Postprint: Revealing the Impact of China-US Basic Research on Technological Innovation Through Innovative Scientific Structure Mapping

Authors: Chen Ting, Leng Fuhai, Li Guopeng, Wang Xiaomei

Date: 2024-05-18T00:00:00+00:00

Abstract

This article employs scientific structure mapping analysis methods to, for the first time, construct an innovative scientific structure map, using patent-cited Essential Science Indicators (ESI) highly cited papers as the data foundation, revealing the impact of world frontier basic research on technological innovation development; identifying hotspot directions that drive technological innovation development, and comparing the performance of China and the United States in these hotspot directions to provide effective data support for scientific decision-making. Analysis reveals: patent-cited papers are primarily concentrated in basic research fields such as new biotechnology and biomedical technology, emerging information technology, energy, and materials that may trigger disruptive technological innovation; China' s patent-cited highly cited papers show significant growth momentum, with a cumulative increase exceeding 60% over five years, yet the quantity remains less than 50% of that of the United States; China' s paper output in many hotspot research directions closely linked to technological innovation exceeds that of the United States, but the number of patent-cited papers is typically lower than that of the United States; the basic research content influencing technological innovation in the United States is more focused on R&D activities that lead industrial transformation, with substantial participation from enterprises.

Full Text

Preamble

Citation Format: Chen T, Leng F H, Li G P, et al. Innovation Science Structure Map: Uncovering influence on fundamental research to technological inno-

vation in China and the United States. Bulletin of Chinese Academy of Sciences, 2024, 39(3): 588-600, doi: 10.16418/j.issn.1000-3045.20230417001.

Innovation Science Structure Map: Uncovering influence on fundamental research to technological innovation in China and the United States

CHEN Ting, LENG Fuhai, LI Guopeng, WANG Xiaomei*

Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190

Abstract

This study employs scientific structure mapping analysis to create, for the first time, an Innovation Science Structure Map based on Essential Science Indicators (ESI) highly cited papers referenced in patents. This approach reveals how cutting-edge fundamental research worldwide influences technological innovation development, identifies hotspot directions that drive technological advancement, and compares the performance of China and the United States in these areas, providing robust data support for scientific decision-making. The analysis finds that papers frequently cited by patents concentrate primarily in emerging biotechnology and biomedical technology, new information technology, and energy and materials—fundamental research fields likely to trigger disruptive technological innovation. China shows remarkable growth in patent-cited highly cited papers, with a five-year cumulative increase exceeding 60%, though the total remains less than half of the United States. In many hotspot research directions closely linked to technological innovation, China produces more papers than the United States, yet typically fewer patent-cited papers. U.S. fundamental research influencing technological innovation tends to focus more on R&D activities that lead industrial transformation, with substantial enterprise participation.

Keywords: innovation map, patent-cited papers, basic research, highly cited papers, technological innovation

1. Data and Methods

This study selected 56,278 papers from the top 1% of highly cited papers in Clarivate Analytics' Essential Science Indicators (ESI) Research Fronts from March 2022 (covering 2016–2021). Using patent-to-paper citation data from a patent analysis platform, we identified 9,767 highly cited papers referenced in patents for analysis. Drawing on scientific structure mapping principles and techniques, we first employed a deep learning model to analyze the complex co-citation networks of these 9,767 highly cited papers, converting network nodes (papers) and relationships (co-citation frequencies between paper pairs) into low-dimensional, dense feature vectors. We then used manifold learning visualization algorithms to project these node vectors onto a two-dimensional plane,

creating an Innovation Science Structure Map that reveals the scientific origins influencing technological innovation. This method effectively preserves the local structure of high-dimensional data, allowing papers with similar citation relationships to cluster closely in the two-dimensional visualization and facilitating identification of research hotspots. Using kernel density algorithms, we automatically identified areas with high paper density, which were further classified as hotspot topics. Notably, the Innovation Science Structure Map constructed here differs from traditional science maps: its nodes represent highly cited papers referenced in patents, focusing specifically on reflecting how fundamental research influences technological innovation.

2. Trends in Basic Research Influencing Technological Innovation

2.1 Constructing the Innovation Science Structure Map to Reveal Scientific Origins of Technological Innovation

To deeply reveal the structure and characteristics of fundamental research influencing technological innovation, we analyzed the relationship distribution of 9,767 patent-cited highly cited papers using deep learning and visualization technologies, constructing the Innovation Science Structure Map [Figure 1: see original paper]. Each node represents one patent-cited highly cited paper; papers with more similar research content appear closer together in the map, naturally clustering into research themes based on similarity. Drawing on the concept of regional density maps from geographic information analysis, we identified high-density areas representing hotspots where numerous similarly-oriented papers influencing technological innovation cluster within a timeframe. Through algorithmic identification and manual interpretation, we labeled 63 research themes and grouped them into three major fields. First, life sciences —the largest area, comprising nearly 60% of the map—includes 33 themes such as bioinformatics and genomics analysis, gene editing and therapy, COVID-19 diagnosis and treatment, Zika and dengue viruses, HIV antibodies and vaccines, transcriptome functional studies, transgenic crops, PD-1/PD-L1 immune checkpoint therapy, immune checkpoint inhibitor side effects, cancer liquid biopsy, leukemia treatment, arthritis treatment, liver cancer targeted therapy, drug pharmacology, neuroscience, and cell biology. Second, physical sciences encompasses 20 themes including quantum computing, two-dimensional van der Waals materials, lithium batteries, perovskite materials, catalysts, and polymer solar cells, spanning condensed matter physics, optics, materials, energy, and chemistry. Third, information science includes 10 themes such as deep learning models and applications, artificial intelligence in healthcare, 5G networks, blockchain, smart city wireless communications, and drone communications.

2.2 National and Institutional Statistics of Patent-Cited Highly Cited Papers

(1) China ranks second globally in patent-cited highly cited papers, with significant growth but a clear gap with the United States. Among the top 10 countries by patent-cited highly cited papers [Figure 2: see original paper], the United States holds a dominant position, accounting for over 50% of all data with 5,363 papers—far exceeding other nations. China ranks second with 2,571 papers, followed by the United Kingdom, Germany, and France. Examining trends across three time windows [Figure 3: see original paper], China' s patent-cited papers grew from 1,563 in the first period (2012–2017) to 2,571 in the third period (2016–2021), an increase exceeding 60%, while U.S. numbers remained relatively stable. While China' s total highly cited papers reach 81.8% of the U.S. total, its patent-cited highly cited papers are only 47.9% of the U.S. figure, indicating a substantial gap.

(2) The Chinese Academy of Sciences ranks second globally and first in China in patent-cited highly cited papers. Among the world' s top 10 institutions [Figure 4: see original paper], the United States has six institutions, China has one, France has two, and the United Kingdom has one. Harvard University leads with 1,058 papers, far surpassing others. The Chinese Academy of Sciences ranks second globally and first in China with 598 patent-cited highly cited papers. Among Chinese institutions in the global top 200 , 10 are listed, including the Chinese Academy of Sciences, seven members of the C9 League, the University of Hong Kong, and Wuhan University. Compared to their rankings in all highly cited papers, seven of these universities show slightly lower positions in patent-cited papers.

3. Analysis of Key Hotspot Topics Influencing Technological Innovation

Using kernel density algorithms, we identified hotspot topics where highly cited papers show significant clustering effects, forming "heat highlands" that likely represent key research themes influencing industrial and technological innovation. Based on paper count (exceeding 50 papers) and heat level (density level 3 or above), combined with manual interpretation, we selected 18 key hotspot topics from the 63 research themes [Figure 5: see original paper]. Life sciences contributed the most hotspot themes (8), followed by physical sciences (7) and information science (3) .

3.1 Life Sciences

Life sciences yielded eight hotspot topics with the highest number of patent-cited highly cited papers. These cluster into three subdomains: gene editing technology, virus treatment, and malignant tumor treatment—all critical to human health with enormous application potential. The United States demonstrates clear advantages in life sciences, with more patent-cited papers than

the world average in all themes except COVID-19 diagnosis and treatment and Zika/dengue viruses.

(1) Gene editing technology. Showing vast potential in health and agriculture, this area includes three hotspot themes: bioinformatics and genomics analysis (70 papers), gene editing and therapy (150 papers), and single-cell transcriptome sequencing (116 papers). Gene editing and therapy achieves the highest average citation frequency per paper at 25.7 patent citations. China outperforms the world average in both paper and patent citation frequencies for COVID-19 diagnosis and treatment, with 129 patent-cited papers, though only 4 involve enterprise participation compared to 30 in the United States. In other life science themes, China has only single-digit corresponding author papers.

(2) Virus treatment. The urgency of virus treatment became especially evident during the recent COVID-19 pandemic. We selected two hotspot themes: COVID-19 diagnosis and treatment, and Zika and dengue viruses. The COVID-19 theme contains the most papers of any theme. China' s performance is notable in COVID-19 diagnosis and treatment, with paper and patent citation frequencies exceeding world averages and surpassing the United States, though U.S. papers show more enterprise involvement.

(3) Malignant tumor treatment. As a leading cause of global morbidity and mortality, cancer has seen major breakthroughs in immunotherapy. We selected three hotspot themes: tumor targeted therapy (67 papers), malignant hematological disease chemotherapy and immunotherapy (69 papers), and immune checkpoint therapy and side effects (143 papers). These themes show relatively high average patent citation frequencies of 13.2, 18.3, and 10.0 respectively. The United States maintains significant advantages in average citation frequencies for gene editing and therapy and immune checkpoint therapy themes.

3.2 Physical Sciences

Physical sciences yielded seven hotspot themes spanning condensed matter physics and optics, and energy and materials.

(1) Condensed matter physics and optics. This includes three hotspot themes: quantum computing (67 papers) focusing on quantum simulation, quantum memory, and quantum algorithms; metasurfaces and metamaterials (121 papers) covering metasurface lenses and photonic chip-based frequency combs; and two-dimensional van der Waals materials (56 papers) including van der Waals heterostructures and "magic angle" graphene. China lags significantly behind the United States in quantum computing (1 vs. 44 corresponding author papers) and has less than 50% of U.S. output in metasurfaces and metamaterials. However, China matches the United States in two-dimensional van der Waals materials (19 papers each), with China' s average patent citation frequency (10.6) far exceeding the U.S. (2.3), indicating strong Chinese capabilities in this emerging material field with vast potential for electronics, energy storage, and environmental remediation.

(2) Energy and materials. This includes four hotspot themes: lithium batteries (81 papers) covering solid-state batteries and electrolyte preparation; perovskite solar cells (105 papers) on material structures and carrier lifetimes; perovskite light-emitting diodes (56 papers) on nanocrystals and device performance; and flexible materials and wearable devices (40 papers) with the second-highest average patent citation frequency (6.8) after deep learning. China leads in perovskite light-emitting diode papers (26 vs. 8 for the U.S.) but trails significantly in flexible materials and lithium batteries. While China' s average citation frequency exceeds world averages in lithium batteries, the United States shows higher frequencies in most other themes. Notably, U.S. papers in quantum computing and lithium batteries show substantial enterprise participation from tech giants like Google, Qualcomm, Microsoft, and Intel, as well as German and Korean R&D firms.

3.3 Information Science

Information science yielded three hotspot themes: deep learning models and applications, artificial intelligence in healthcare, and blockchain.

China' s average citation frequencies fall below world averages in all three themes, while the United States exceeds them. China produces more papers than the United States in deep learning models and applications (88 vs. 29) and blockchain (22 vs. 9), but U.S. influence in both paper and patent citations significantly surpasses China' s. U.S. papers show greater enterprise involvement, including teams from Google and Meta, while Chinese papers are dominated by universities and research institutes.

(1) Deep learning models and applications. This theme includes 185 highly cited papers on applications such as autonomous driving, remote sensing, image processing, and human activity detection, achieving the highest patent citation frequency (10.7) outside life sciences.

(2) Artificial intelligence in healthcare. With 126 papers, this theme focuses on deep learning applications in CT, MRI, and X-ray imaging, where the United States holds substantial advantages.

(3) Blockchain. This theme includes 56 papers across applications like IoT, smart grids, vehicle networks, and healthcare data. China shows strong performance in deep learning models and blockchain research but lags in citation impact.

4. Conclusions and Implications

4.1 Conclusions

Comparing Chinese and U.S. fundamental research influencing technological innovation reveals several key findings:

First, fundamental research affecting technological innovation concentrates in

emerging biotechnology and biomedical technology, new information technology, and energy and materials—fields likely to trigger disruptive innovation. Biotechnology and biomedicine show the highest coverage, including gene editing, single-cell sequencing, virus treatments (COVID-19, Zika, HIV), and disease treatments for cancer, diabetes, and Alzheimer' s. Energy and materials include lithium batteries, perovskite cells, 2D materials, and flexible/wearable devices. New information technology encompasses AI healthcare, quantum computing, 5G networks, and blockchain.

Second, China' s patent-cited highly cited papers show remarkable growth but remain substantially behind the United States. While China' s five-year growth exceeds 60%, its total is less than half of the U.S. figure. China leads in some hotspot themes including deep learning models, 2D van der Waals materials, perovskite cells and LEDs, and COVID-19 diagnosis and treatment. The United States shows clear advantages in AI healthcare, blockchain, quantum computing, lithium batteries, flexible materials, bioinformatics, gene editing, immune checkpoint therapy, and malignant hematological disease treatment.

Third, China's impact remains lower than the United States despite higher paper counts in some areas. In many hotspot themes closely linked to technological innovation, China produces more highly cited papers and receives more paper citations than the United States, but typically fewer patent citations. China excels in COVID-19 research with citation frequencies above world averages, and its 2D van der Waals materials and lithium battery research also exceed global benchmarks. However, the United States generally surpasses world averages across most themes.

Fourth, U.S. enterprises demonstrate high participation in fundamental research influencing technological innovation. Notably, U.S. hotspot themes with high patent citation frequencies involve substantial enterprise participation in deep learning, lithium batteries, and quantum computing, with early-stage involvement from tech giants. This contrasts with China, where patent-cited highly cited papers are almost exclusively from universities and research institutes.

4.2 Recommendations

Based on these findings, we propose three recommendations for China' s fundamental research development:

First, strengthen existing advantageous fields. Government and relevant institutions should sustain investment, encourage rapid industrialization and commercialization of fundamental research already showing technological impact, and drive economic growth through technological innovation.

Second, strategically deploy fields likely to cause disruptive technological innovation. Increase R&D investment, cultivate and attract top talent in these areas, and facilitate the transfer of fundamental research to technological outcomes.

Third, drawing from U.S. experience with early enterprise involvement in funda-

mental innovation, China should further incentivize capable enterprises to participate in potentially disruptive fundamental research through tax incentives, R&D subsidies, and policy support. Provide more venture capital support to reduce investment risks and attract private capital into technological innovation, particularly in early-stage R&D, to better promote rapid application and deployment of innovative achievements.

References

1. Pan J F, Du P. Strengthening the knowledge foundation of building a strong science and technology country—Starting from the concept and characteristics of basic research. China Science Daily, 2021-11-22(01). (in Chinese)
2. Jefferson R. Comment: Turning science into social outcomes. Nature, 2017, 548: S8.
3. Wang X M. Mapping Science Structure 2021. Beijing: Institutes of Science and Development, Chinese Academy of Sciences, 2021. (in Chinese)
4. Wang X M, Han T, Li G P, et al. Mapping Science Structure 2021. Beijing: Science Press, 2017.
5. Chen T, Li G P, Wang X M. Optimizing science knowledge mapping methods to map the global science structure graph. Library and Information Service. 2022, 66(21): 107-119. (in Chinese)
6. Chen T, Li G P, Deng Q P, et al. Using network embedding to obtain a richer and more stable network layout for a large scale bibliometric network. Journal of Data and Information Science, 2021, 6(1): 154-177.

Note: Figure translations are in progress. See original paper for figures.