu data highlighted public discourse centered heavily on subjective entrepreneurial capabilities, particularly team management and business operations.

The most striking finding is the profound emphasis on subjective entrepreneurial capabilities within public discussion, accounting for 81.38% of related topics. This underscores that young entrepreneurs prioritize internal factors—such as team management, operational skills, preparation, and foundational experience—for success. This contrasts with the more balanced focus in academic literature, which also extensively covers external factors like policy and education, revealing a potential experiential gap between formal understanding and ground-level realities.

By integrating these dual perspectives within the innovation ecosystem framework, three primary influencing factors on youth entrepreneurship were identified: (1) Subjective Entrepreneurial Capabilities, (2) the Macro-Policy Environment, and (3) Objective Social Resources. This holistic view emphasizes that while academic research often prioritizes policy and overarching frameworks, public discourse gravitates towards the practicalities of execution and individual agency.

Theoretical implications

This study offers several contributions to the theoretical understanding of youth entrepreneurship, particularly within the context of innovation ecosystems and the interplay between academic discourse and public perception. Firstly, by systematically integrating bibliometric analysis of scholarly literature with LDA topic modeling of large-scale online public discourse, our research provides a novel methodological framework for examining the multi-faceted nature of youth entrepreneurship. This framework allows for the direct mapping of emergent themes from public discourse onto established theoretical constructs, thereby strengthening the empirical grounding of theoretical discussions. This dual-perspective approach moves beyond singular analyses to offer a more holistic and nuanced view, revealing potential alignments and, more critically, disconnects between what is researched academically and what is perceived as salient by young individuals engaging with entrepreneurship.

Secondly, our findings enrich the application of the innovation ecosystem framework in the youth entrepreneurship domain. While the framework generally outlines key actors and elements, our empirical identification and categorization of influencing factors into (1) Subjective Entrepreneurial Capabilities, (2) the Macro-Policy Environment, and (3) Objective Social Resources – derived from both academic trends and public concerns – provide a more contextualized and empirically grounded specification of this framework for the Chinese youth context. The overwhelming emphasis on subjective capabilities in public discourse (Zhihu data) compared to the more balanced, often policy-centric, view in academic literature offers a critical insight into the perceived locus of entrepreneurial success and challenges, prompting a re-evaluation of how different ecosystem components are weighted from diverse perspectives. This suggests that theoretical models of entrepreneurial ecosystems might benefit from explicitly

incorporating and differentially weighting the perceived importance of various components from the viewpoint of the entrepreneurs themselves, particularly in relation to their developmental stage and cultural context.

Thirdly, by highlighting the dynamic interplay and potential experiential gap between formal academic understanding and the ground-level realities faced by young entrepreneurs, this study underscores the need for theories on entrepreneurial learning and policy effectiveness to more actively consider the channels and content of information that resonate with target youth populations. The limited discussion of macro-policy on user-driven platforms like Zhihu, despite its prominence in academic research, points to a theoretical gap in understanding policy awareness, absorption, and perceived relevance among young entrepreneurs, suggesting avenues for future theoretical development in entrepreneurial policy communication and engagement.

Recommendations

Based on the findings of our dual analysis of academic literature and public discourse, and viewed through the innovation ecosystem framework, we propose the following recommendations to foster a more supportive ecosystem for youth entrepreneurship in China:

- (1) For Universities: Enhance Practical Entrepreneurial Capability Development

Our LDA analysis of Zhihu data unequivocally demonstrated that “Subjective Entrepreneurial Capabilities” (e.g., Team Management, Business Operations) are perceived as the most critical factors by young individuals (accounting for 81.38% of topic proportion). While our bibliometric analysis shows “Entrepreneurial Education” as a consistent research theme, the emphasis from the public perspective suggests a need for more practice-oriented training. Therefore, universities should intensify their focus on cultivating students’ practical innovation and entrepreneurship abilities. This includes strengthening entrepreneurship education programs with an emphasis on experiential learning, practical skills development (e.g., financial literacy, market analysis, operational management), and team building. Providing structured mentorship and coaching opportunities, connecting students with experienced entrepreneurs and industry experts, can further bridge the gap between theoretical knowledge and real-world application.

- (2) For Government: Strengthen Policy Support and Improve Policy Awareness

The bibliometric analysis consistently highlighted “Entrepreneurial Policy” and “Innovation and Entrepreneurship” (often policy-driven) as significant themes in academic research. However, our LDA analysis of Zhihu data revealed limited discussion around macro-policy topics, suggesting a potential gap in awareness or a mismatch in communication channels, as discussed in Section 5. Therefore, the government should not only continue to refine and strengthen policies that promote youth entrepreneurship—such as simplifying regulations, providing targeted financial incentives (addressing resource limitations identified in both literature and public discourse), and fostering a supportive regulatory environment—but also significantly enhance efforts to increase awareness and accessibility of these support programs and resources among young entrepreneurs through channels they actively use.

- (3) For Industry: Foster an Enabling Environment and Collaborative Platforms

Both our literature review (Section 2.1) and the innovation ecosystem framework emphasize the importance of resources and

inter-actor collaboration. The LDA analysis also touched upon “Industry Selection” and “Market Trends” as objective social factors. To address this, industries should actively contribute to creating a healthy competitive environment and a vibrant atmosphere for innovation. This involves creating more opportunities for young entrepreneurs to access essential resources (e.g., early-stage funding, market information, supply chains), gain practical experience (e.g., internships, pilot projects), and connect with potential partners and investors. Promoting robust collaboration between universities, research institutions, and industry stakeholders is crucial to facilitate knowledge transfer, co-develop talent, and accelerate the translation of innovative ideas into viable businesses.

Limitations and future research

This study uses text quantification to analyze factors and trends influencing youth entrepreneurship, integrating bibliometrics, Python-based web scraping, and LDA modeling to construct an ecosystem evaluation framework. However, limitations exist: the study primarily synthesized academic research without a systematic analysis of state policies, and the research sample for public perception was limited to “college student entrepreneurship” discussions on Zhihu, potentially narrowing the scope of general youth perceptions.

Future research should aim to broaden its theoretical base beyond a primary focus on the innovation ecosystem to include other relevant entrepreneurship theories. Expanding sample sizes for both academic literature (e.g., including more international databases or a wider range of Chinese academic sources if quality can be assured) and public discourse (e.g., incorporating data from other social media platforms, forums, or direct surveys) is recommended. Crucially, future studies should include more diverse youth groups (e.g., rural youth not in college, vocational graduates, young employees transitioning to entrepreneurship) to enhance the applicability and generalizability of findings.

Furthermore, given China’s role as an emerging and developing nation and the global imperative for sustainable development, future research would significantly benefit from exploring how youth entrepreneurship aligns with and contributes to the United Nations Sustainable Development Goals (SDGs). This could involve investigating entrepreneurial ventures specifically addressing challenges related to poverty (SDG 1), quality education (SDG 4), decent work and economic growth (SDG 8), innovation and infrastructure (SDG 9), and climate action (SDG 13). Such a focus would not only provide a clearer framework for sustainability-oriented entrepreneurship research but also enhance the practical relevance of findings for policymakers and educators aiming to foster impactful youth ventures.

Building on the findings of this study, a key proposition for future research is that strategies explicitly integrating the development of subjective entrepreneurial capabilities—as prioritized by young entrepreneurs—with comprehensive macro-policy support and robust objective social resources within the innovation ecosystem will significantly enhance entrepreneurial success and accelerate progress towards specific Sustainable Development Goals in emerging economies like China. By advancing these directions, future studies can contribute more comprehensively to understanding and supporting youth entrepreneurship in a dynamically evolving economic landscape.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 1 February 2025; Accepted: 30 September 2025;

Published online: 20 November 2025

References

- Adner R (2006) Match your innovation strategy to your innovation ecosystem. *Harv Bus Rev* 84:98
- Aiping Y, Xiaoyun Z (2022). On the administrative promotion mechanisms of Hong Kong and Macao Youth's integration into the development of mainland cities in Greater Bay Area—a case study of the construction of the Hong Kong and Macao Youth Innovation and Entrepreneurship Bases. *Youth Explor.* 2:100–112. <https://doi.org/10.13583/j.cnki.issn1004-3780.2022.02.0010>
- Ajzen I (1991) The theory of planned behavior. *Organ Behav Hum Decis Process* 50(2):179–211
- Ana IV, Nelu F (2019) Impact of personality traits and entrepreneurship education on entrepreneurial intentions of business and engineering students. *Sustainability* 11(4):1192
- Augustino FA, Kisawike B, Ng'elenge H (2023) The role of innovation spaces towards the youth's engagement in entrepreneurship; a case of innovation spaces in the Iringa Region. *Asian J Econ Bus Account* 23(22):408–421
- Barney J (1991) Firm resources and sustained competitive advantage. *J Manag* 17(1):99–120
- Becker GS (1964) Human capital: a theoretical and empirical analysis, with special reference to education. National Bureau of Economic Research, New York, pp 15–67
- Campion a ED, Campion MA (2025) A review of text analysis in human resource management research: Methodological diversity, constructs identified, and validation best practices. *Hum Resour Manag Rev* 35(2):101078
- Campo TL, Amar SP, Olivero VE (2022) Interaction of potential and effective entrepreneurial capabilities in adolescents: modeling youth entrepreneurship structure using structural equation modeling. *J Innov Entrepreneurship* 11(1):WAJCI
- Chang YW, Lai YH (2021) What attracts young talent from Taiwan to start businesses in mainland China? A fuzzy analytic hierarchy process study. *Technol Anal Strategic Manag* 35(4):394–408. <https://doi.org/10.1080/09537325.2021.1976747>
- Cheyuan L, Zeguano C (2018) The interpretation of the Iceberg model and its lifting path of environmental cognition inhibiting the willingness of Taiwanese youths to work and start a business in China's mainland. *Contemp Youth Res* 06:94–99
- Cheyuan L, Zeguano C, Huihui W (2019) A study on the Intention of Taiwan Youth Coming to Chinese mainland for employment and entrepreneurship from the perspective of national identity: based on the investigation of eight universities in Taiwan. *J Chin Youth Soc Sci* 38(03):57–65. <https://doi.org/10.16034/j.cnki.10-1318/c.2019.03.009>
- Chuan G, Ning Y, Limei Z (2021) The evolutionary mechanism of youth returning to their hometowns for entrepreneurship under the interaction of social networks and industrial ecology. *Rural Econ* 9:128–134
- Elisa CB, Aline RS, Jorge T (2021) Fifty years of university-industry collaboration: a global bibliometrics overview. *Sci Public Policy* 48(2):177–199. <https://doi.org/10.1093/scipol/scaa077>
- Fang X, Zhu J (2020) Comparative analysis on Chinese and foreign innovation performance research—based on Bibliometrics and science mapping analysis. *Chin Manag Stud* 14(2):493–527. <https://doi.org/10.1108/CMS-03-2019-0101>
- Francisco G, Pedro S, Eduardo S (2023) What is going on in entrepreneurship research? A bibliometric and SNA analysis. *J Bus Res* 158:113624
- Gosztonyi M, Csákné Filep J, Zsigmond-Heinczinger S (2022) Comparison of applicability of startup life cycle theories based on natural language processing. *East Eur Econ* 60(6):511–539. <https://doi.org/10.1080/00128775.2022.2111312>
- Guobin D (2023). Quantitative study on the policy of youth returning to their hometowns for entrepreneurship in China - from the perspective of policy tools. *Social Sci* 1:143–152
- Haiyang Y, Entai W, Qi L, Jianan W (2024) Intelligent retrieval and comprehension of entrepreneurship education resources based on semantic summarization of knowledge graphs. *IEEE Trans Learn Technol* 17:1198–1209. <https://doi.org/10.1109/TLT.2024.3364155>
- Iserberg D (2011) The entrepreneurship ecosystem strategy as a new paradigm for economic policy: principles for cultivating entrepreneurship. Institute of International and European Affairs, Dublin, Ireland, pp. 1–13
- Jiaqing W, Deshui Z (2022) How internet use spurs entrepreneurial activities? Evidence from China. *Math Prob Eng* 2022:1–10
- Kim MH, Kim K, Peters BG (2025) Perceiving policy entrepreneurship: A Q-methodology study on the attitudes of mid-level managers in developing countries. *Public Policy Admin* 0(0) <https://doi.org/10.1177/0952076251328182>
- Longfei L (2019) Implicit logic of youth returning home to start their own businesses in the context of rural revitalization: a multi-case study from the perspective of personal meaning construction. *China Youth Study* 10:62–68. <https://doi.org/10.19633/j.cnki.11-2579/d.20190924.008>
- Maria SR, Stefano S (2017) Institutional entrepreneurship and change in consumer protection policy in the telecommunications sector: innovations in the text-based analysis approach. *Policy Soc* 36(4):611–631. <https://doi.org/10.1080/14494035.2017.1382085>
- Meidong J, Haibin W, Ji T (2022) The bibliometrics and visualization analysis on the research of maker education based on internet plus in China. 2022 *International Conference on Big Data, Information and Computer Network (BDICN)*, Sanya, China. 351–355, <https://doi.org/10.1109/BDICN55575.2022.00071>
- Mingfeng T, Yafang W (2021) Research on resilient community construction and governance based on the perspective of innovation ecosystem. *Social Sci Hunan* 1:96–103
- Moore JF (1993) Predators and prey: a new ecology of competition. *Harv Bus Rev* 71:75–86
- Moral-Muñoz JA, Herrera-Viedma E, Santisteban-Espejo A, Cobo MJ (2020) Software tools for conducting bibliometric analysis in science: An up-to-date review. *Profesional De La información* 29(1). <https://doi.org/10.3145/epi.2020.ene.03>
- North DC (1990) Institutions, institutional change and economic performance. Cambridge University Press, Cambridge
- Parker GG, Van Alstyne MW, Choudary SP (2016). Platform revolution: how networked markets are transforming the economy and how to make them work for you. W. W. Norton & Company, New York, pp 7–50
- Pillai TR, Ahamat A (2018) Social-cultural capital in youth entrepreneurship ecosystem: Southeast Asia. *J Enterprising Commun: People Places Glob Econ* 12(2):232–255. <https://doi.org/10.1108/JEC-08-2017-0063>
- Research Group from Guangzhou University (2021) We need “movers”: a discussion on the issue of linking youth entrepreneurship resources. *Youth Explor* 1:78–91
- Ries E (2011) The lean startup: how today's entrepreneurs use continuous innovation to create radically successful businesses. Crown Currency, New York, pp 75–149
- Ruzhong C, Changkiu L, Guihong C (2011) Research on the evolution law of creative industry innovation ecosystem based on organisational ecology theory. *Sci Technol Progr Policy* 28(3):64–68
- Sid AB, Wan RWN, Omar S, Amalina M (2025) Extending the theory of planned behavior in predicting entrepreneurial intention among university students: The role of perceived relational support. *Int J Manag Educ* 23(2):101168. <https://doi.org/10.1016/j.ijme.2025.101168>
- Stam E (2015) Entrepreneurial ecosystems and regional policy: a sympathetic critique. *Eur Plan Stud* 23(9):1759–1769. <https://doi.org/10.1080/09654313.2015.1061484>
- Svotwa TD, Jaiyeoba O, Roberts LM, et al. (2022) Perceived Access to finance, entrepreneurial self-efficacy, attitude toward entrepreneurship, entrepreneurial ability, and entrepreneurial intentions: a botswana youth perspective. *SAGE Open* 12(2)
- Szivas E (2001) Entrance into tourism entrepreneurship: A UK case study. *Tour Hospitality Res* 3(2):163–172. <http://www.jstor.org/stable/23746723>
- Tie W, Linling X (2022) Research on the synergistic effect of influencing factors in technological catch-up of latecomer enterprises: based on the perspective of innovation ecosystem. *J Guangxi Univ (Philos Soc Sci)* 44(04):123–130. <https://doi.org/10.13624/j.cnki.jgups.2022.04.016>
- Timmons JA, Spinelli S (2008) New venture creation: entrepreneurship for the 21st century, 8th edn. McGraw-Hill/Irwin, Boston, pp 76–105
- Timothy RH, Richard FJH, Keyvan V, Hovig T, Vern LG, Milo SW, Sarah K, P DJ (2019) Topic modeling in management research: rendering new theory from textual data. *Acad Manag Ann* 13:2. <https://doi.org/10.5465/annals.2017.0099>
- Tommaso M, Giuseppe C, Lucio C (2014) Are youth really different? New beliefs for old practices in entrepreneurship. *Int J Entrepreneurship Innov Manag* 18(2):233–259. 10.1504/IJEIM.2014.062881
- Toschi L, Ughetto E, Fronzetti CA (2023) The identity of social impact venture capitalists: exploring social linguistic positioning and linguistic distinctiveness through text mining *Small Bus Econ* 60:1249–1280. <https://doi.org/10.1007/s11187-022-00655-0>
- Weidong G, Junxia H (2021) Impact of key resource acquisition on the growth performance of new venture under the background of mass entrepreneurship and innovation mediation effects based on dynamic capabilities. *J Cap Univ Econ Bus* 23(03):86–100. <https://doi.org/10.13504/j.cnki.issn1008-2700.2021.03.008>
- Xiaohong D (2017) Internet entrepreneurship problems of returned youth and the support of 'political energy'—taking hunan as an example. *China Youth Study* 5:91–96. <https://doi.org/10.19633/j.cnki.11-2579/d.2017.05.013>
- Xu Q, Wang J, Li P (2024) Willingness valued more than ability in partner choice: Insights into behavioral and ERP data. *Psychophysiology* 61(7):e14558

- Yan C (2024) Study on the impact of digital literacy on the entrepreneurial intention of youth groups in the age of digital intelligence. *Chin Soc Sci Youth* 43(4):118–129
- Yinghui L (2020) Strong ties dominance and dual resource acquisition: an analysis of youth groups social network use in the process of entrepreneurship. *China Youth Study* 7:96–104. <https://doi.org/10.19633/j.cnki.11-2579/d.2020.0107>
- Yuan C-H, Wang D, Hong L, Zou Y, Wen J (2022) Coming back home to start up a business? A comparison between youth from rural and urban backgrounds in China. *Front Psychol* 13 <https://doi.org/10.3389/fpsyg.2022.962419>
- Yuhua C, Yongzhi Y (2019) A comparative study of regional innovative personnel policies in China: content analysis based on policy tools and system failure. *Sci Technol Manag Res* 39(10):55–65
- Zeping R, Xuesong B, Yuxin L, et al. (2022) China youth entrepreneurship development report. *China Youth Stud* 2:85–100
- Zhenxing H (2016) The Evaluation of Vc supply and demand ability in six exemplary NHTIDZS. *Sci Technol Prog Policy* 33(05):33–39
- Zhou C, Peng H, Li B (2022) How risk prevention mechanisms regulate serial entrepreneurs to achieve sustainable entrepreneurship—a policy text analysis. *Sustainability* 14(21):13904. <https://doi.org/10.3390/su142113904>

Acknowledgements

We acknowledge Jiayi Huo for her initial contributions to proofreading and formatting of the manuscript. Founded by Science Research Project of Hebei Education Department (JCZX2023005).

Author contributions

X.F.: Conceptualization; Methodology; Supervision. Y.L.: Formal Analysis; Writing – Original Draft. L.Z.: Validation; Writing – Review & Editing. S.Z.: Data Curation; Writing – Review & Editing. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Ethical approval

The study does not involve human participants or their data.

Informed consent

The study does not involve human participants or their data.

Additional information

Correspondence and requests for materials should be addressed to Linan Zhang.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025