Trend Analysis of Innovative Talents Based on the Data Visualization Using the Chinese Social Science Data

Xiaoling Yu[1,2(✉)] and Takaya Yuizono[1]

[1] Graduate School of Advanced Science and Technology, Japan Advanced Institute of Science and Technology, Nomi 9231292, Japan
s1920402@jaist.ac.jp

[2] School of Foreign Language, Dalian Polytechnic University, Dalian 116034, China

Abstract. In this study, we analyzed the literature data of innovative talents recorded in Chinese Social Sciences Citation Index using the KH Coder software and examined the academic development, research contents, and research trends of the innovative talents research in China. By analyzing the changes in the frequency of high-frequency words and the co-occurrence network diagram reflecting the visualization of relationships, we found that the research on innovative talents in China has always been focused on how to cultivate innovative talents, but the attention to "creativity" and "thinking" has decreased. At the same time, the research on innovation and entrepreneurship education, which focuses on the cultivation of innovative talents, has become a new field and trend in the research on the innovative talents in China.

Keywords: Innovation talents · CSSCI literature · KH coder · High frequency vocabulary · Co-occurrence network diagram

1 Introduction

Innovation is the activity of creating new values from resources with new ideas. In modern societies, science and technology have continued to advance and innovate, driving economic development. Innovation is an engine of growth and an important 21st-century competency [1]. With the advent of the fourth industrial revolution, the new knowledge-intensive industries will soon replace the traditional industrial economy [2]. The development of knowledge-based economy is inseparable from innovation. Innovative ability is considered to be a highly valuable ability in a knowledge society [3], and is one of the important abilities that talents should possess in the 21st century [4]. Talents with innovative abilities have become a powerful support for sustainable economic development. In this context, academic research on innovative talents has drawn the attention of a number of countries. A careful analysis and summarization of the research on innovative talents in China is an important part of the research on the innovative talents in the world. Therefore, in this study, we used the text mining software KH Coder to analyze the co-word visualization of Chinese Social Sciences Citation Index (CSSCI)
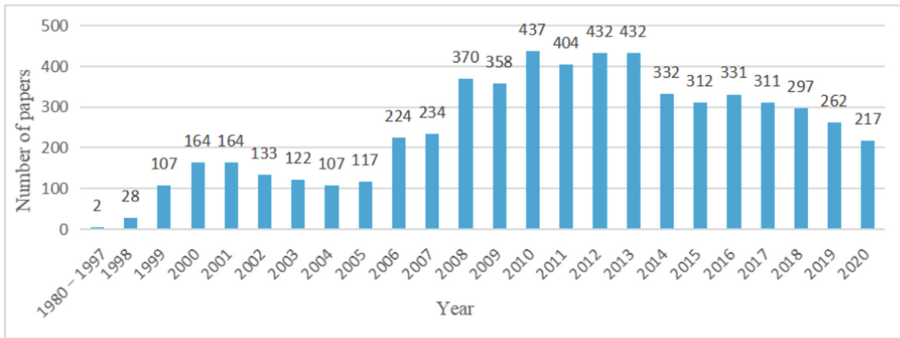
Fig. 1. Annual distribution of CSSCI papers on innovative human talents

papers related to innovative talents, summarize the research conducted on the innovative talents in China over the past 40 years, and analyze the academic development, research contents, and research trends of innovative talents research.

2   Research Methodology

First, the keywords "innovative talents," "innovation-oriented talents," and "innovative human resources" were searched on the China Knowledge Network, and a total of 5,992 CSSCI papers were retrieved by the end of 2020 (retrieved on June 1, 2021). The annual distribution of the papers is presented in Fig. 1. Among them, there were 2 papers published in 1980–1997 and 28 papers published in 1998, which increased year by year since then and reached a peak of 437 papers published in 2010. From 2010 to 2014, the number of papers published annually remained above 400, and decreased to 217 in 2020. Second, from the CSSCI papers published each year, the five most cited papers were selected and imported into Text format files, and KH Coder was used for word selection and pre-processing to ensure the accuracy of the extracted data. Then, KH Coder is used to calculate the frequency of high-frequency keywords and draw a co-occurrence network diagram reflecting the visualization of keyword relationships. Finally, the academic development and research trends of the innovative talents research in China are discussed with the change in the high-frequency words and keyword co-occurrence network diagram.

3   Research Results

3.1   The Extracted Words

From the 5,992 CSSCI papers searched, the five most cited papers per year were selected to obtain a total of 117 research papers related to innovative talents (selected as of June 1, 2021). These papers were loaded into a Text format file for preliminary text processing by using KH Coder. Following the processing of the preliminary text, we found out the total number of words in the text of Chinese innovative talents research literature to be 515,244 and the number of dissimilar words in the text to be 19,666. This text contains a certain number of words that can be used for text data mining via the use of KH Coder.

Table 1. Extracted words

| Ranking	Before 2000			2000–2010			2011–2020			Total		
	Extracted words	Times	Appearance percentage	Extracted words	Times	Appearance percentage	Extracted words	Times	Appearance percentage	Extracted words	Times	Appearance percentage
1	Education	846	1.33%	Education	2246	1.16%	Education	3961	1.55%	Education	7053	1.37%
2	Student	650	1.02%	Innovation	2037	1.05%	Talent	2786	1.09%	Innovation	5402	1.05%
3	Innovation	622	0.98%	Talent	2028	1.05%	Innovation	2743	1.07%	Talent	5202	1.01%
4	Knowledge	555	0.87%	Student	1712	0.88%	Student	2111	0.83%	Student	4473	0.87%
5	Ability	416	0.65%	Ability	1465	0.76%	Cultivation	1607	0.63%	Ability	3369	0.66%
6	Talent	388	0.61%	Teaching	1329	0.69%	Cultivate	1523	0.60%	Teaching	3131	0.61%
7	Teaching	361	0.57%	Cultivation	1202	0.62%	Ability	1488	0.58%	Cultivation	2986	0.58%
8	Thinking	353	0.56%	Cultivate	1060	0.55%	Teaching	1441	0.56%	Cultivate	2831	0.55%
9	Problem	310	0.49%	University	1044	0.54%	Entrepreneurship	1362	0.53%	Knowledge	2382	0.46%
10	Creativity	293	0.46%	Knowledge	920	0.47%	Innovate	1326	0.52%	University	2223	0.43%

3.2 High Frequency Vocabulary

The results of the survey on high-frequency words using KH Coder's high-frequency word counting function are shown in Table 1. The top ten most frequently appeared words were "education," "innovation," "talent," "student," "ability," "teaching," "cultivation," "cultivate," "knowledge," and "university." Of these words, "education" was mentioned 7,053 times, "innovation" 5,402 times, "talent" 5,202 times, "student" 4,473 times, and "ability" 3,369 times.

Subsequently, the high-frequency words statistics function of KH Coder was used to analyze the papers published before 2000, 2001–2010, and 2011–2020, respectively, and it was found that the most frequently appeared keywords did not change much in different periods. Regardless of the periods, the six keywords, namely "education," "students," "innovation," "ability," "talent," and "teaching," were among the top ten frequency keywords. In all three periods, "education" consistently topped the list of high-frequency words, with "innovation" occupying third place. Another noteworthy change is that "entrepreneurship" did not appear before 2000, was rarely mentioned between 2001–2010, and rose directly to 10th place in the list of the most frequently appeared keywords of 2011–2020. In contrast, "thinking" and "creativity," which occupied 8th and 10th places in the list of high-frequency words before 2000, showed decreasing fluctuations.

3.3 Co-occurrence Network Diagram for Keywords

The co-occurrence network has been traditionally used in content analysis to statistically express the data [5]. The keyword co-occurrence network diagram is drawn by KH Coder to form a visualization of the relationship between keywords, and the relevance of the keywords is represented by circles and lines. The larger the area of the circle, the more frequently the keyword appears in the text, and the thicker the line connecting the two circles, the higher the relevance of the two keywords. In this paper, the keyword co-occurrence network diagram analyzed using KH Coder is presented in Fig. 2. The co-occurrence network diagram of this study consists of seven main parts, each of which is composed of circles of different colors.
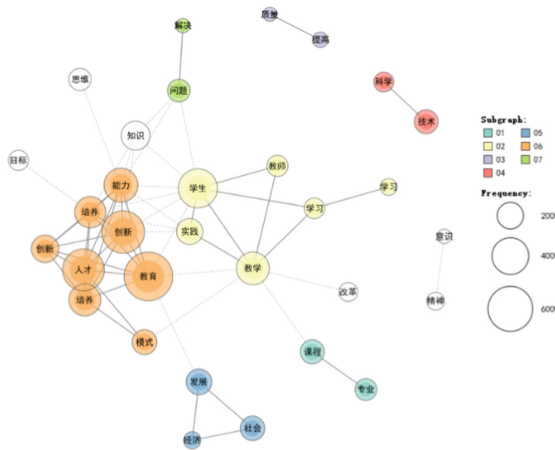
Fig. 2. Co-occurrence network of words (1982–2020)

As can be seen in Fig. 2, the group with the most highly frequently appearing words is the first group. It consists of eight keywords: "education," "innovation," "talent," "cultivation," "cultivate," "ability," "innovate," and "model." Of these keywords, the keywords "education," "cultivation," "ability," "innovation," and "talent" are the closest to "innovation," and "talent", and most closely related. The group with the second most highly frequently appearing words is the second group, consisting of six keywords. These include "student," "teaching," "practice," "teacher," "study," and "studying." The third group consists the following keywords: "development," "society," and "economy."

At the same time, we draw keyword co-occurrence network diagrams for different time periods of the literature to examine the visualized relationships between the keywords. In the co-occurrence network diagram for the pre 2000 literature (Fig. 3), "education," "ability," "cultivation," "innovate," "cultivation," "spirit," and "quality" are the keywords that are close to "innovation" and "talent," and these are classified as a group. In the co-occurrence network diagram for the 2001–2010 literature (Fig. 4), "education," "cultivation," "innovate," "cultivate," and "ability" are the key words that are close to "innovation" and "talent," and these are classified as a group. Further, in the co-occurrence network diagram for the 2011 to 2020 literature, "cultivation," "cultivate," "innovate," "model," "university," and "goal" are the keywords that are close to "innovation" and "talent" (Fig. 5), and these are classified as a group. It can be seen that "innovative talents" is closely related to "cultivation," "training," and "innovate," and is always classified as belonging to the same group. This confirms that the research on innovative talents in China is mainly focused on the direction of how to cultivate innovative talents. Meanwhile, "ability" was always in the same group, with "innovation" and "talent" before 2000 and during 2001–2010, but was classified as belonging to a group with "students," "knowledge," and "practice" during 2011–2020.
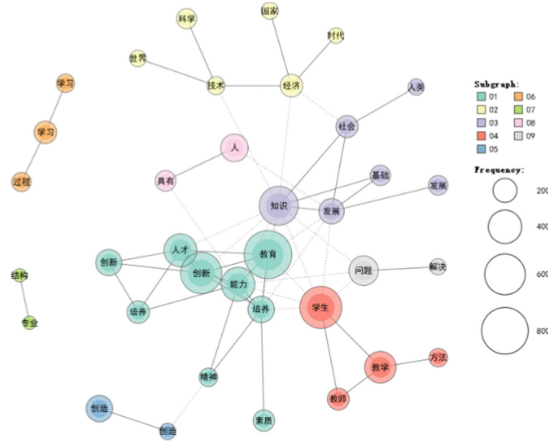
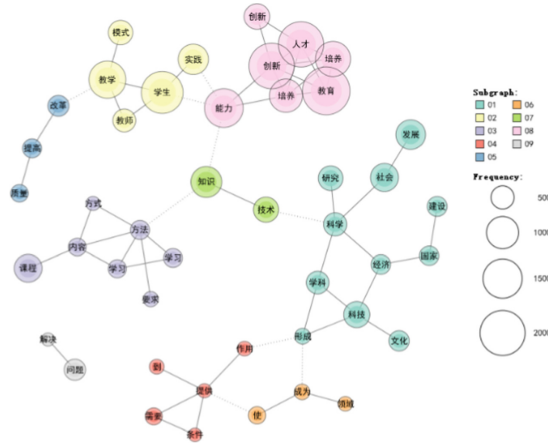**Fig. 3.** Co-occurrence network of words (before 2000)

Fig. 4. Co-occurrence network of words (2001–2010)

3.4 Co-occurrence Network Diagram for Centrality Keywords

The high-frequency word centrality analysis of the paper was conducted using KH Coder, and the top three centrality keywords were "teaching," "education," and "ability." These three words were in the list of the top ten frequency keywords. It can be seen that these three key words are the three most important words in the literature. Therefore, we analyzed the three centrality keywords in the keyword co-occurrence network diagram.

Of these, "teaching" is closely related to "students," "practice," "teacher," and "learning," connected via thick solid lines, while it is less related to "education" and "mode," connected via dashed lines. "Education" is connected to "innovation," "talent," "ability," "cultivation," "innovate," and "cultivate" via thick solid lines, and to "student,"
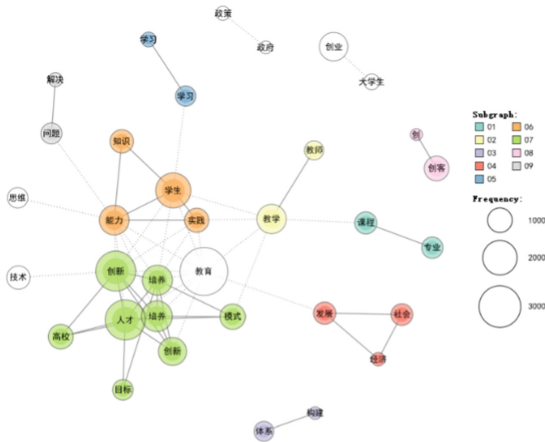
Fig. 5. Co-occurrence network of words (2011–2020)

"teaching," and "development" via dashed lines. "Ability" is firmly connected to "education," "cultivation," "talent," "cultivation," and "cultivate" via thick solid lines, and to "student," "practice," "problem," "knowledge," and "thinking" via dotted lines.

4   Conclusion and Discussion

4.1   Decline in Attention Toward "Creativity" and "Thinking"

A noticeable change in the frequency of high-frequency words is the decrease in the frequency of "creativity" and "thinking." The keyword "thinking" ranked 8th (353 times, 0.56%) until 2000, dropped to 38th (372 times, 0.19%) from 2001 to 2010, and then rose up to 26th (678 times, 0.27%) from 2011 to 2020. "Creativity" ranked 10th (293 times, 0.46%) until 2000, falling in rank to 121st (136 times, 0.07%) from 2001 to 2010, and then rose up to 69th (298 time, 0.12%) from 2011 to 2020. Basing on the distribution of CSSCI papers on Chinese innovative talents, the research on Chinese innovative talents started in the 1980s. The early research on innovative talents mainly focused on the importance of innovative talents, innovation education, and innovative talents cultivation [6]. Creativity enhancement and innovative thinking are important parts of innovative talents training, thus, "creativity" and "thinking" appeared more frequently before 2000. After 2000, the research on innovative talents in China gradually developed into a multi-perspective of practical teaching, innovative talents cultivation, and international comparison of innovative talents [6]. "Creativity" and "thinking" were given relatively less attention, and eventually, their frequency decreased.

4.2   The Main Research Direction of "Cultivating Innovative Talents Through University Education"

In terms of the centrality of keywords, "teaching" and "education" were the two most central keywords. At the same time, the frequency of the occurrence of the sub-words

made them stand in the sixth and first places in the list of high frequency words. It can be seen that "teaching" and "education" are two very important key words in the literature. Based on the co-occurrence network diagram of keywords, the keyword "education" is closely related to "innovation," "talent," and "cultivation." "Teaching" is closely related to "practice." Moreover, based on the keyword co-occurrence network diagram in different periods, "innovative talents" and "cultivation" are always classified as belonging to the same group. It can be seen that the question of how to cultivate innovative talents through education always receives attention. Education, especially university education, is indispensable to cultivate innovative talents. Nowadays, higher education is trying to cultivate students with strong innovative ability [7]. In China, research and practice related to the fostering of innovative talents via university education has also spawned a large number of research results, as evidenced by the frequency of most frequently appeared keywords and the positional relationship of the co-occurrence network diagram in the Chinese CSSCI literature. In comparison to the other research directions, the question of how to cultivate innovative talents through university education still constitutes the main research direction of innovative talents research in China.

4.3 The New Research Direction of "Innovation and Entrepreneurship Education"

In terms of the frequency of keywords, there is a significant change in "entrepreneurship." While "entrepreneurship" did not appear in the literature before 2000 and has been mentioned only 66 times during 2001–2010, it rose to tenth place during 2011–2020, with 1,362 mentions. In the keyword network co-occurrence diagram for the period 2011–2020, it can be seen that "entrepreneurship" and "students" are connected via dashed lines and establish a connection gradually. The change of the keyword "entrepreneurship" shows that innovation and entrepreneurship education have become a new trend in the study of the innovative talents in China. This change is closely related to "mass entrepreneurship and innovation" proposed by the Chinese Premier Li Keqiang at the 2014 Summer Davos Conference. At the conference, Li Keqiang proposed to turn the "mass entrepreneurship and innovation" strategy into a new engine for China's economic and social development under the new normal. In 2015, the era of "mass entrepreneurship and innovation" began officially, and scholars started to focus on the research field of innovation and entrepreneurship education and carried out in-depth research [8]. A number of higher education and vocational schools in China also encourage students to actively participate in innovation activities and entrepreneurial practices, and innovation and entrepreneurship education and innovative talents cultivation have come to flourish. Therefore, research on innovation and entrepreneurship education, focusing on the cultivation of innovative talents, has become a new field and trend in the study of the innovative talents in China.

In this study, above descriptions are leaded by using the top five most-cited papers, but more papers should be considered toward more deepen understanding in the future research.

Disclosure Statement.   The author declares no conflict of interest.

References

1. Morad, S., Ragonis, N., Barak, M., 2021, An integrative conceptual model of innovation and innovative thinking based on a synthesis of a literature review. Thinking Skills and Creativity, 40, 100824.
2. Jingyu, Y., Yi, S., 2018, Research on developing the cultivation of innovative talents in China under the fourth industrial revolution. 2018 International Conference on Management Science & Engineering, 130–137.
3. Scott, C. L., 2015, The futures of Learning 3: What kind of pedagogies for the 21st century? Education Research and Foresight, 1–21.
4. Shawn, J., Micah, L., 2016, Additive innovation in design thinking and making. International Journal of Engineering Education, 32(3), 1438–1444.
5. Higuchi, K., 2016, A two-step approach to quantitative content analysis: KH Coder tutorial using Anne of Green Gables (Part I). Ritsumeikan Social Science Review, 52(3), 77–91.
6. Yayuan, L., Ruixue, W., Na, L., 2018, Innovative talent research: Combing and foresight of thirty years of academic History. Higher Education Management, 12(3), 116–124. (In Chinese)
7. Barak, M., Usher, M., 2019, The innovation profile of nanotechnology team projects of face-to-face and online learners. Computers & Education, 137, 1–11.
8. Bai, Yu., 2016, A review of research on the cultivation model of top innovative talents in first-class universities. Education and Teaching Forum, (9), 1–2. (In Chinese)