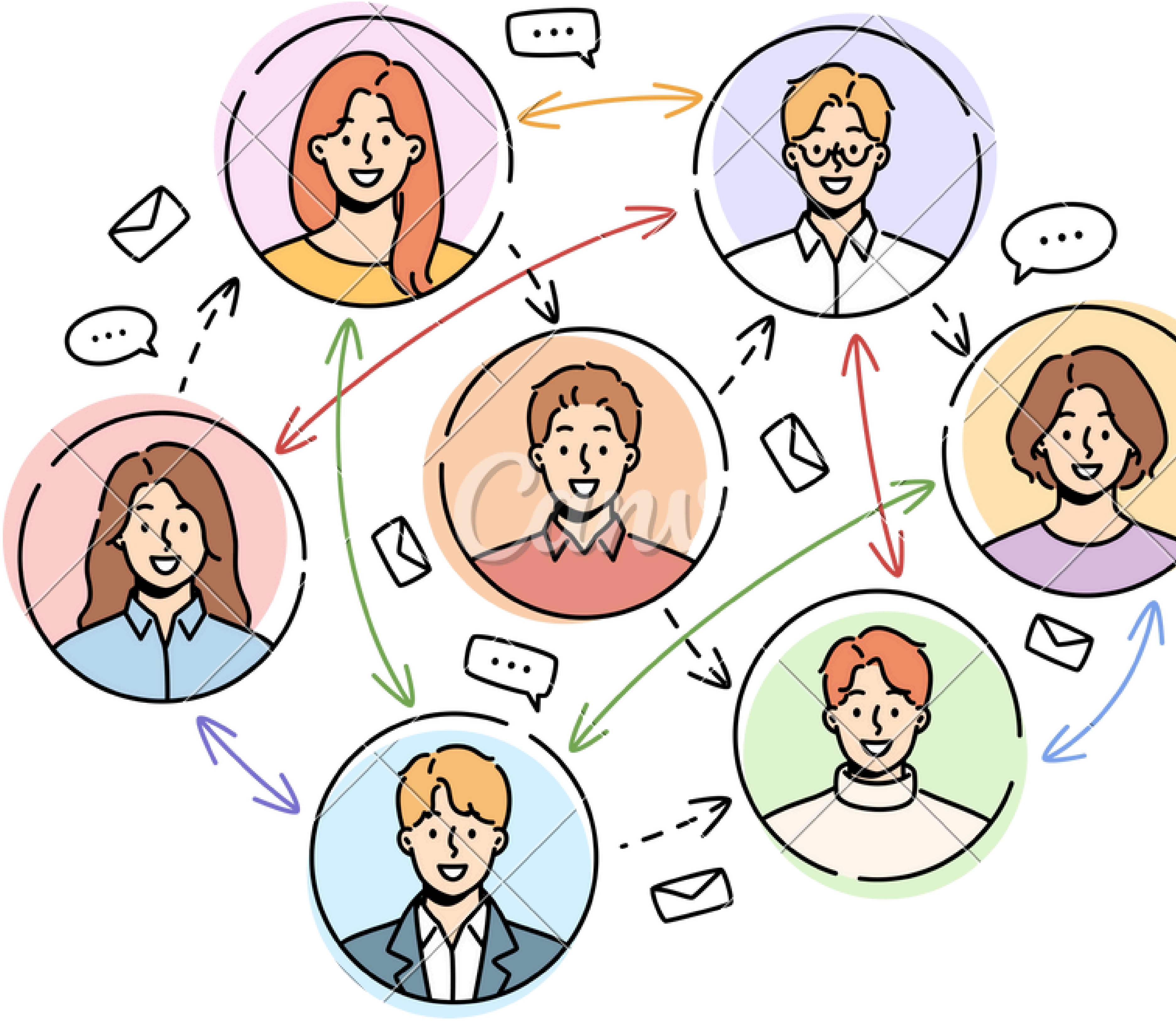
Decoding Shifts in Academic Networks

US-China Collaboration Dynamics in the GPT Era

Presenters: Zhang Ruoxi, Lei Xi, Yang Shu



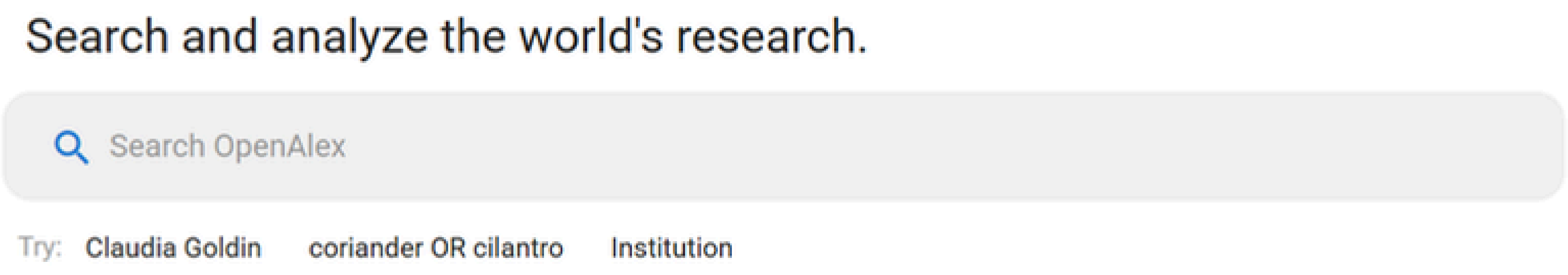
Research Questions

- How has the emergence of GPT technologies reshaped co-authorship and citation networks between Chinese and American research institutions?

- What significant differences can be observed in their collaboration paradigms, disciplinary structures, and knowledge flow characteristics before and after the rise of GPT models?

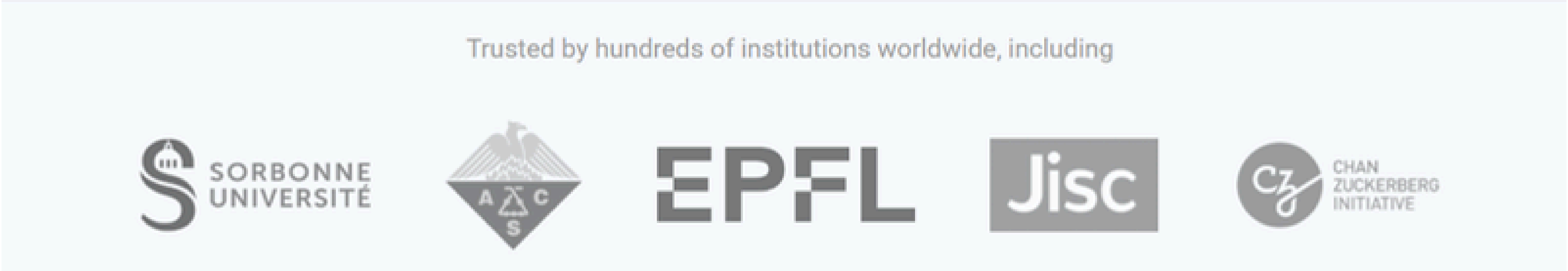
Dataset

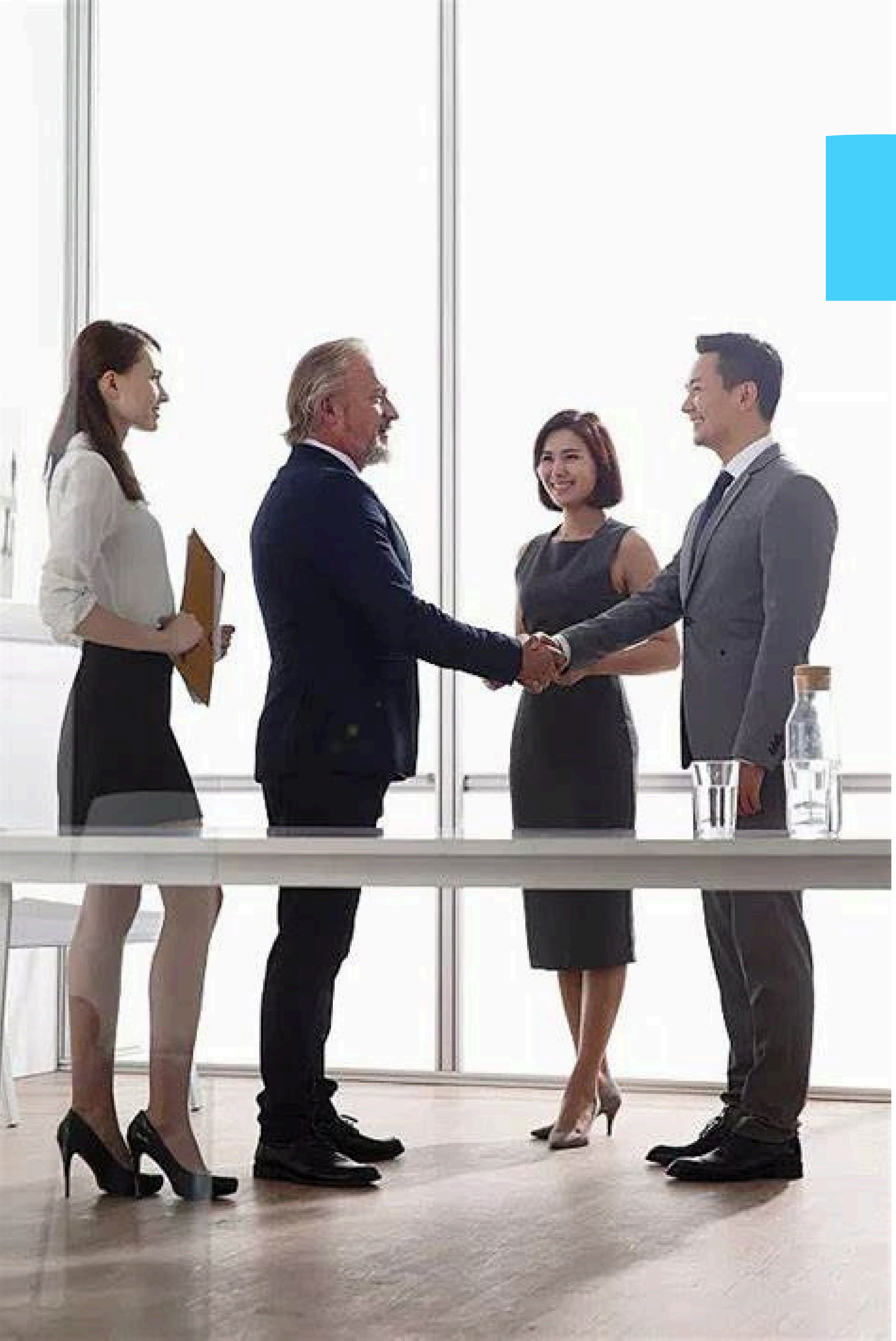
We utilize the OpenAlex dataset to investigate academic publications in the field of computer science from 2015 to 2025, focusing specifically on institutions from China and the US.



Inclusion Criteria:

- Papers in the field of **Computer Science.**
- Affiliation limited to **top institutions:**
  - U.S. (CS Ranking Top30 universities)
  - China (universities involved in the 985 project. Top 30 universities)
- Time Span: **2015-2025**
- Exclude:
  - **Retracted publications**
  - **Non-research content:** such as editorials or metadata-only records
- Only Include:
  - **Impactful papers** (citation percentile value > 0.5)



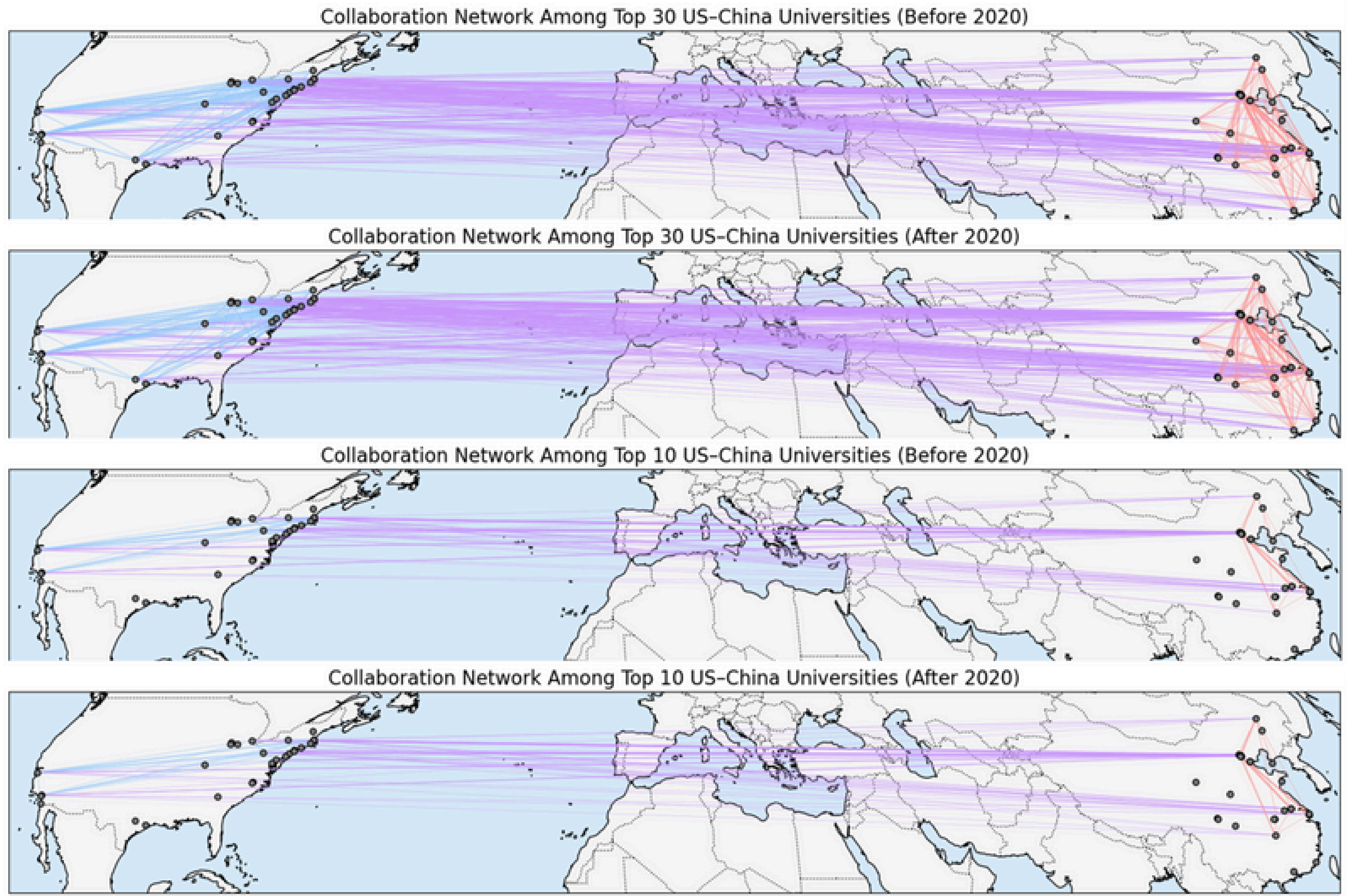


1

About Collaboration Network

- Our collaboration network is constructed based on **co-authorship relations among scholars**, reflecting real academic partnerships.

- We explore the network from three main perspectives:
  - 🌍 **Geographic distribution**:
    What regional patterns emerge in academic collaboration?
  - 📅 **Temporal evolution**:
    How has collaboration changed **before and after 2020**?
  - 👥 **Top 300 scholar network structure**:
    What clustering patterns and community structures are observed among the most prolific scholars?

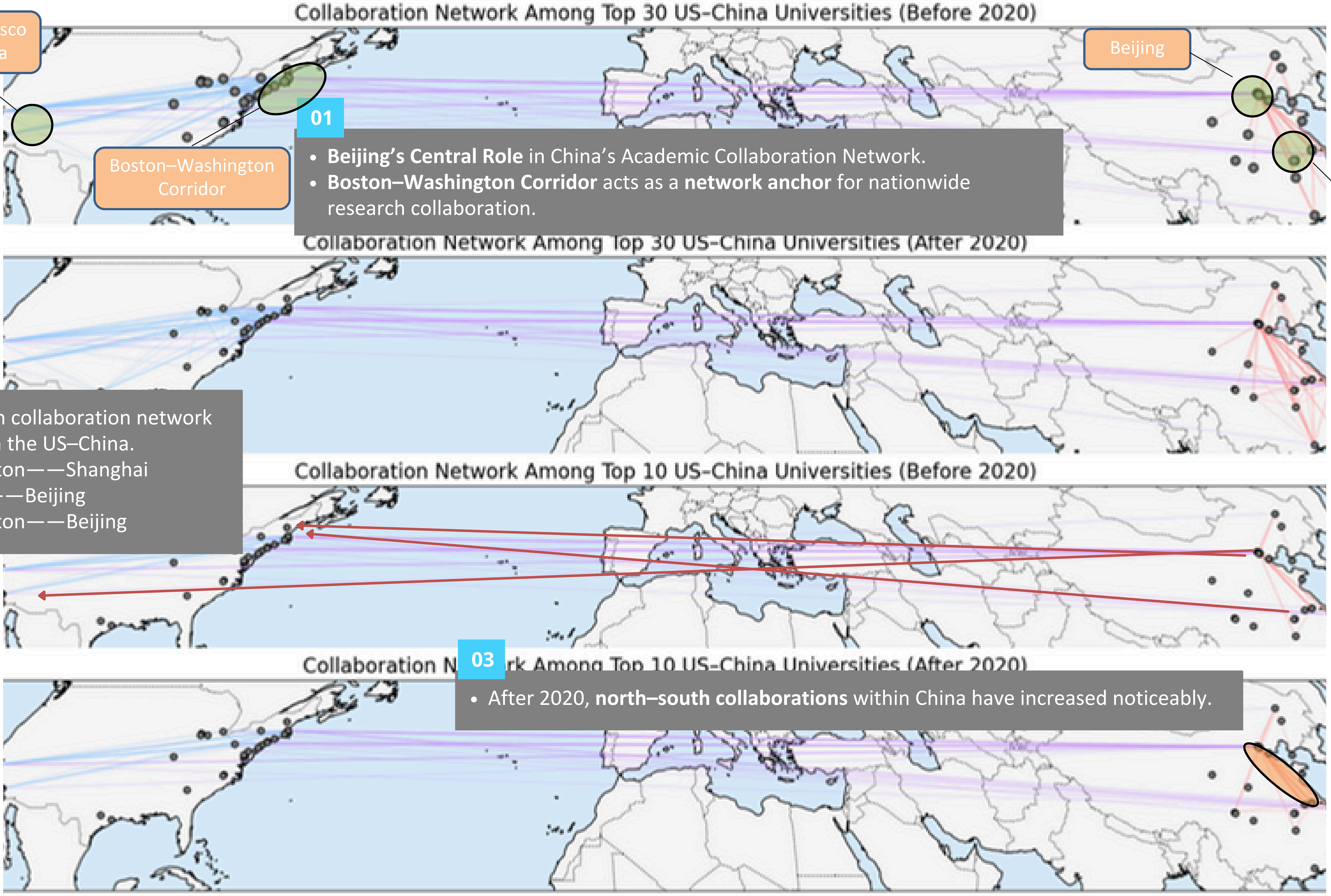


01

Number of collaboration links rose from **1044 to 1090** after 2020.

02

US–China collaborations outnumber domestic ones.



01

- **Beijing's Central Role** in China's Academic Collaboration Network.
- **Boston–Washington Corridor** acts as a **network anchor** for nationwide research collaboration.

02

- The main collaboration network between the US–China.
  - Boston——Shanghai
  - LA——Beijing
  - Boston——Beijing

03

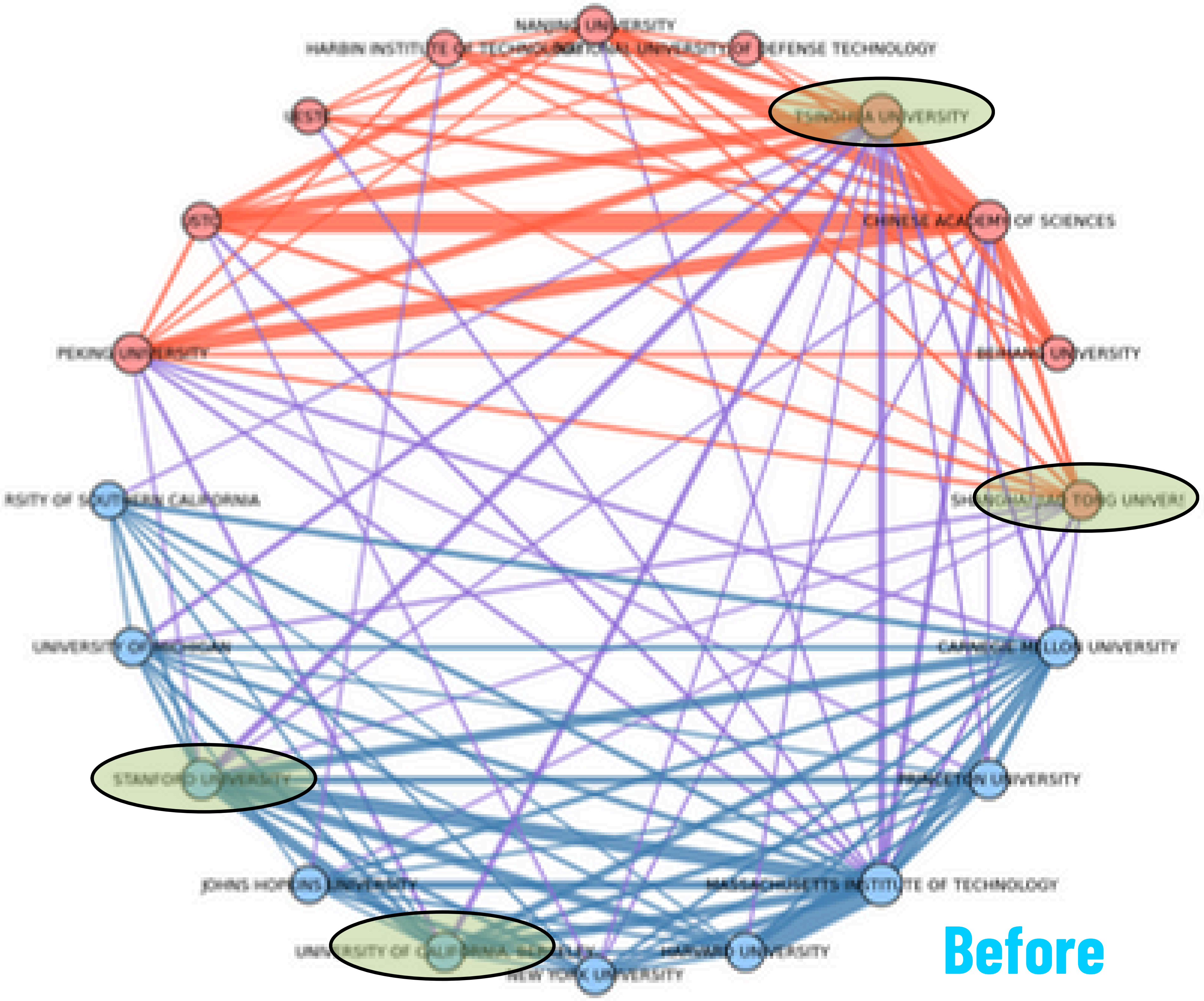
- After 2020, **north–south collaborations** within China have increased noticeably.

Collaboration Network Among Top 10 US-China Universities

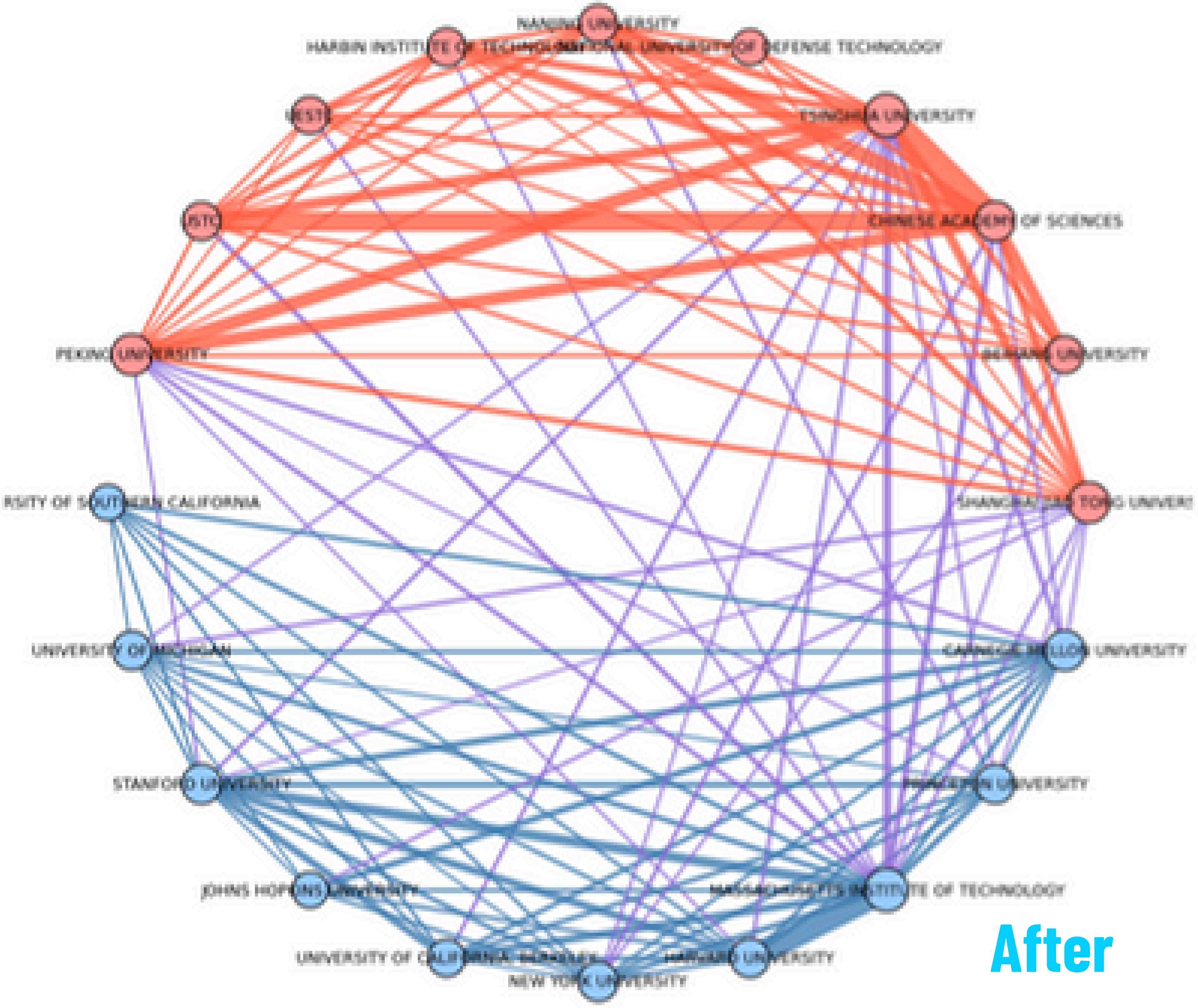
01

- **US–China strong ties decline**: Purple edges become sparser after 2020.
- **Top universities' external links weaken**: Tsinghua, SJTU, Stanford, and UC Berkeley show thinner edges.

Deep collaborations among top-tier universities are diminishing.



Before

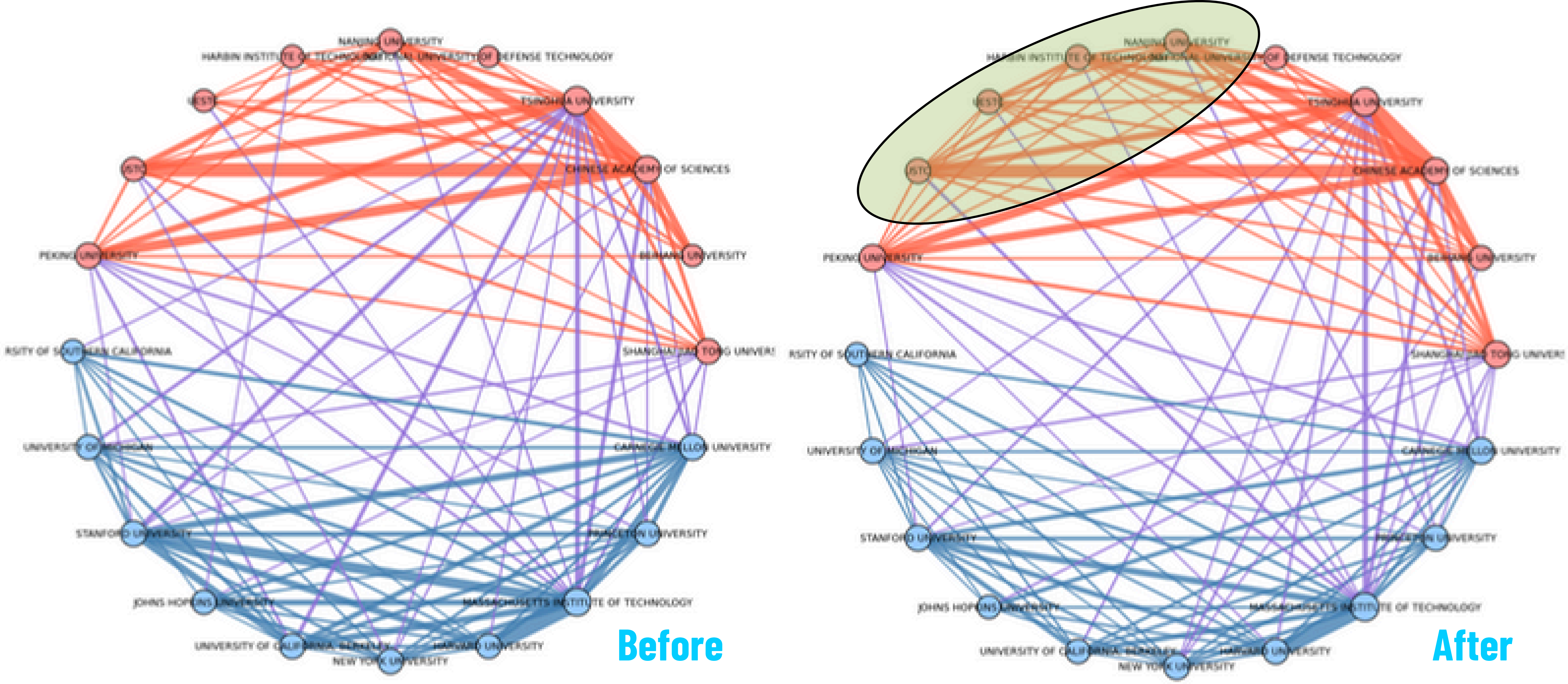


After

Collaboration Network Among Top 10 US-China Universities

02

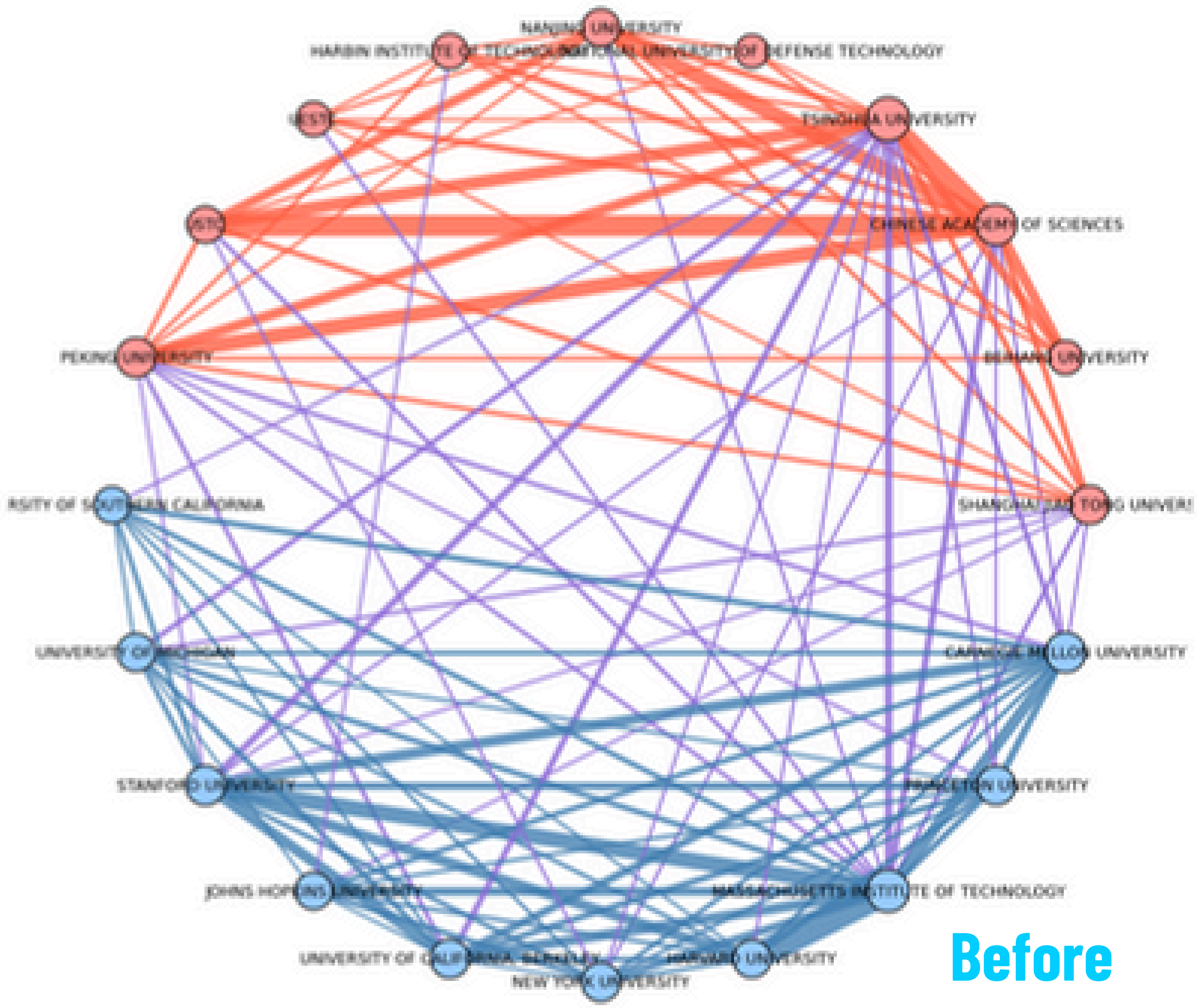
- **US internal collaboration remains stable**: Blue edges show little change in density.
- **Chinese domestic collaboration intensifies**: Red edges among Chinese universities become thicker, showing stronger internal ties.

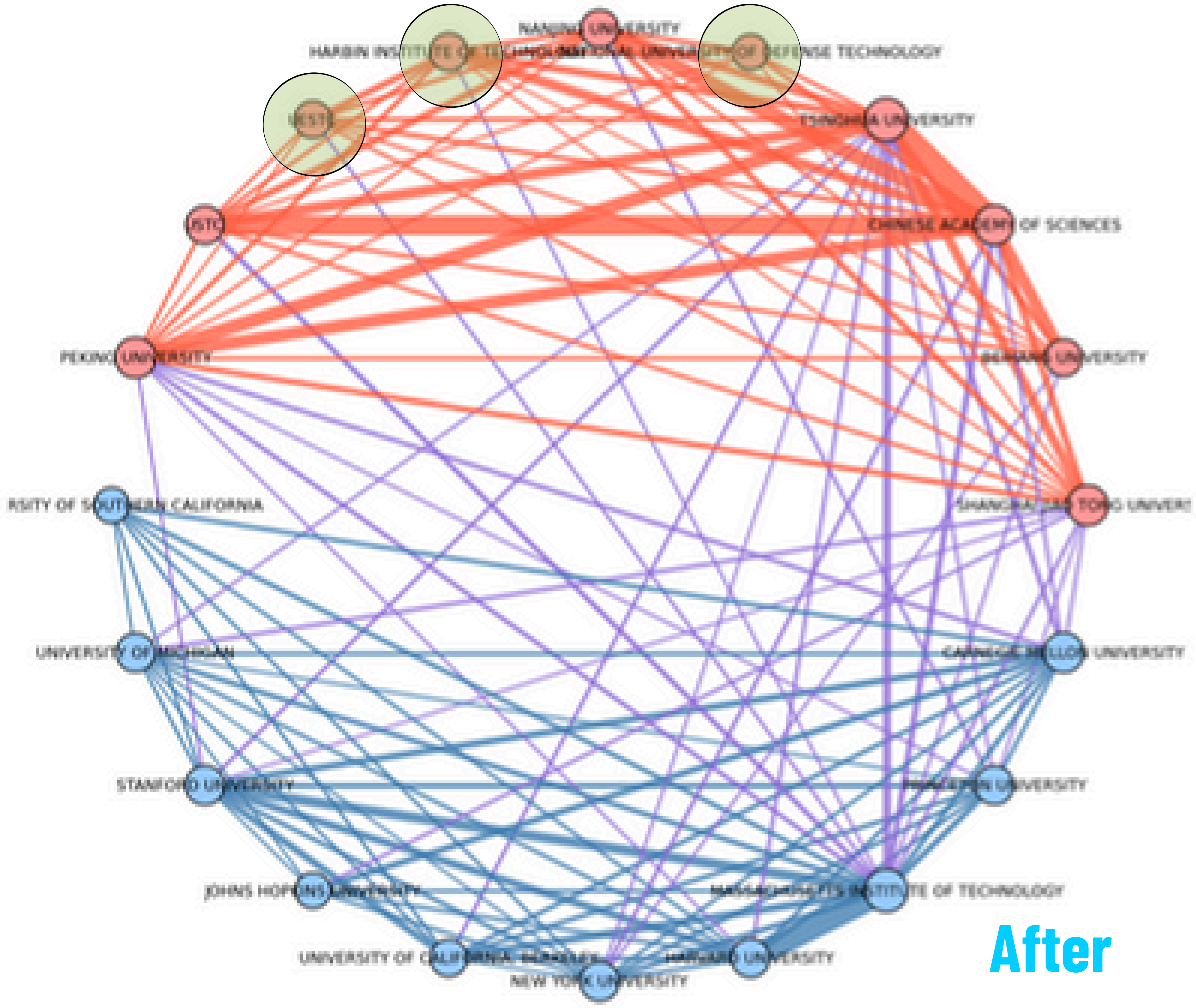


Collaboration Network Among Top 10 US-China Universities

Fun Fact

- **Defense-affiliated Chinese universities (e.g. HIT, UESTC, NUDT)** have **very limited international ties**, consistent with expectations.

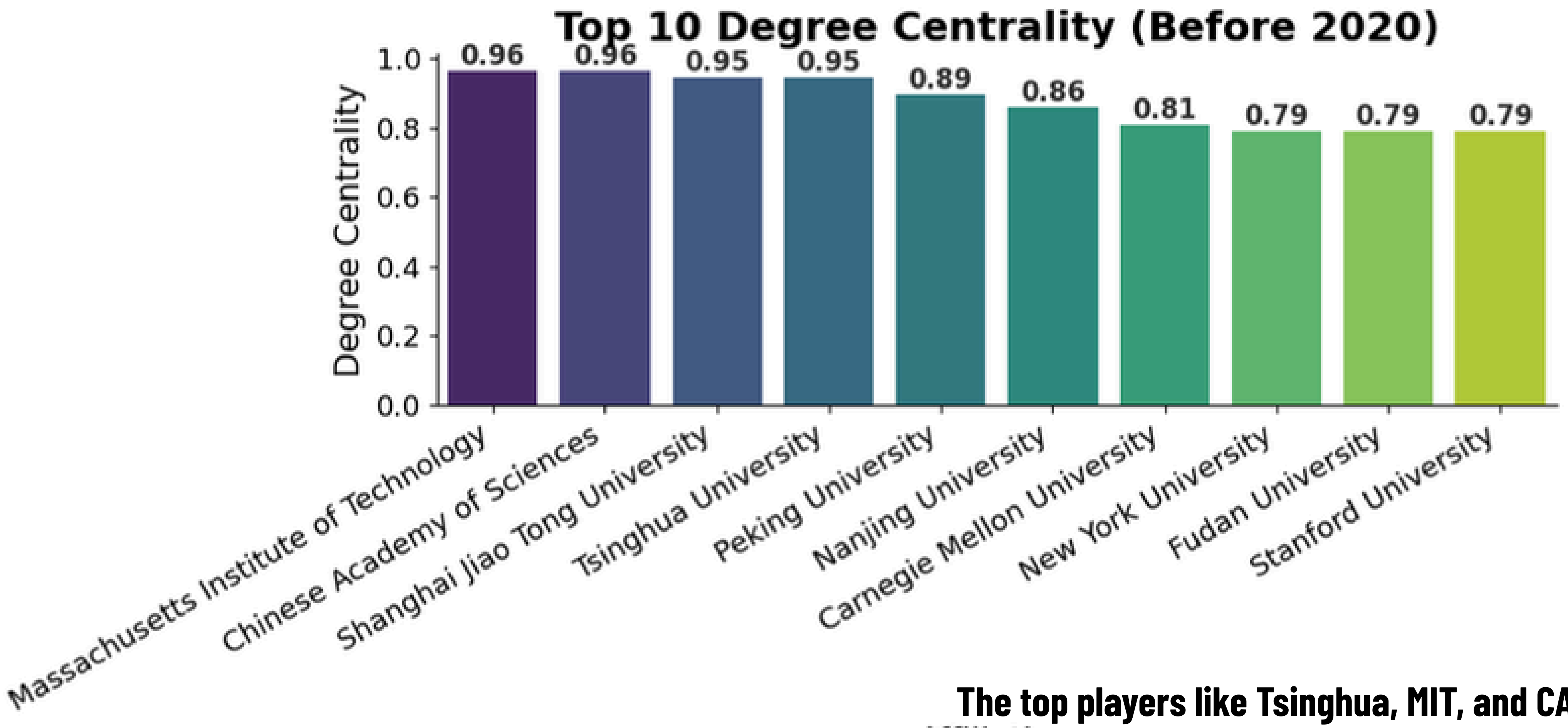


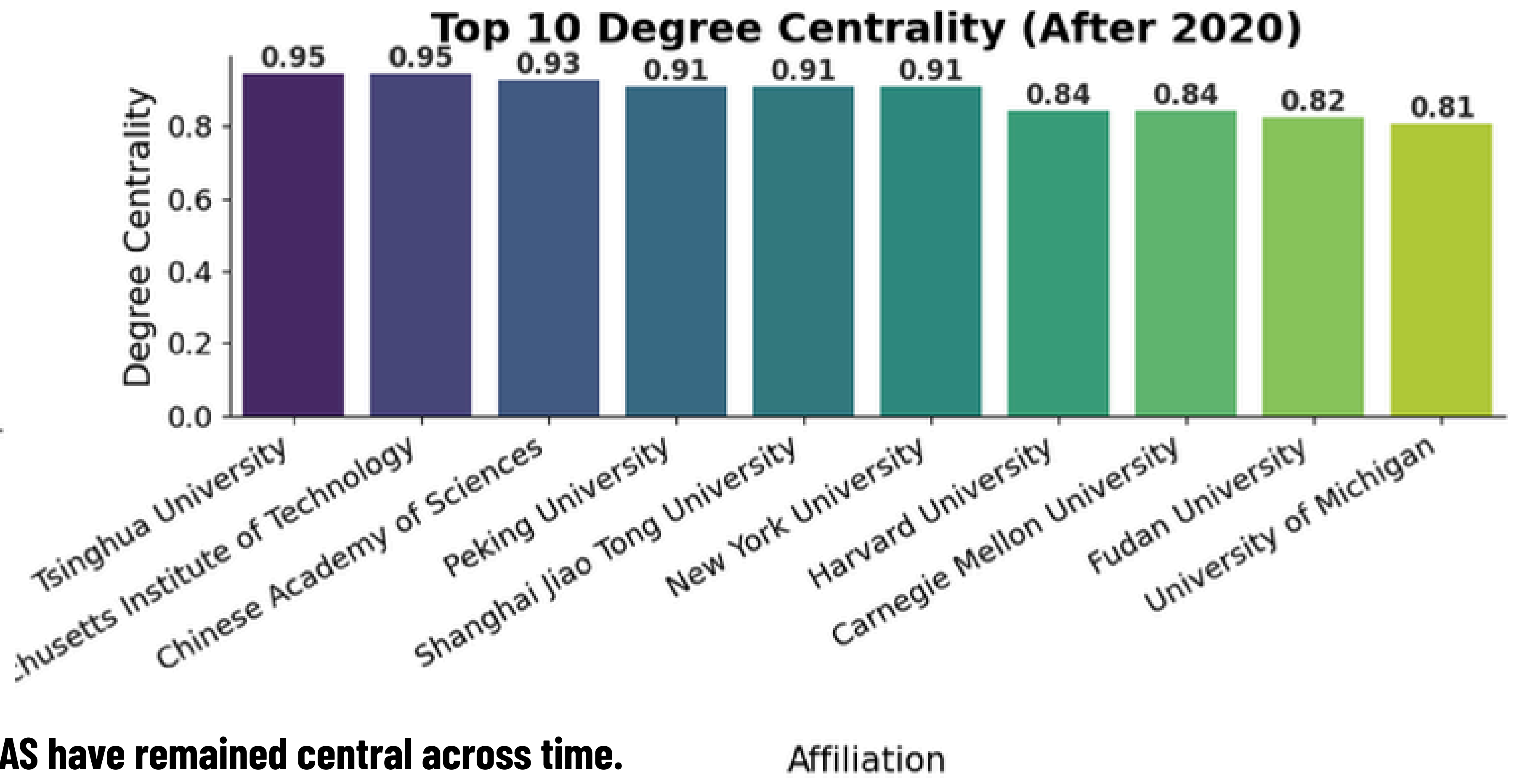


co-authorship

Degree Centrality– Before vs After 2020

Degree centrality measures how many direct connections a node has, indicating its level of activity or importance within a network.





The top players like Tsinghua, MIT, and CAS have remained central across time.

Some U.S. universities (e.g., Harvard, Michigan) gained influence after 2020.

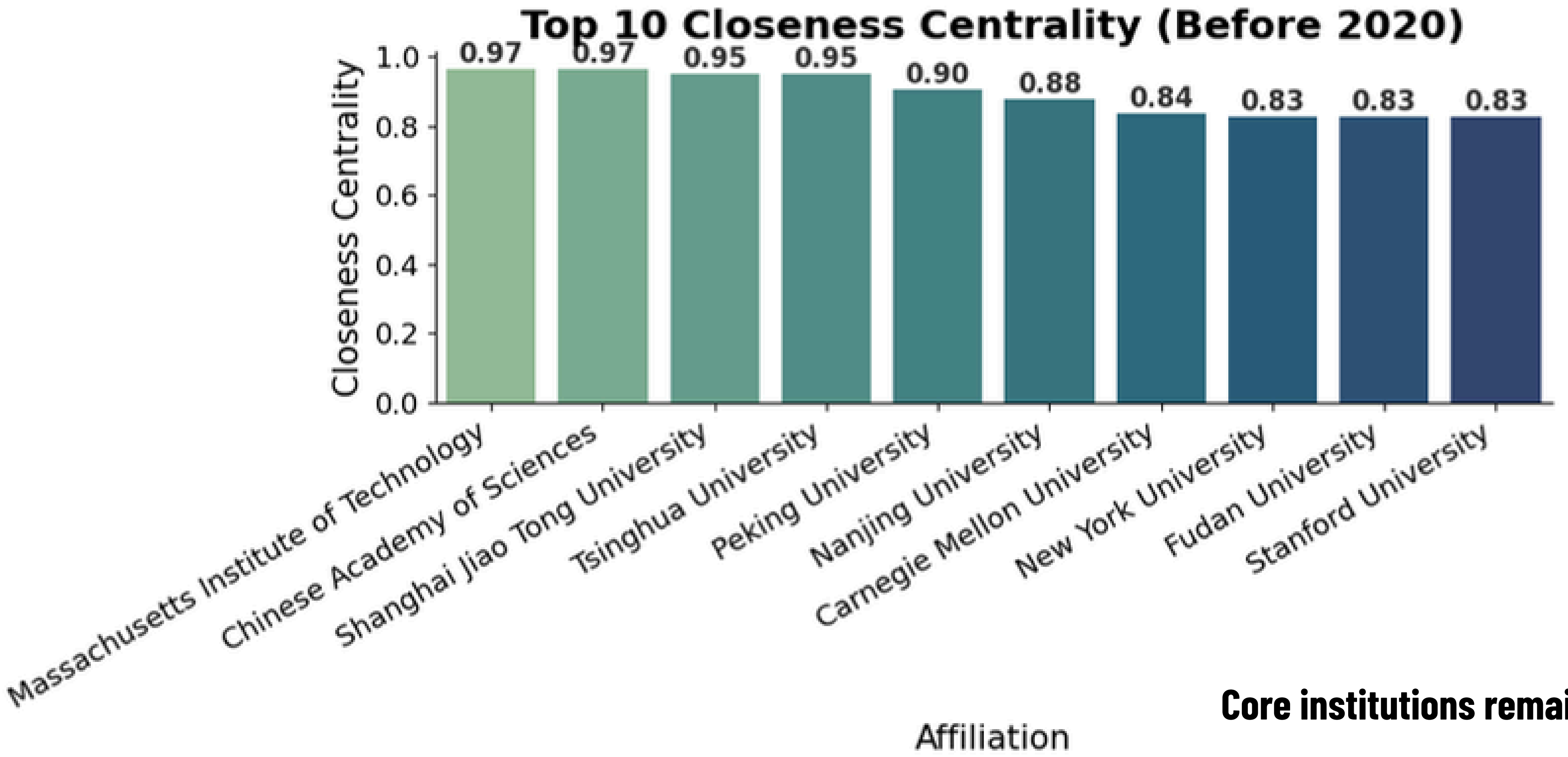
Others (e.g., Stanford, Nanjing) dropped out, showing shifts in collaborative engagement..

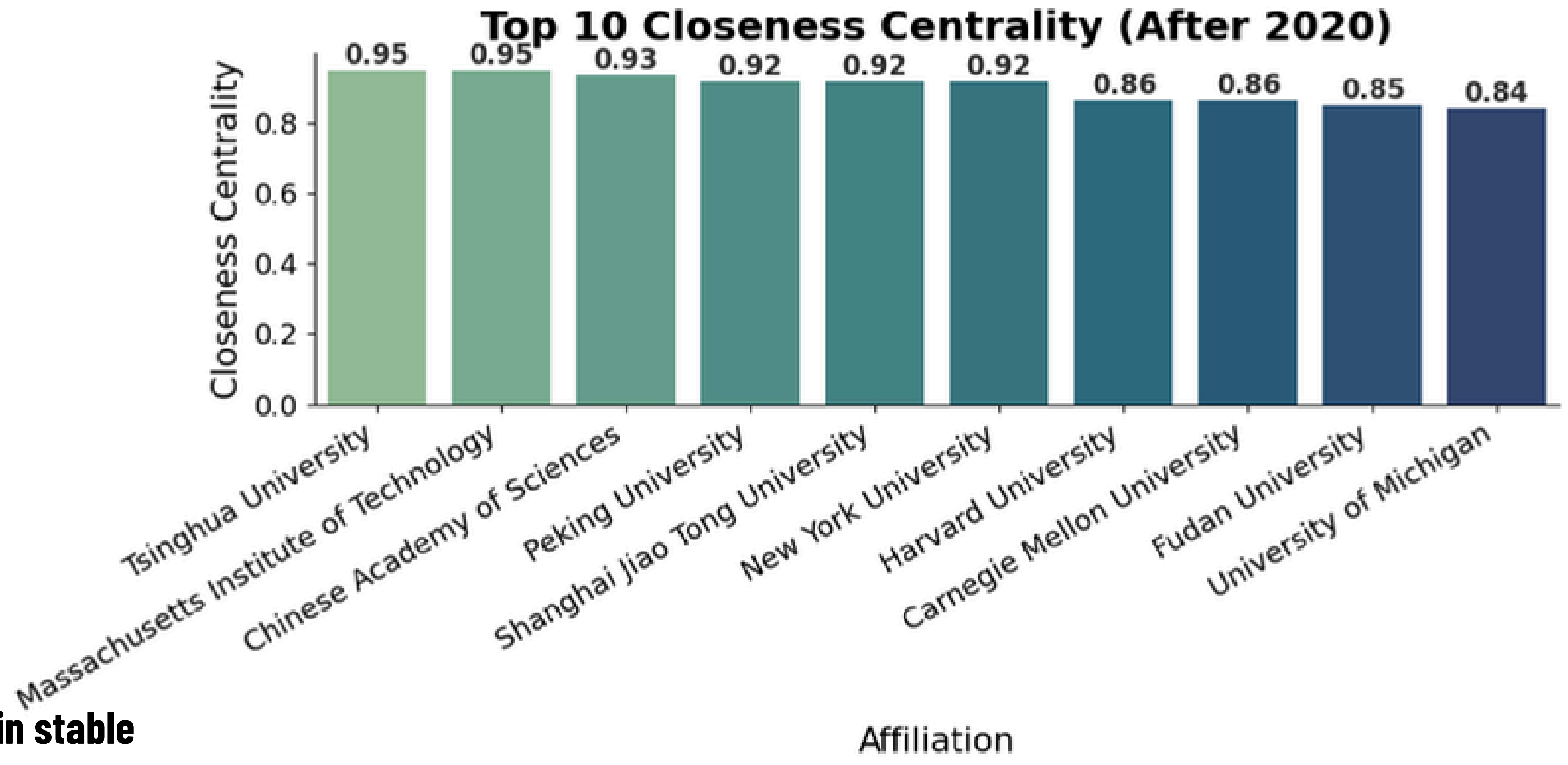
The network shows signs of greater balance, with degree centrality values becoming more even across top institutions.

co-authorship

Closeness Centrality– Before vs After 2020

Closeness centrality measures how close a node is to all other nodes in the network, based on the shortest path distances.





Core institutions remain stable

Tsinghua, MIT, CAS, SJTU, and Peking University stay in the top 5 across both periods.

New U.S. entries after 2020

Harvard and University of Michigan joined the top 10, showing increased connectivity.

Stanford and Nanjing dropped out

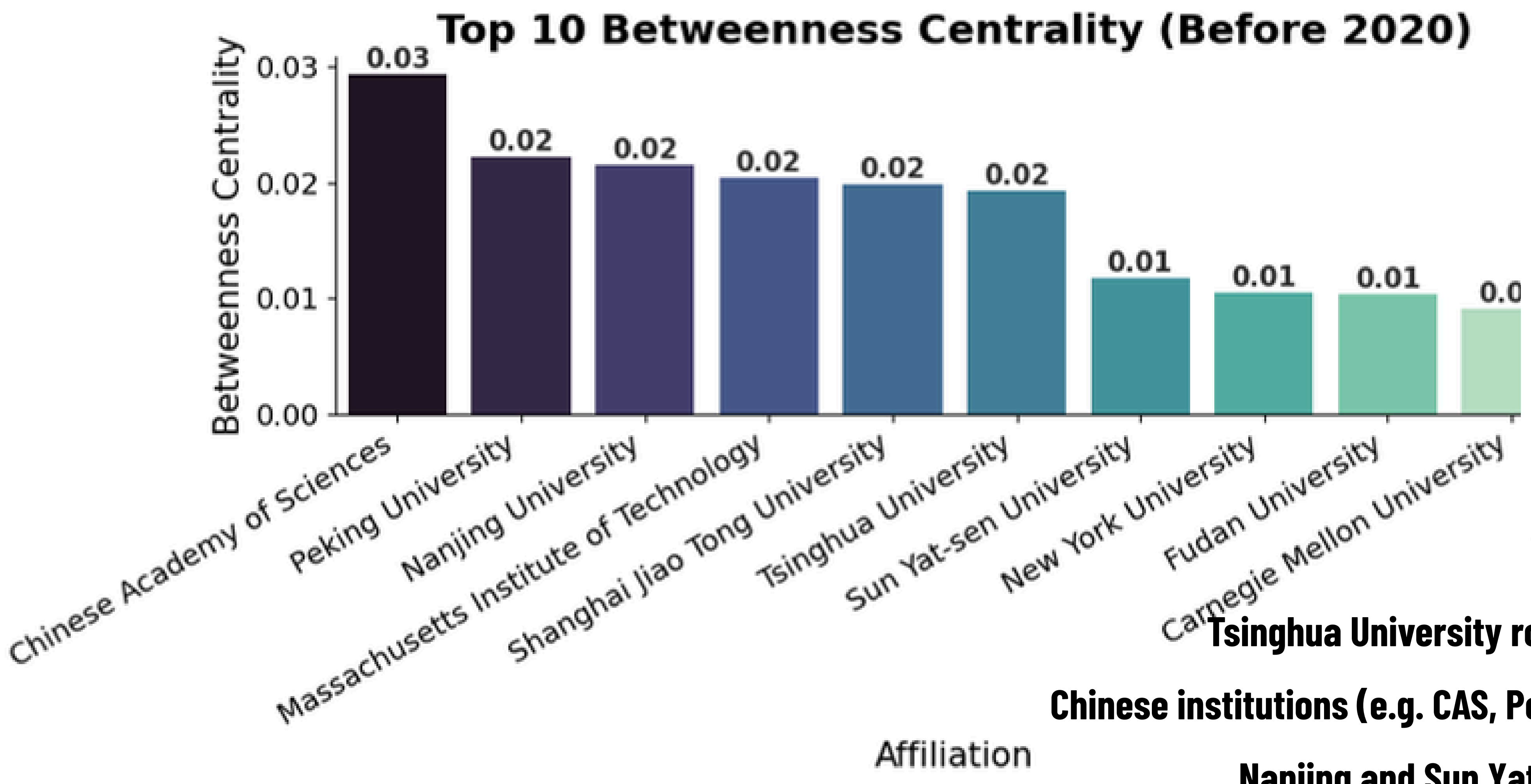
These universities were central before 2020 but lost relative prominence afterward..

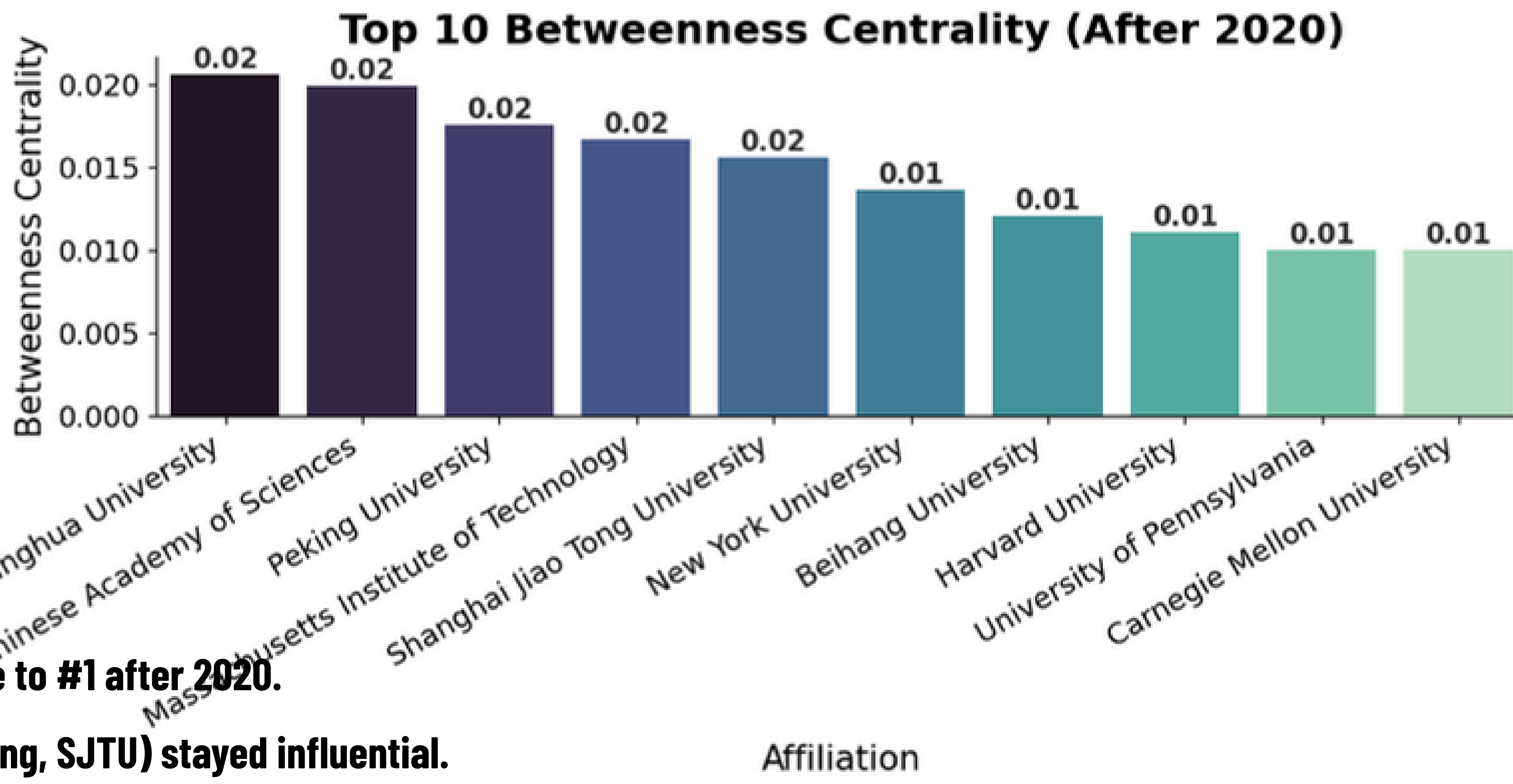
Scores became more even: post-2020 closeness centrality is more balanced, suggesting a more distributed network.

co-authorship

Betweenness Centrality Before vs After 2020

Betweenness centrality measures how often a node lies on the shortest paths between other nodes, indicating its role as a bridge or intermediary in the network.





Tsinghua University rose to #1 after 2020.

Chinese institutions (e.g. CAS, Peking, SJTU) stayed influential.

Nanjing and Sun Yat-sen dropped out.

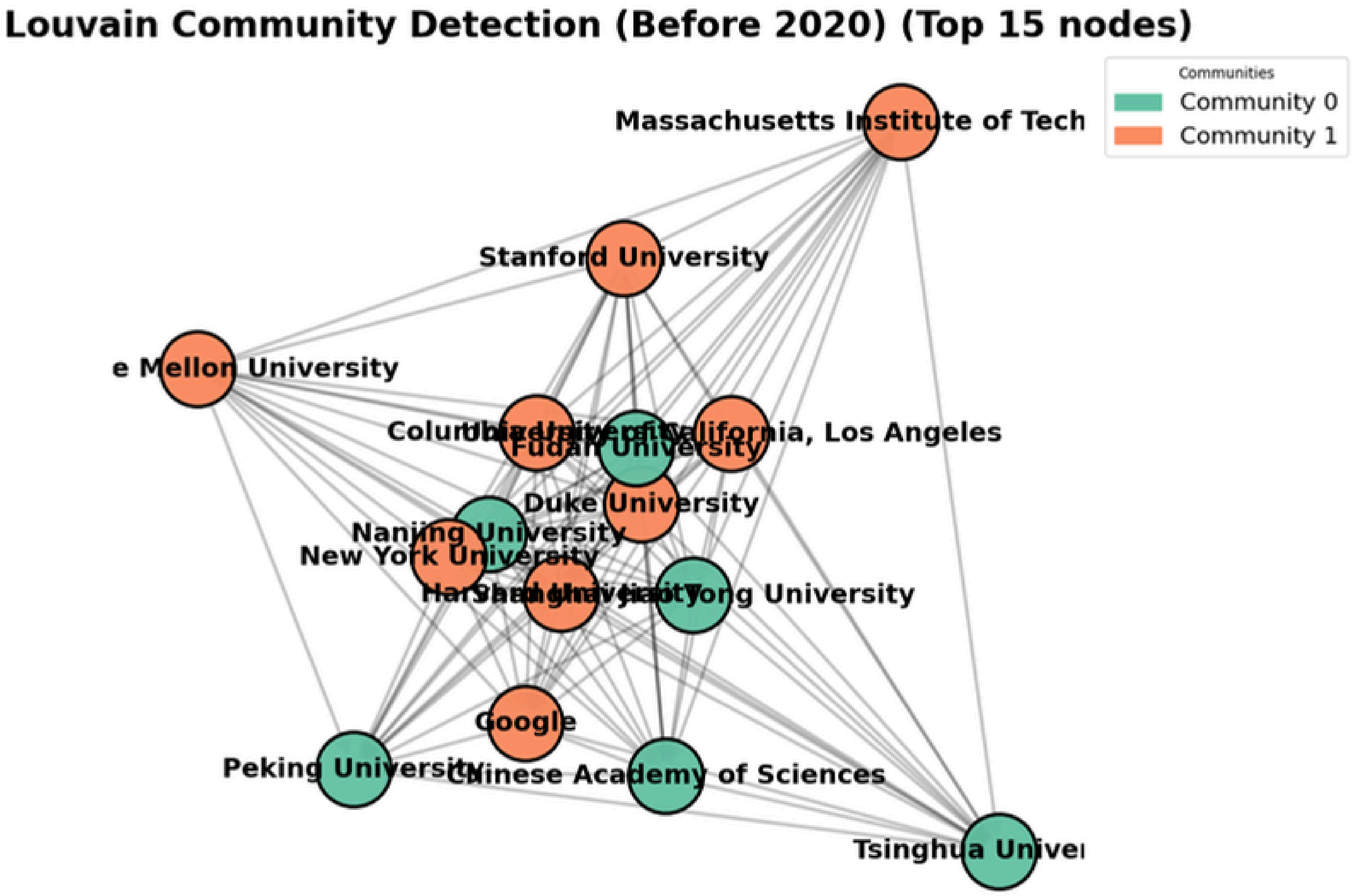
New entries: Beihang, Harvard, UPenn.

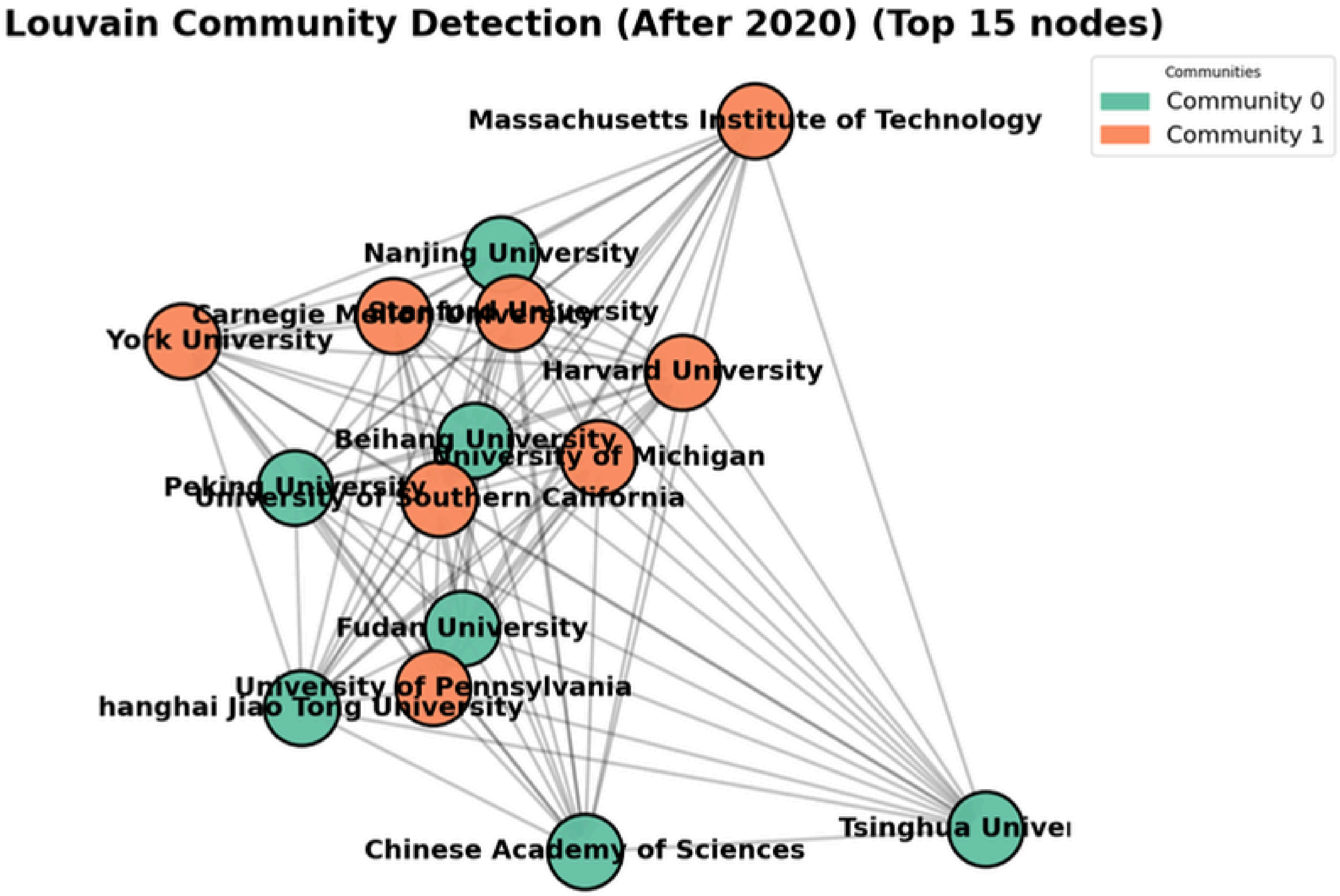
- **Overall decrease in betweenness centrality scores**

it suggests that the academic collaboration network has become less centralized, meaning that the flow of information or collaboration is no longer dominated by a few key institutions. Instead, more institutions are participating in bridging roles, indicating a more distributed structure.

co-authorship Louvain Community Before vs After 2020

Louvain Community Detection identifies clusters of nodes in a network by maximizing modularity, revealing groups with dense internal connections and sparse external links.





- ## After 2020, More Mixed Communities

After 2020, Chinese and US universities appear more frequently within the same communities.

Institutions like New York University and the University of Pennsylvania are now grouped with Chinese universities such as Tsinghua and Fudan.

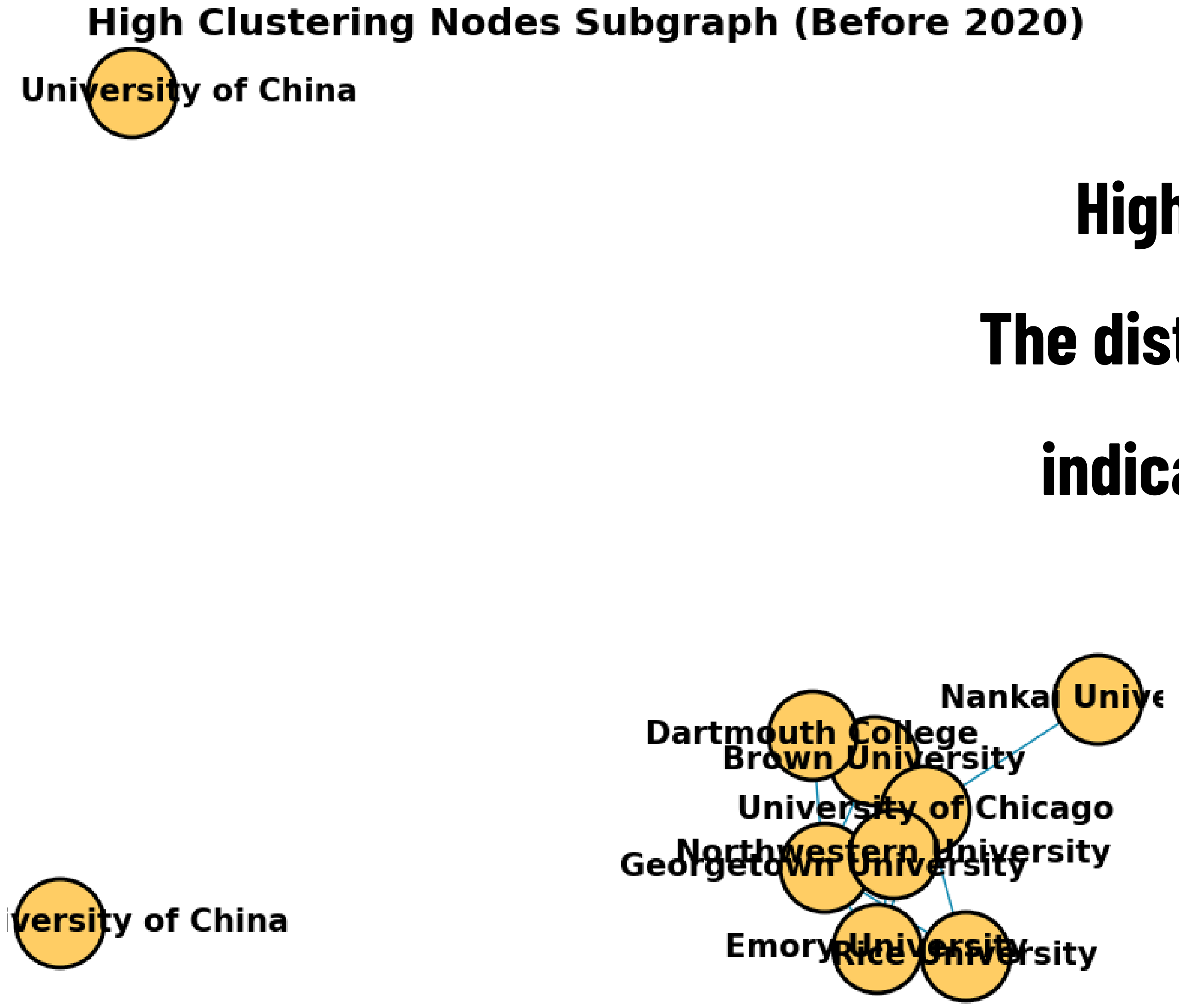
This indicates stronger cross-national collaboration.

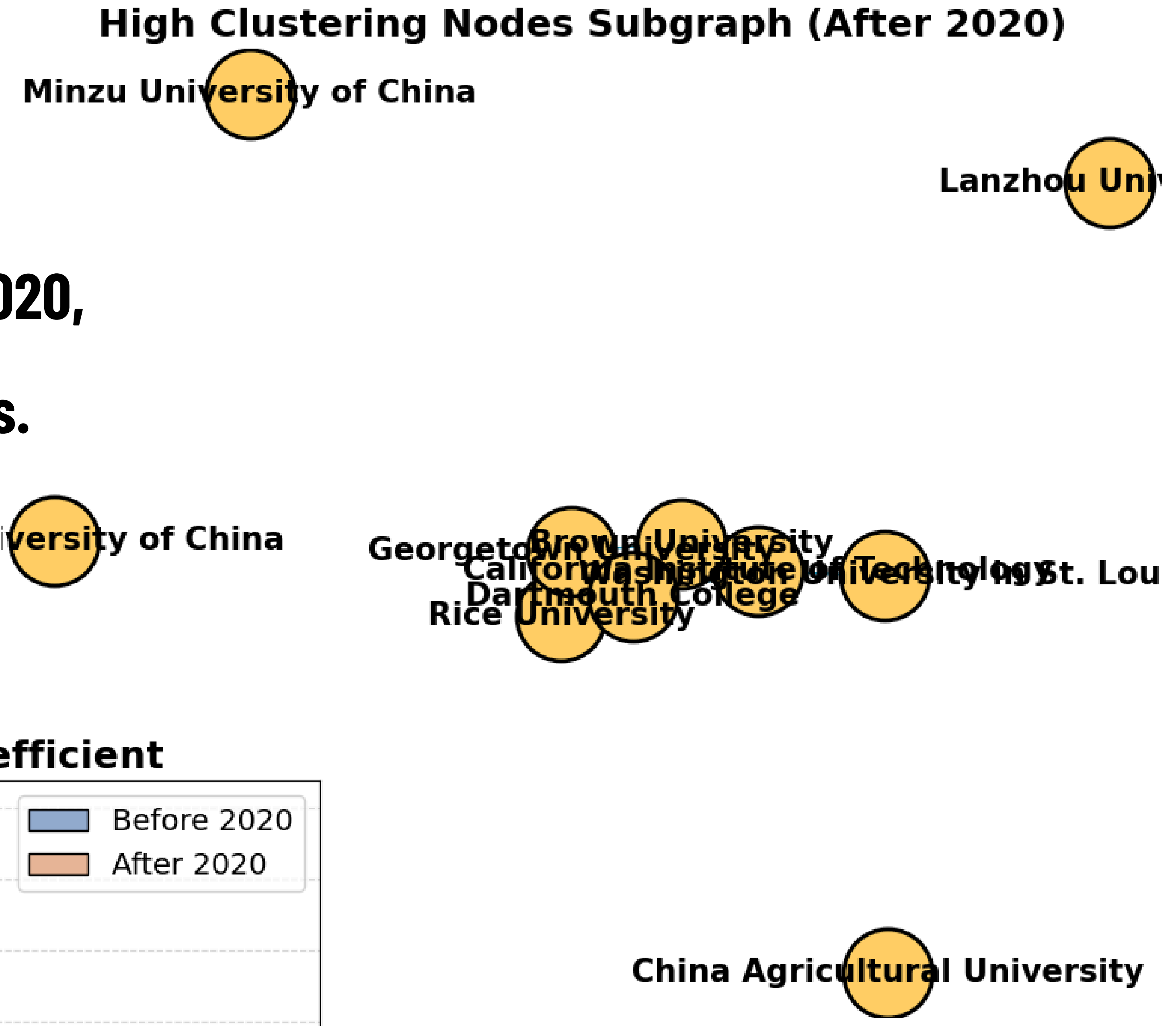
- ## Tighter Local Clusters

Post-2020, institutions within the same community—often from the same country or region—form denser internal connections, reflecting stronger domestic collaboration and tighter local clusters.

co-authorship

Changes in Clustering Coefficient and Network Structure

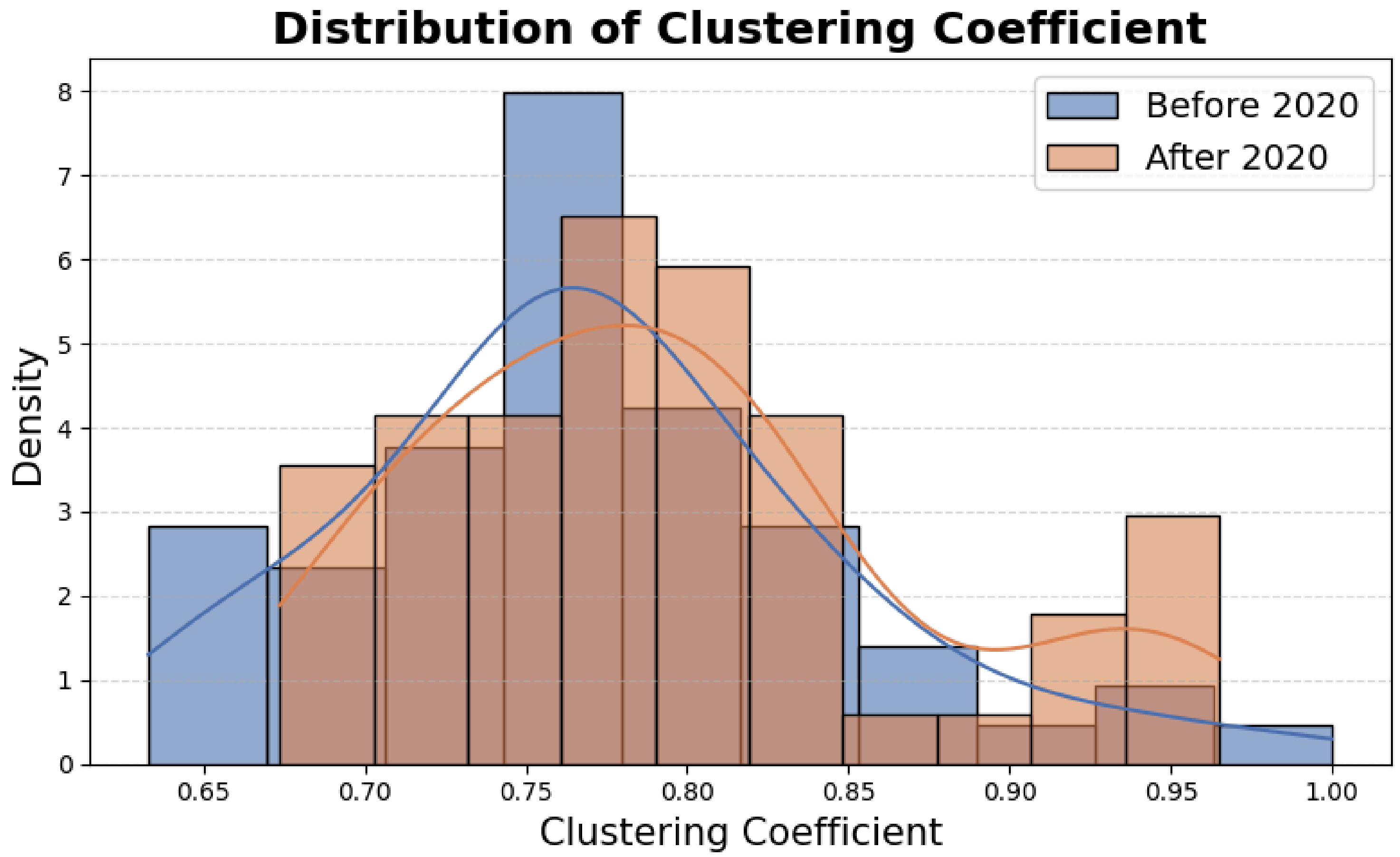




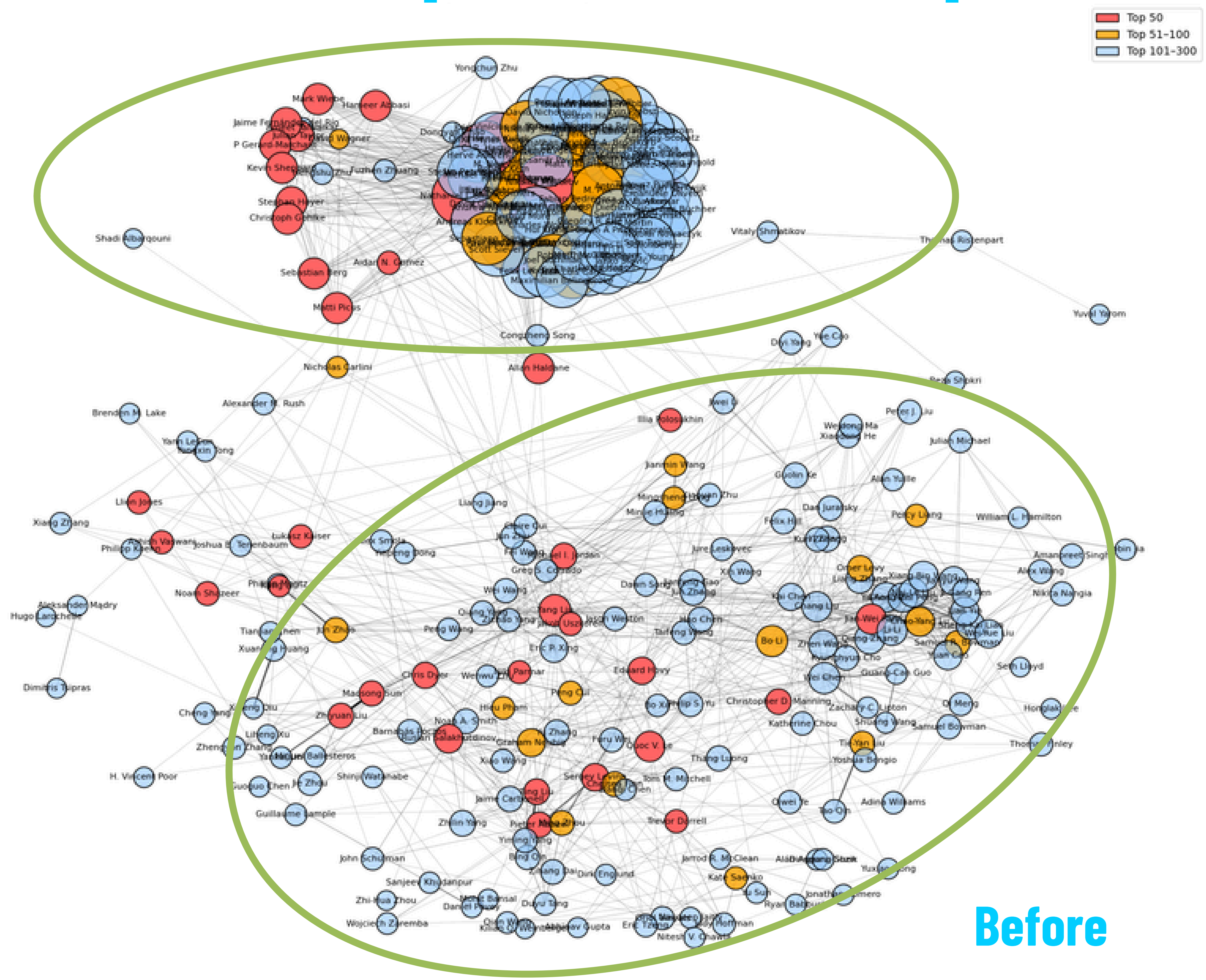
Higher Clustering Coefficients After 2020

The distribution shifted slightly right after 2020,

indicating more tightly local collaborations.



Co-authorship Network of Top 300 Authors

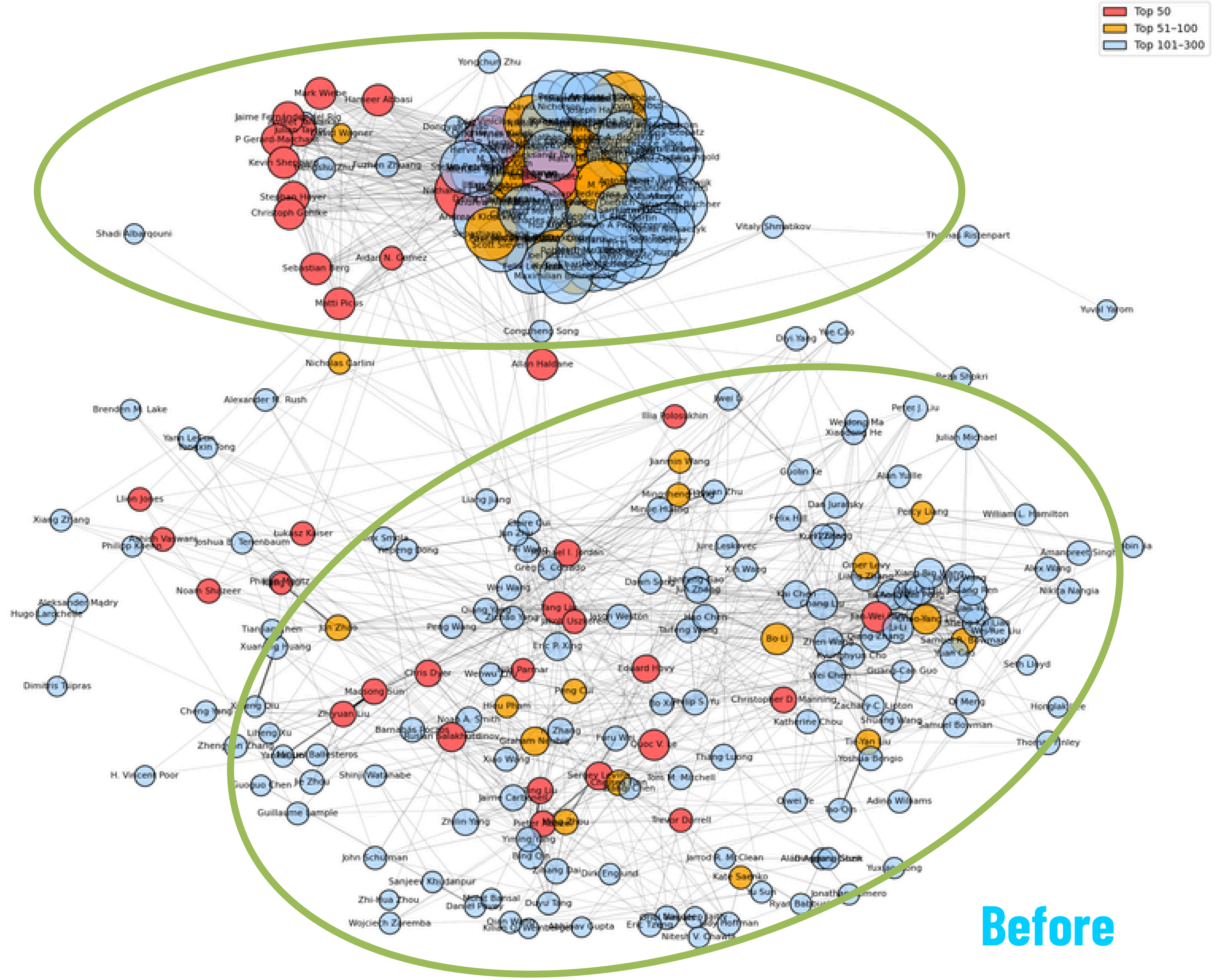


Before

Network Structure Summary

- **Nodes represent scholars**
- **Edges indicate co-authorships**
- **Node size** = number of co-authors (**degree**)
- **Edge thickness** = collaboration intensity (more co-authored papers)
- **Node color indicates citation rank**:
  - **Red** = Top 50 scholars
  - **Orange** = Top 51–100 scholars
  - **Blue** = Top 101–300 scholars
- **Two major communities emerge** *(via Girvan–Newman split)*
  - A densely connected cluster (mostly mid-ranked Chinese institutions)
  - A more dispersed cluster with many top-ranked US scholars

Co-authorship Network of Top 300 Authors



Before

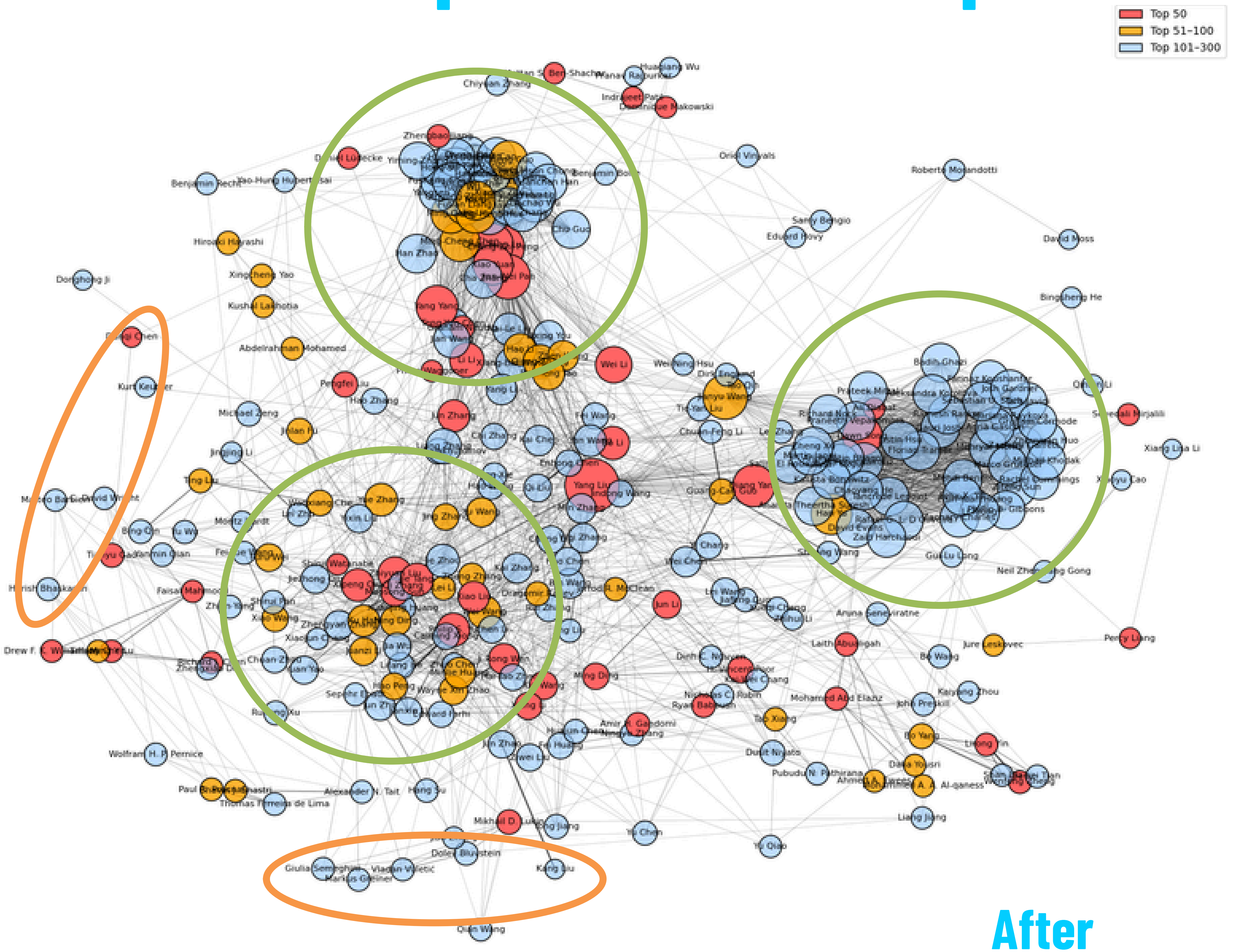
Top Community: Densely Connected Core

- Forms a **tightly knit "core network"** with near-complete internal connectivity.
- Includes scholars from **all ranking tiers.**
- Large central nodes suggest scholars with high degree—likely influential connectors or coordinators.
- **Top 50 scholars** tend to have **fewer collaborators**, but their influence comes from **high-impact publications**.
- **Top 101–300 scholars** adopt a **broad-collaboration strategy**, publishing more papers with more co-authors to build influence.

Bottom Community: Loosely Structured & Dispersed

- More **fragmented and decentralized**, with weaker internal cohesion.
- Top 50 scholars (red nodes) are spread across the periphery, not forming a single tight block.
- Characterized by bridge-builders and cross-institutional cooperation.
- Likely reflects a **more international and mobile collaboration style**, rather than a localized academic cluster.

Co-authorship Network of Top 300 Authors



After

Key Changes in the Co-authorship Network After 2020

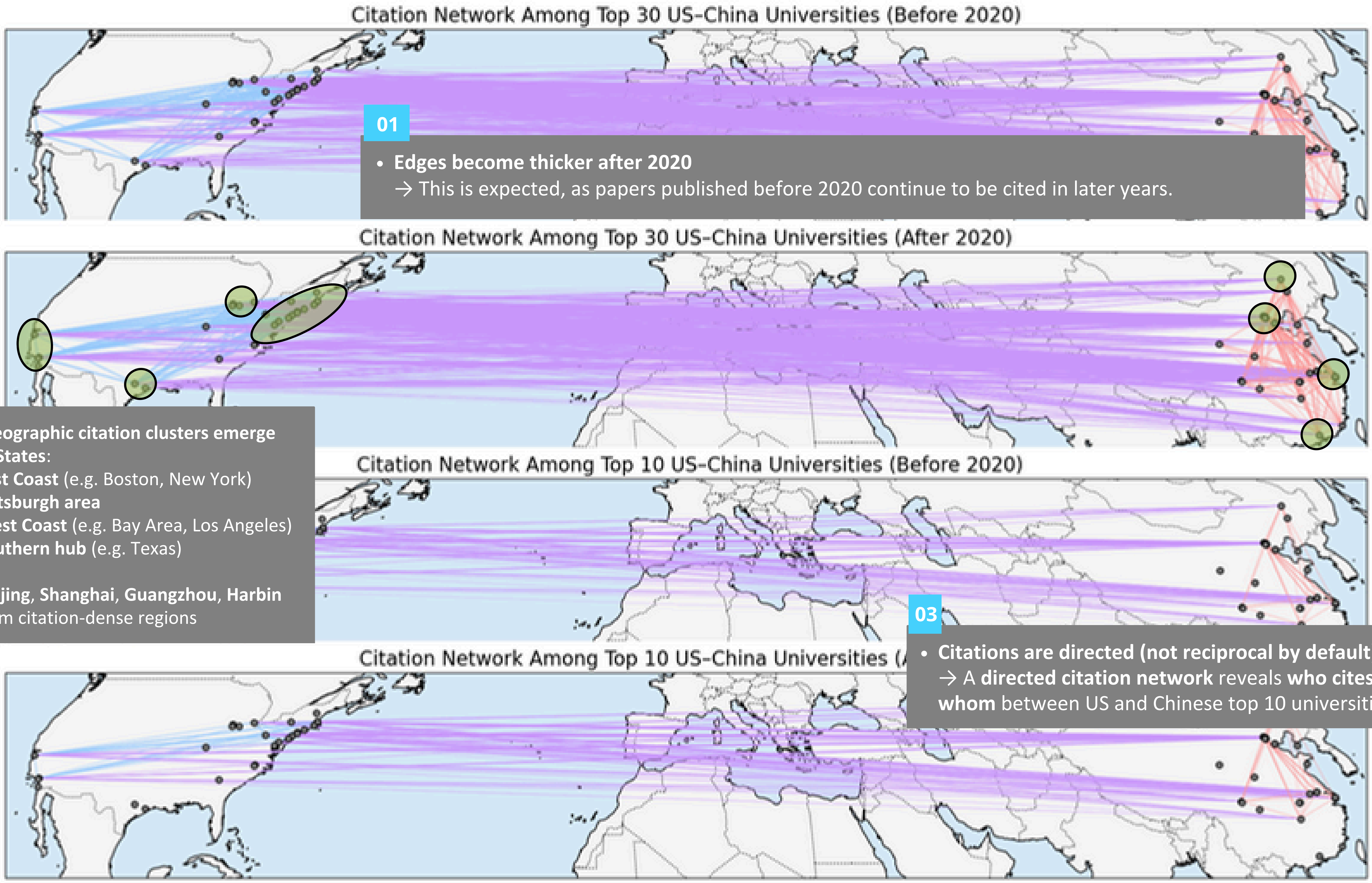
- **The overall network becomes denser but more decentralized**
  → Multiple local clusters emerge; no longer dominated by a single core.
- **Top 50 scholars are more dispersed**
  → **Red nodes are widely spread**, with some located at the periphery.
  → No unified leadership core among top-tier scholars.
- **Top 101–300 scholars are more active and central**
  → Blue nodes play key roles as **connectors across different sub-communities.**
- **More isolated or weakly connected nodes appear at the edges**
  → May indicate the entry of **new researchers or interdisciplinary scholars**
  → These scholars are still in the process of integrating into the main network.

2

About Citation Network

- Our citation network is built by tracing **which papers cite which**, reflecting the **flow of academic influence** between institutions and countries.

- We explore citation patterns from two key perspectives:
  - 🌍 **Geographic distribution**:
    How are citations concentrated across different regions and institutions?
  - 📅 **US–China academic impact over time**:
    How has the **citation quality and influence** shifted between the US and China **before and after 2020**?



01

- **Edges become thicker after 2020**
  → This is expected, as papers published before 2020 continue to be cited in later years.

02

Distinct geographic citation clusters emerge

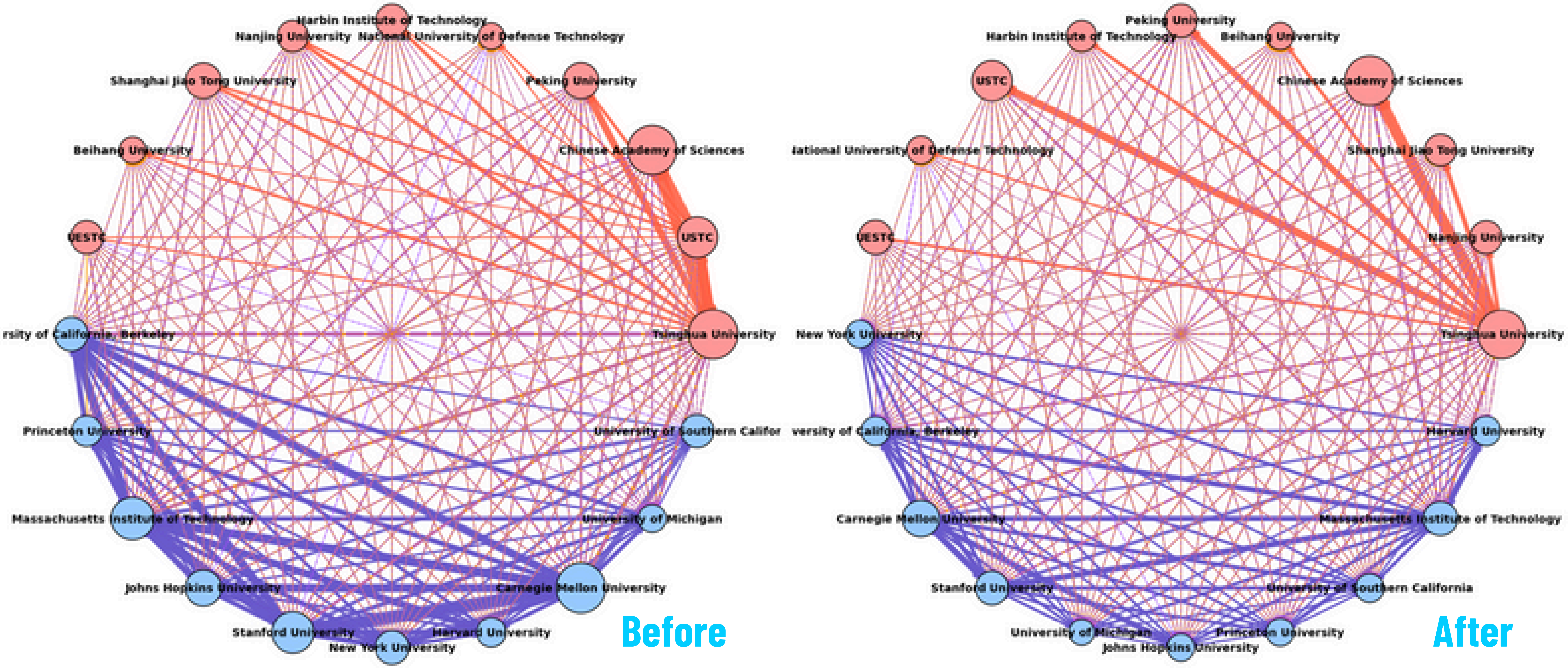
- **United States**:
  - **East Coast** (e.g. Boston, New York)
  - **Pittsburgh area**
  - **West Coast** (e.g. Bay Area, Los Angeles)
  - **Southern hub** (e.g. Texas)
- **China**:
  - **Beijing, Shanghai, Guangzhou, Harbin** form citation-dense regions

03

- **Citations are directed (not reciprocal by default)**
  → A **directed citation network** reveals **who cites whom** between US and Chinese top 10 universities.

Citation Network Among Top 10 US-China Universities (A First Look)

- **Red edges**: Citations **within Chinese universities**
- **Blue edges**: Citations **within US universities**
- **Purple edges**: Citations **from China to the US**
- **Orange edges**: Citations **from the US to China**



Citation Network Among Top 10 US-China Universities

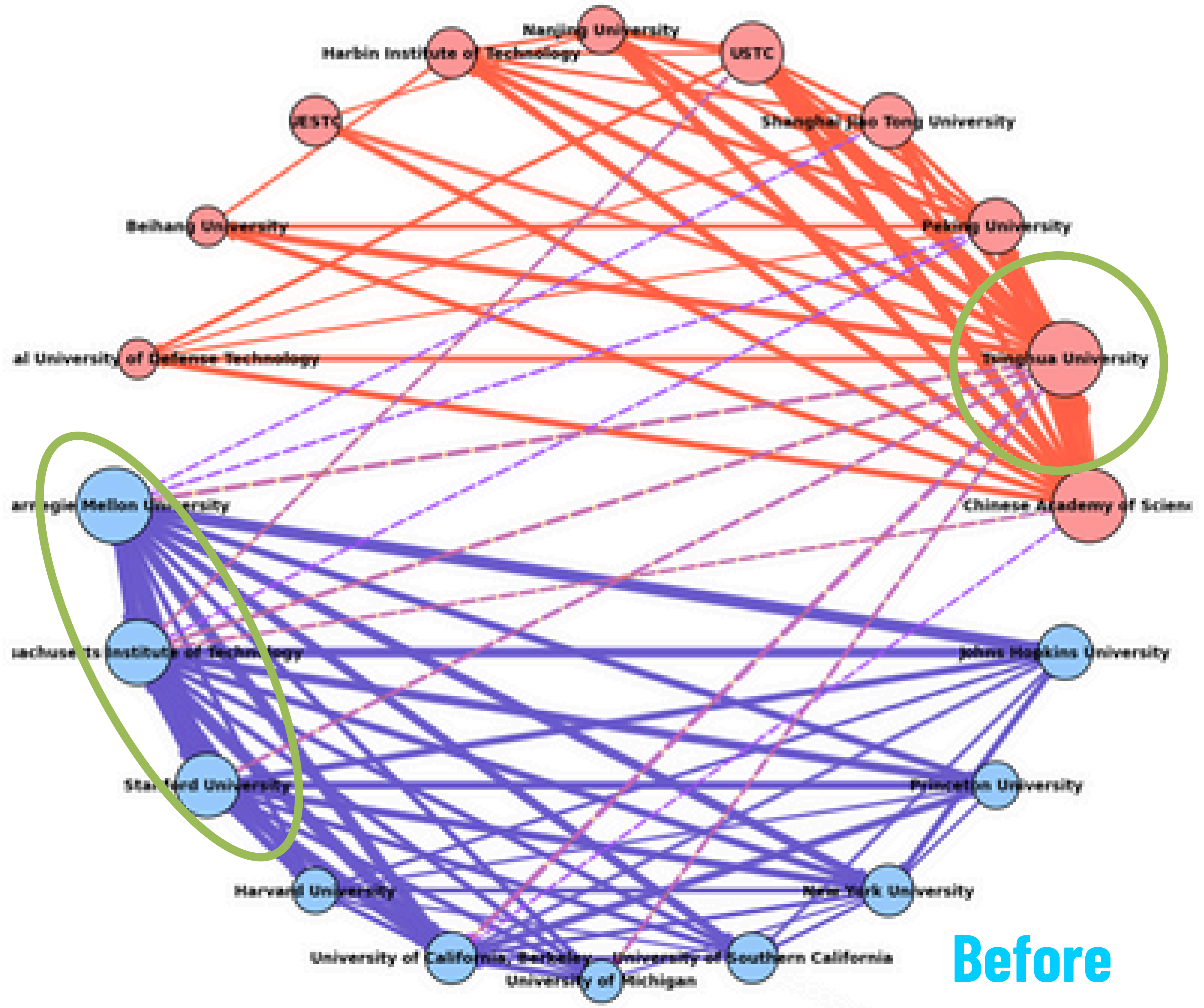
The trend

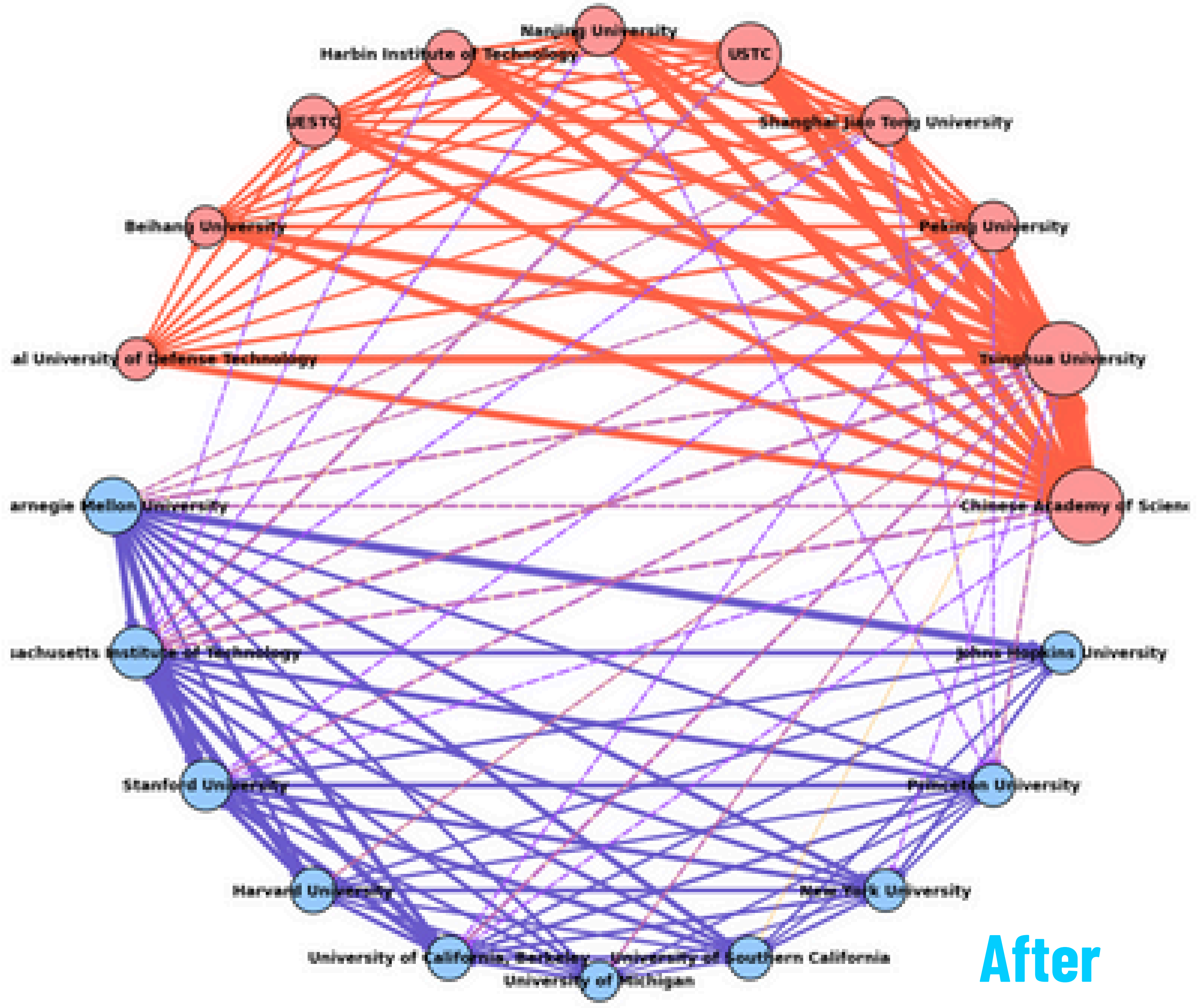
- **China's self-citation increased**, while **US internal citations decreased**
- → Denser red edges among Chinese universities; lighter blue edges among US schools.
- **Before 2020: Citations mainly flowed from China to the US**
- **After 2020: US universities increasingly cite Chinese work**
- → Indicates **rising academic recognition and quality** of Chinese research.

Citation Network Among Top 10 US-China Universities

Who is leading the way

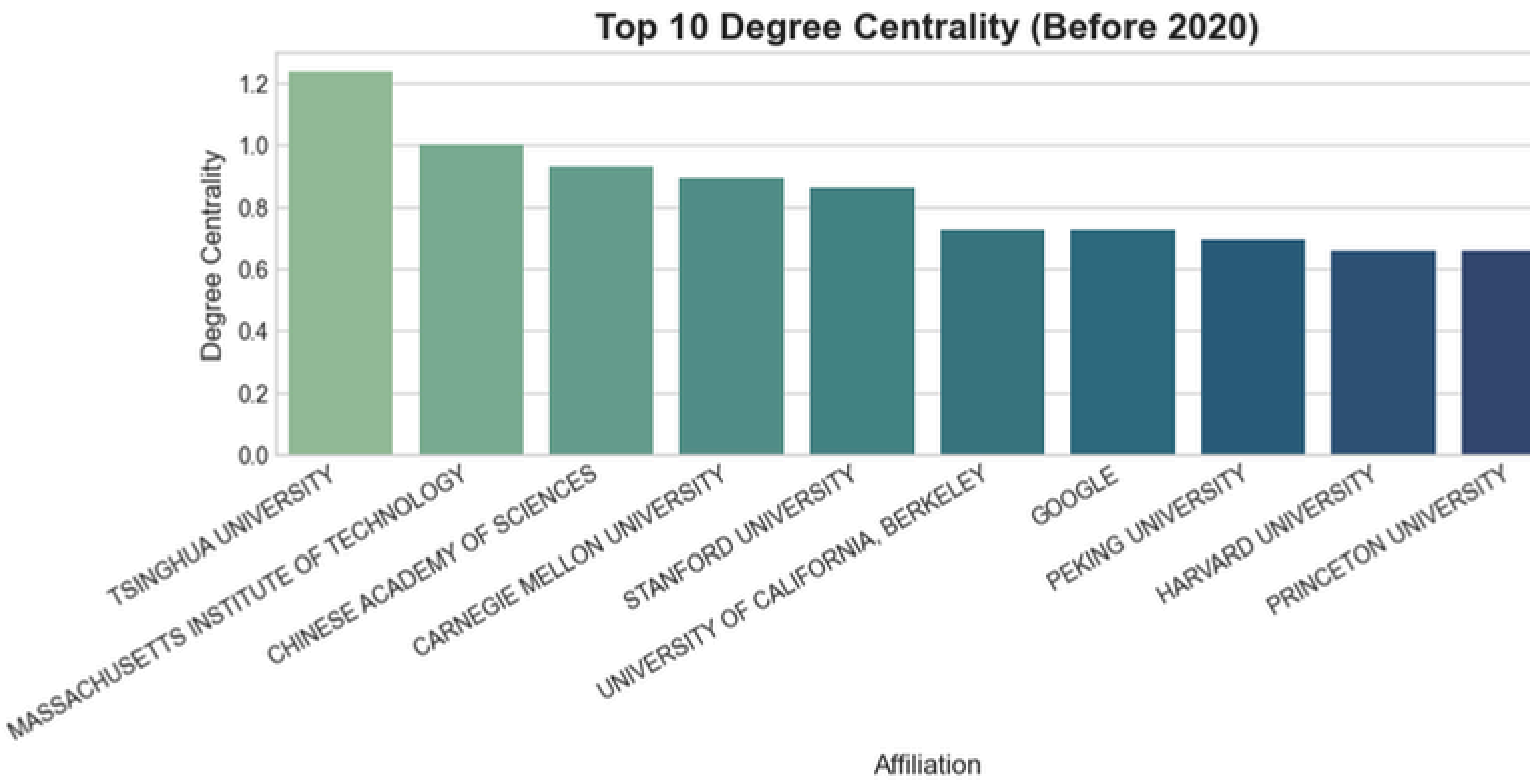
- 🇨🇳 **Tsinghua University stands out as the citation leader in China**
  → Frequently cited by both domestic and international peers.
- 🇺🇸 **MIT, CMU, and Stanford are highly cited within the US and abroad**
  → Consistently central in the international citation network.

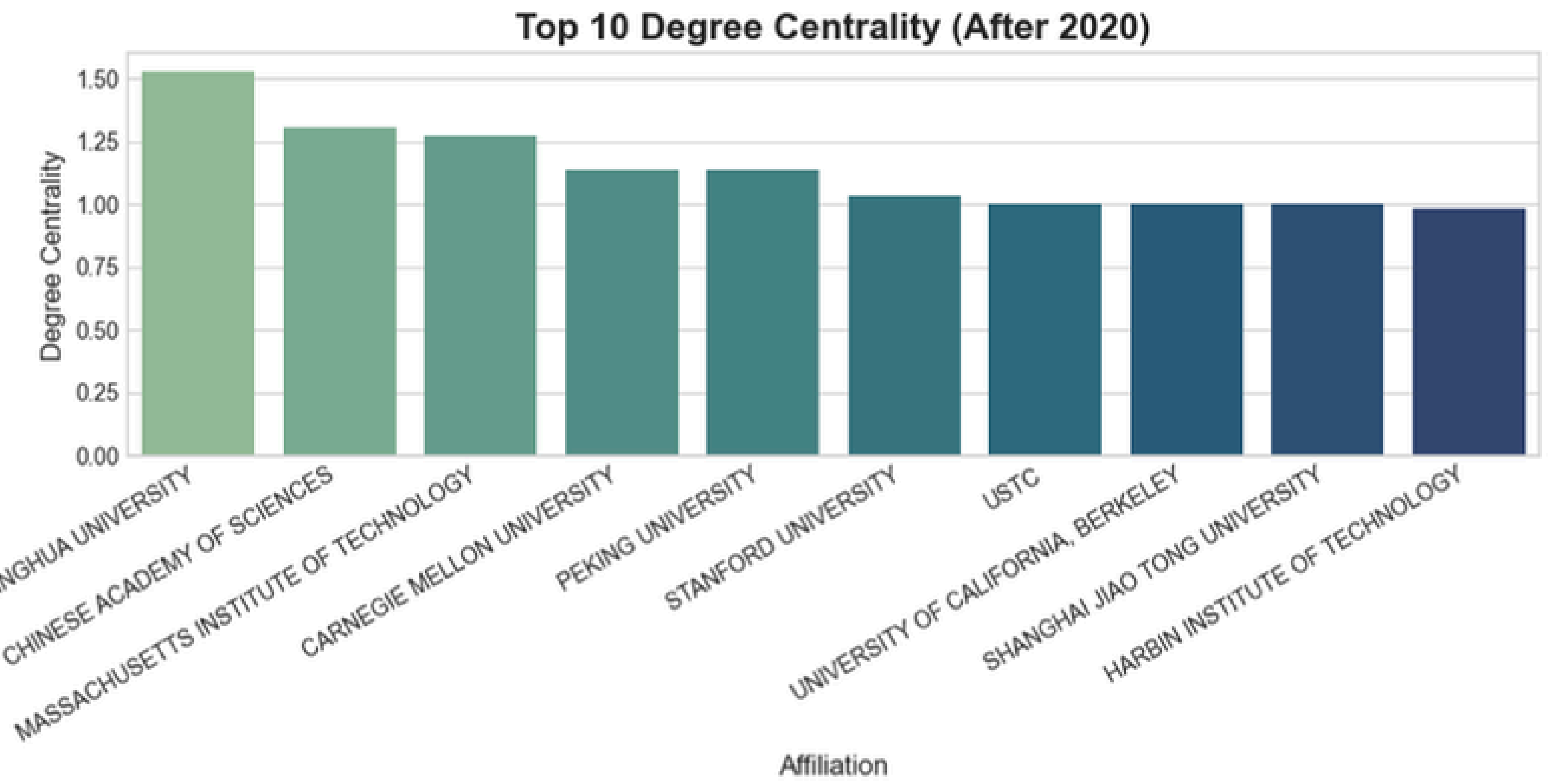




Citation

Degree Centrality– Before vs After 2020





- **Rise of Chinese institutions after 2020**

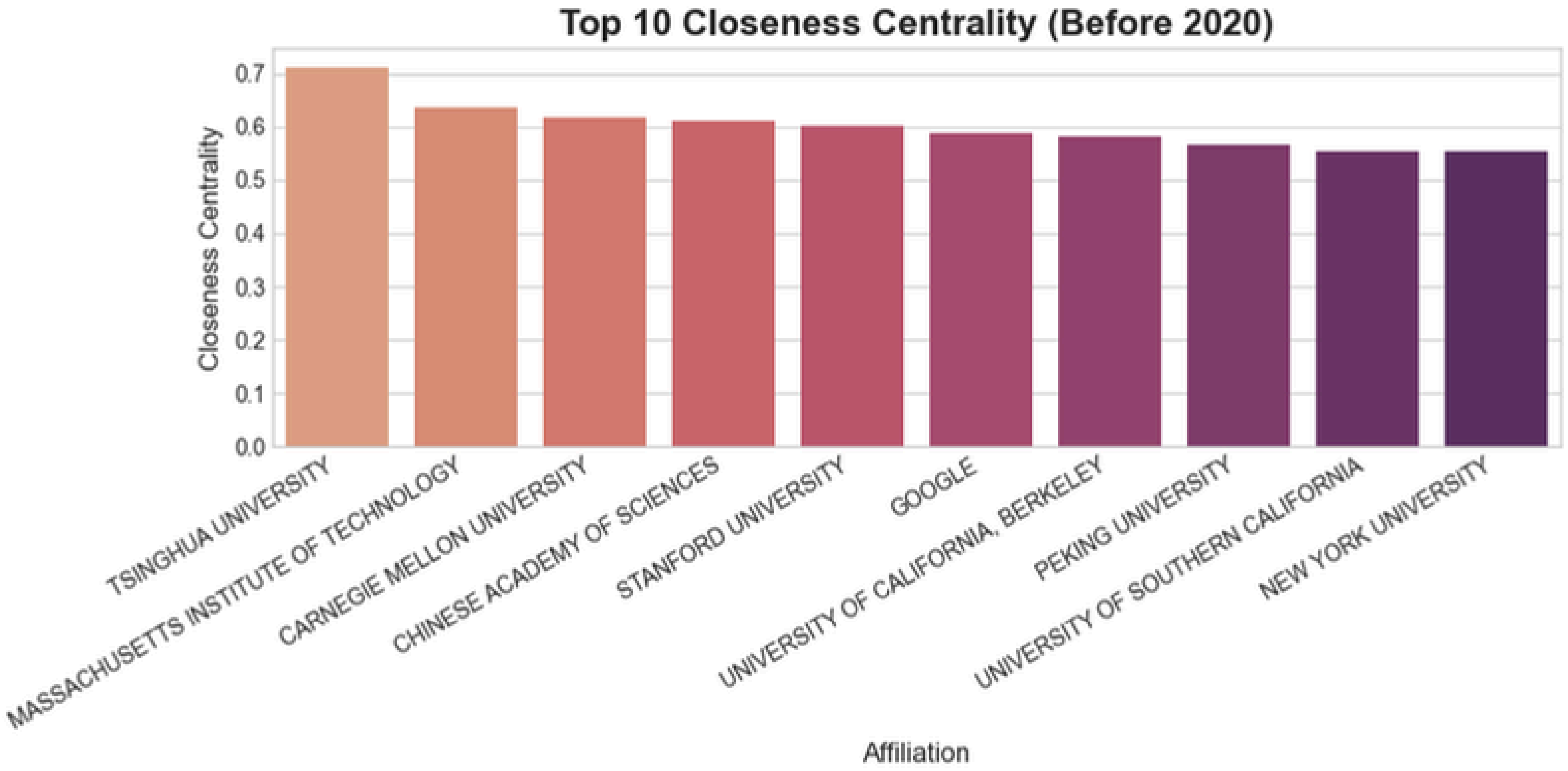
After 2020, 7 out of the top 10 institutions are based in China reflect broader and growing participation from Chinese universities in international networks.

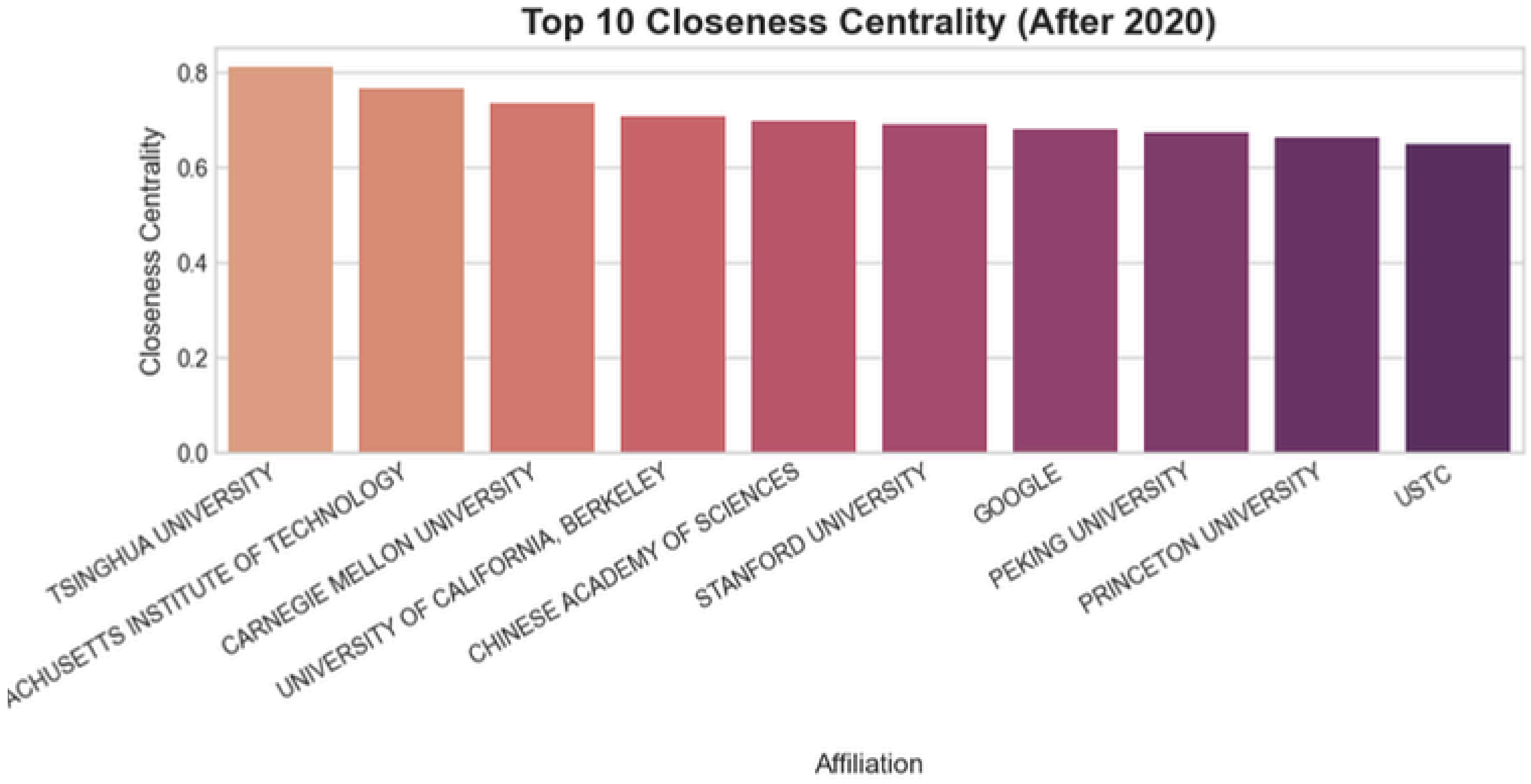
- **Overall increase in degree centrality values**

This points to increased collaboration overall, especially among leading institutions.

Citation

Closeness Centrality – Before vs After 2020





- **Higher scores after 2020**

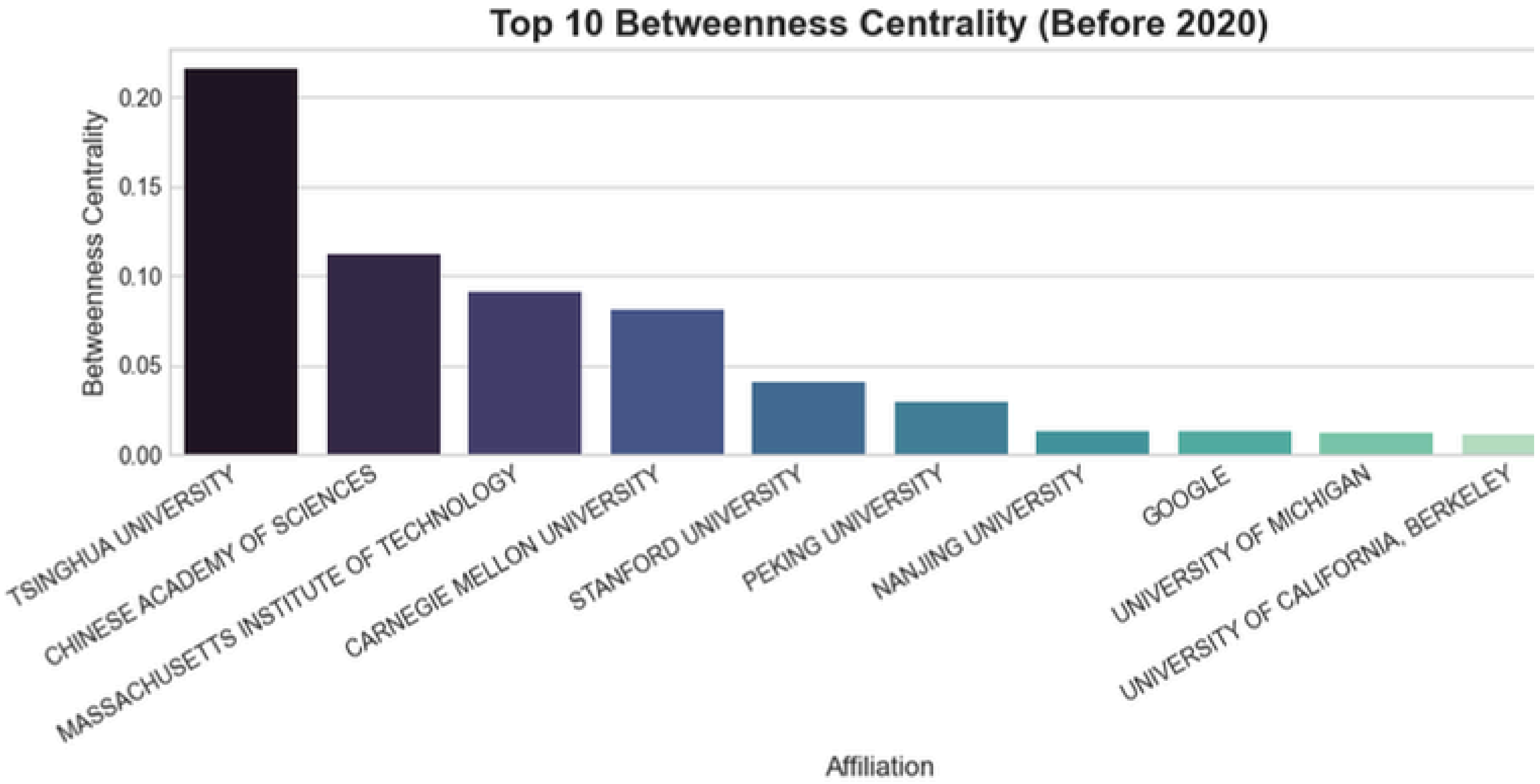
This means the network became more connected and information flows more quickly.

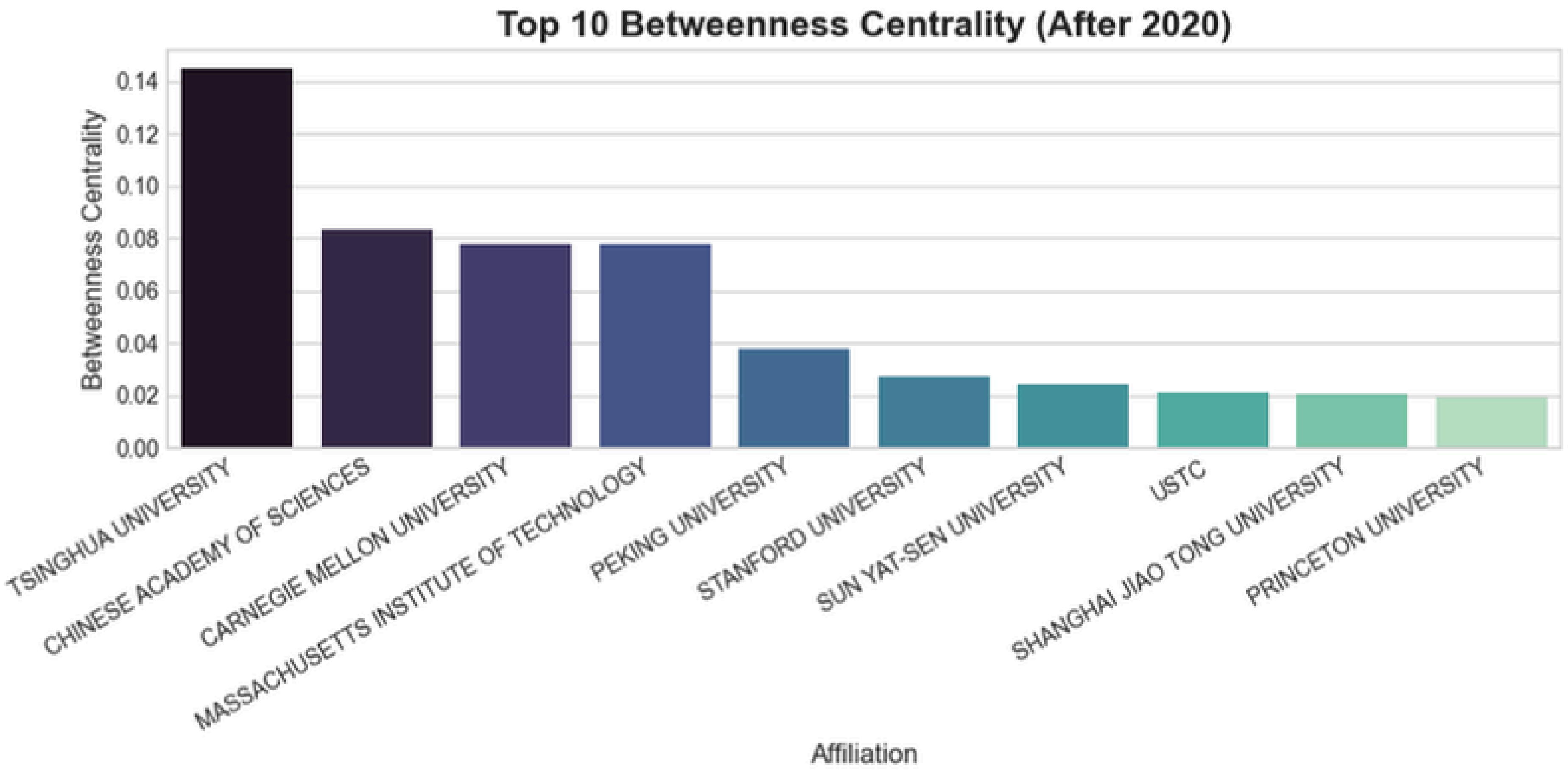
Google appears in both lists
Google kept a central position in the network.

Strong U.S. presence continues
MIT, Carnegie Mellon, Stanford, and UC Berkeley stayed in the top 10 in both periods.
They remain important and well-connected globally.

Citation

Betweenness Centrality – Before vs After 2020





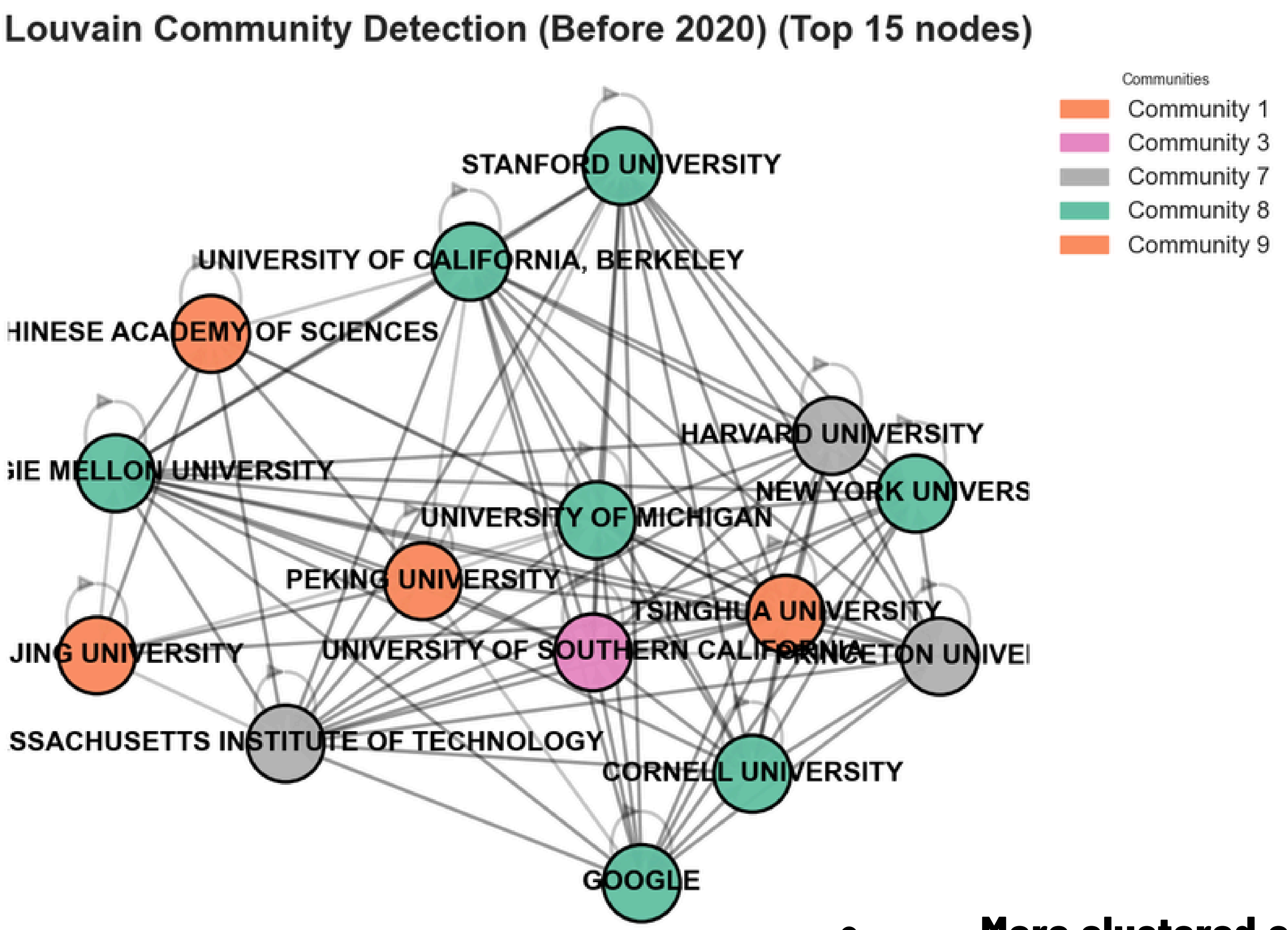
- **General decline in centralization**

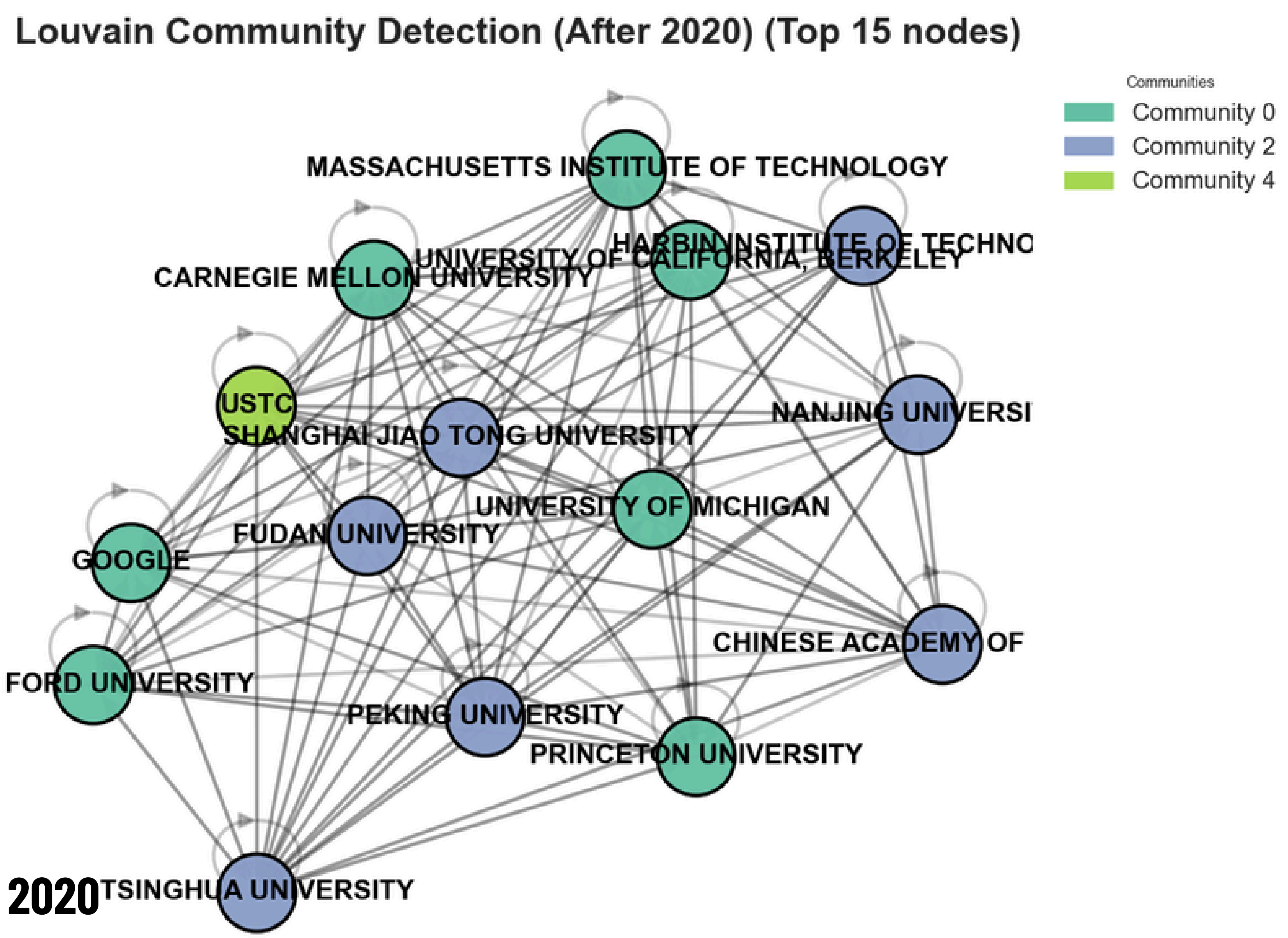
The top scores are lower after 2020 (max ~0.14 vs. ~0.22), suggesting a more distributed network with fewer “bridge” institutions dominating.

Tsinghua University still leads
Tsinghua ranks #1 in both periods, showing it plays the biggest role in connecting different parts of the collaboration network.

Chinese institutions increase in number
After 2020, 6 of the top 10 are from China (e.g. Tsinghua, CAS, Peking, USTC, SJTU, Sun Yat-sen).
This shows China has gained stronger control over network flow and connection across research communities.

Citation Louvain Community Detection – Before vs After 2020





- **More fragmented before 2020**

The network had 5 communities.

Chinese and US institutions were more mixed.

- **More clustered after 2020**

The network reduced to 3 communities, showing stronger internal connections.

Chinese institutions (Tsinghua, Fudan, SJTU, etc.) formed one tight community.

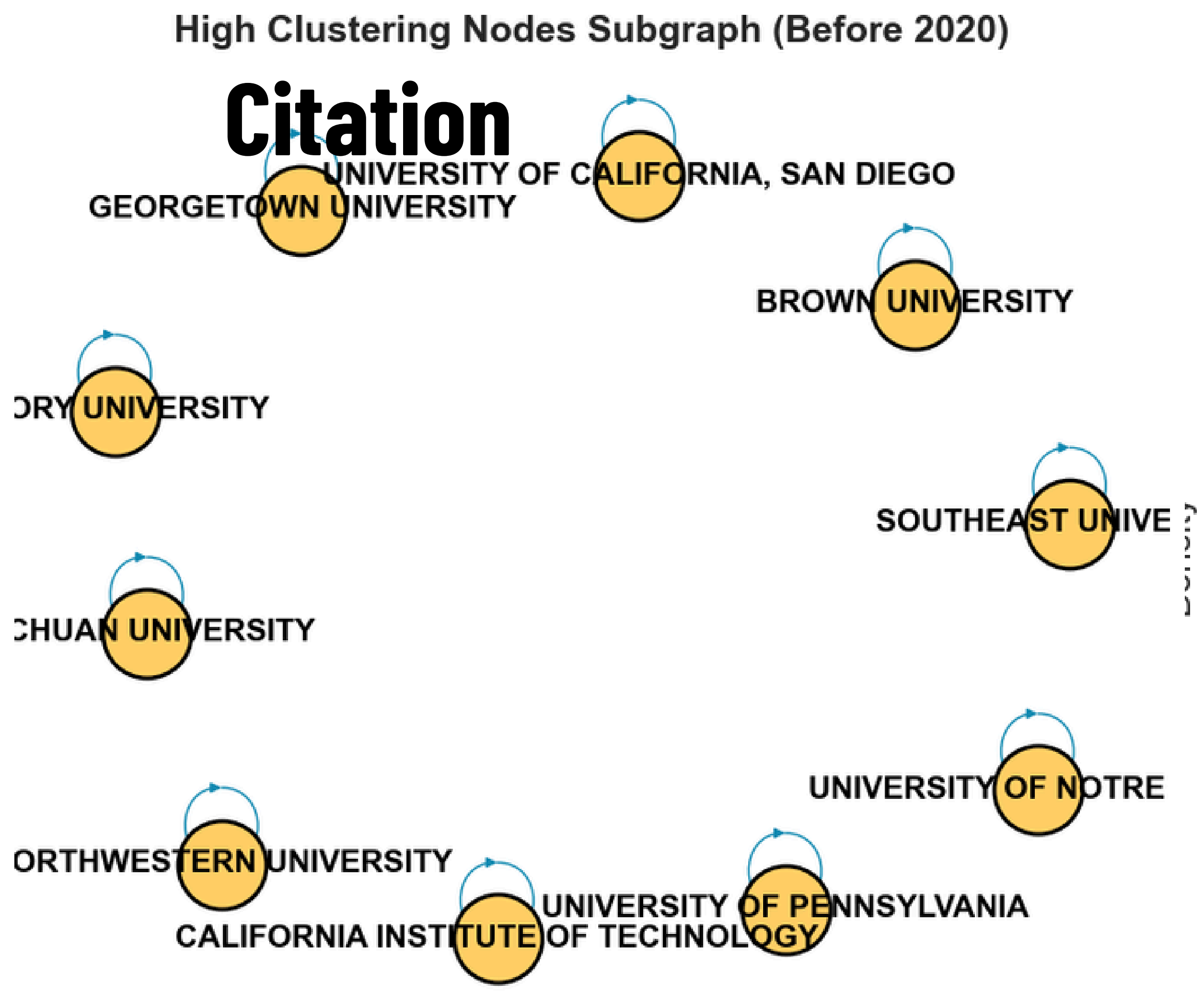
U.S. institutions (MIT, CMU, UC Berkeley) clustered together in another.

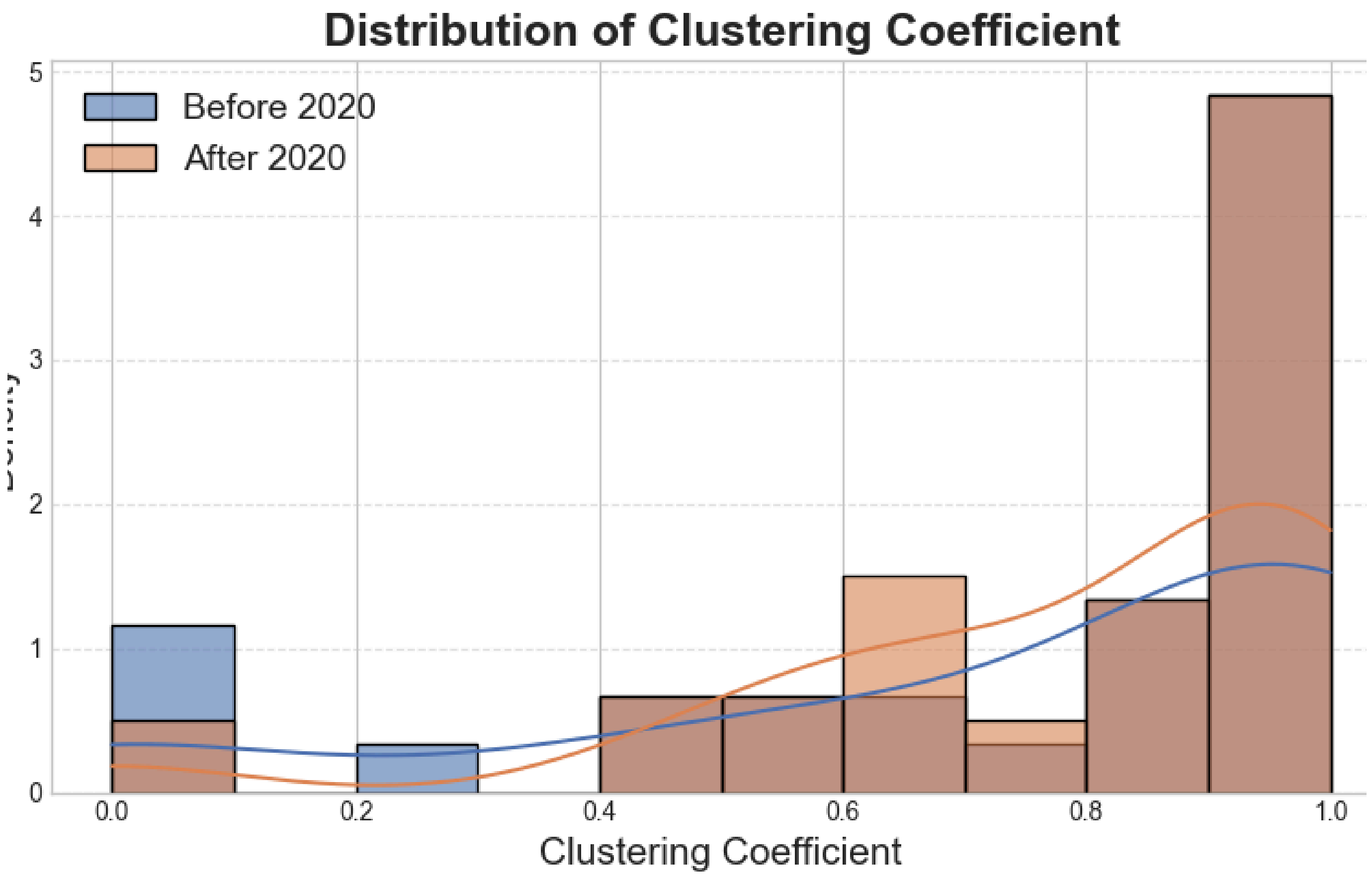
- **Stronger regional separation**

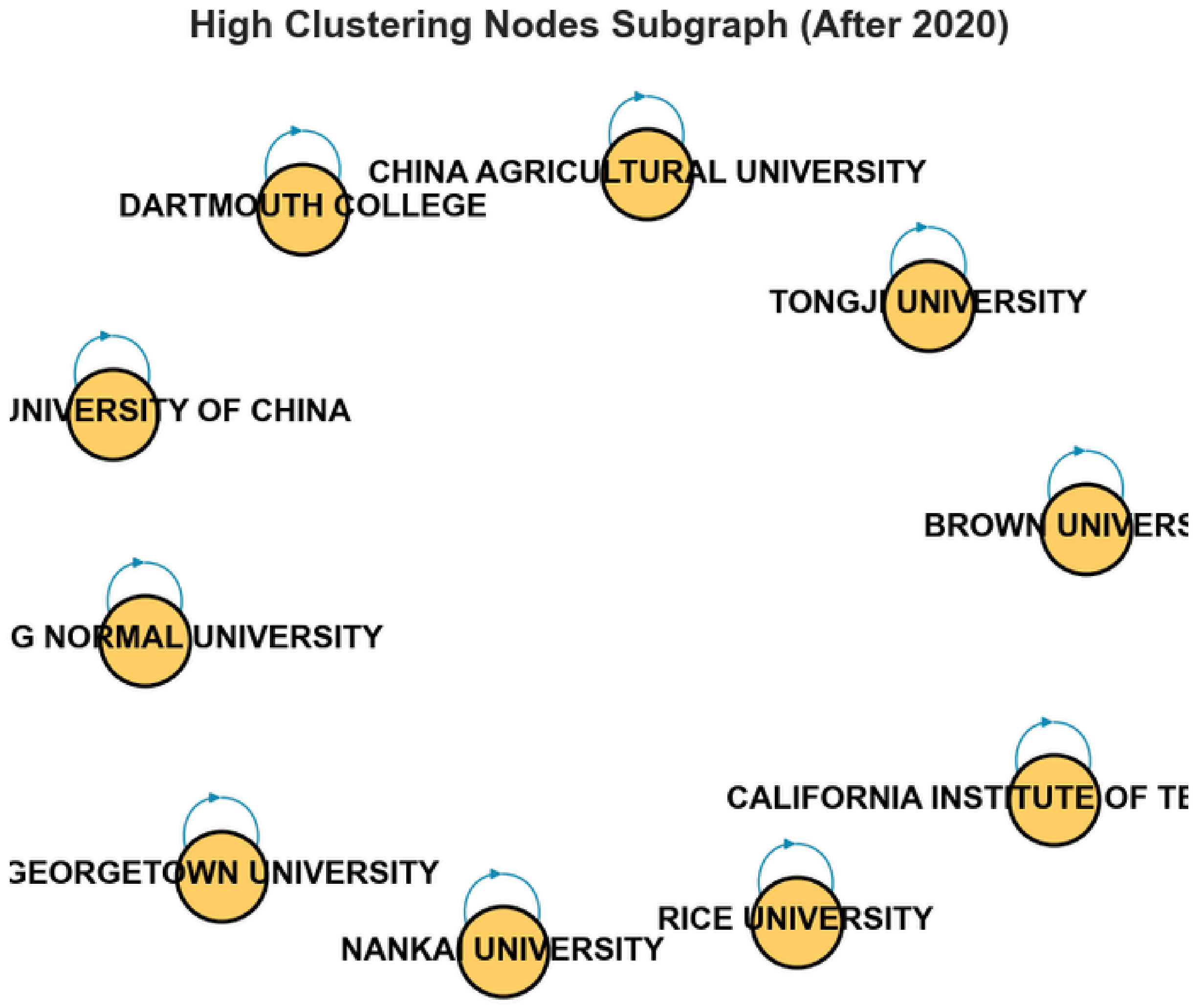
Collaboration is now more regionally concentrated.

There is less overlap between Chinese and U.S. institutions than before.

Citation







Clustering Coefficient – Distribution and High-Clustering Subgraphs

After 2020, more institutions have high clustering coefficients (close to 1.0).

This suggests stronger local collaboration: institutions tend to collaborate with tightly connected groups.

The distribution curve shifts right, indicating a more locally clustered academic network rather than broad, global networks.

3

About Interdisciplinary Network

- Our co-occurrence network is built by tracing which disciplines co-author with Computer Science, reflecting the flow of interdisciplinary research between institutions and countries.

- We explore collaboration patterns from two key perspectives:
  - 🌏 Static Distribution :
    How are CS–X collaborations concentrated across top-10 “hot” fields vs. bottom-5 “cold” fields in China and the US?
  - 🔄 Trend Analysis (2016–20 vs. 2021–25):
    Which fields are gaining or losing CS co-occurrence over time, and how do these shifts differ between China and the US?

Whats The Top 10 fileds?

01 Fileds

Computer science
Engineering
Medicine
Mathematics
Physics
Chemistry
Biology
Business
Economics
Psychology
Sociology
Environmental science
Geography
Geology
Materials science
Philosophy
History
Art
Political science

- More closely related would be some of the typical science and engineering disciplines

02 Hypothesis

Computer science
Physics
Mathematics
Engineering
Medicine
Economics
Business
Chemistry
Biology
Materials science

01 Fileds

Computer science
Engineering
Medicine
Mathematics
Physics
Chemistry
Biology
Business
Economics
Psychology
Sociology
Environmental science Geography
Geology
Materials science
Philosophy
History
Art
Political science

02 Hypothesis

Computer science
Physics
Mathematics
Engineering
Medicine
Chemistry
Biology
Materials science
Economics
Business

03 CN_TOP10

1. Computer science
2. Mathematics
3. Physics
4. Engineering
5. Chemistry
6. Psychology
7. Philosophy
8. Biology
9. Political science
10. Economics

04 US_TOP10

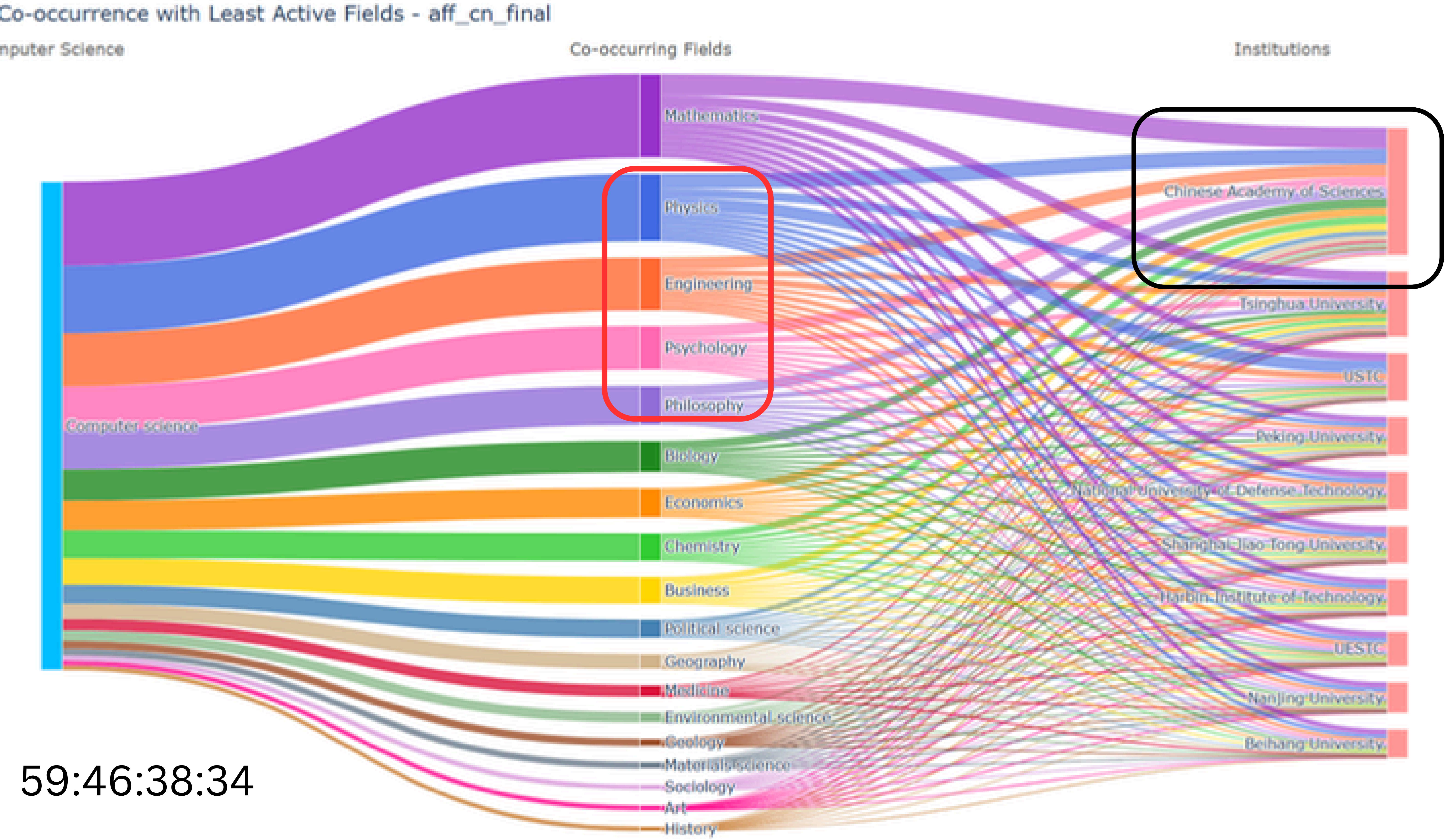
1. Computer science（41%）
2. Mathematics
3. Physics
4. Psychology
5. Philosophy
6. Biology
7. Engineering
8. Chemistry
9. Economics
10. Medicine

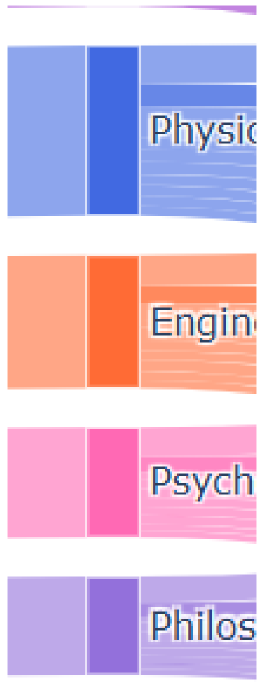
CN：40.1%-43.3%

US：40.0%-41.4%

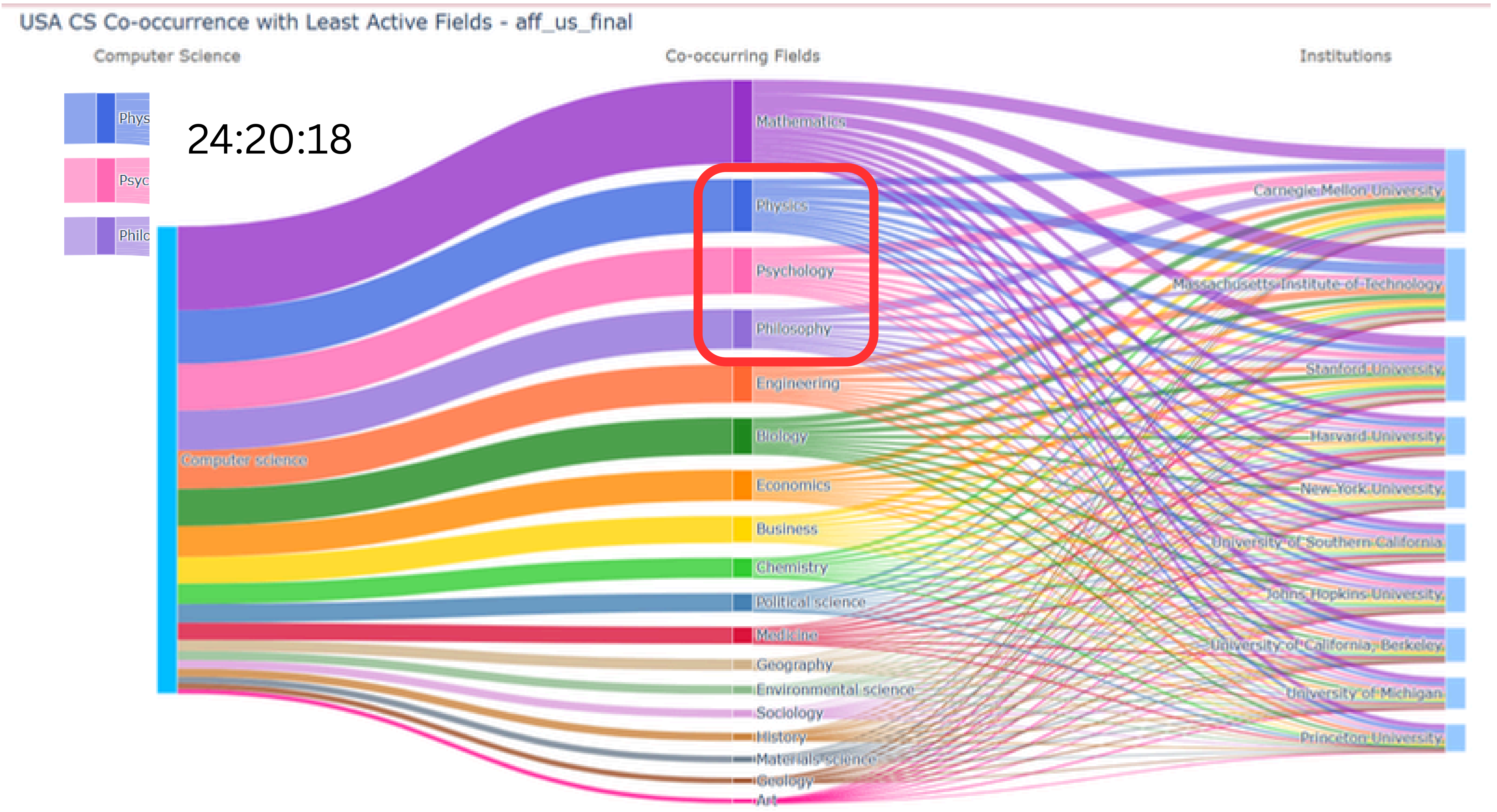
- The **psychology** and **philosophy** are far more closely connected to computer science than we ever imagined.
- Between 2016–2020 and 2021–2025, China’s CS co-occurrence rose from 40.1% to 43.3%, while the US’s increased from 40.0% to 41.4%, so both grew but China’s growth was faster.

- Stands out as the overwhelmingly dominant institution.

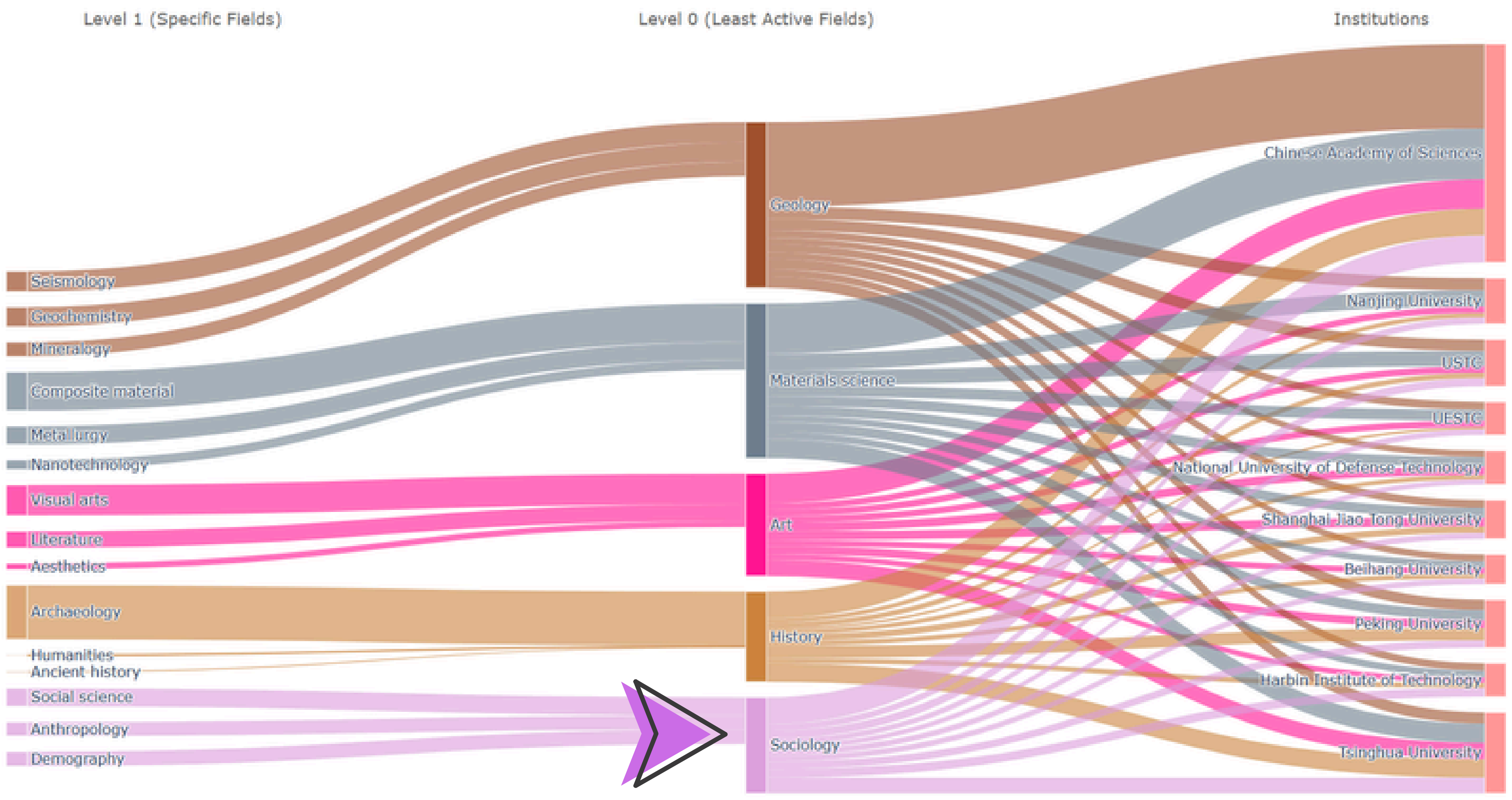


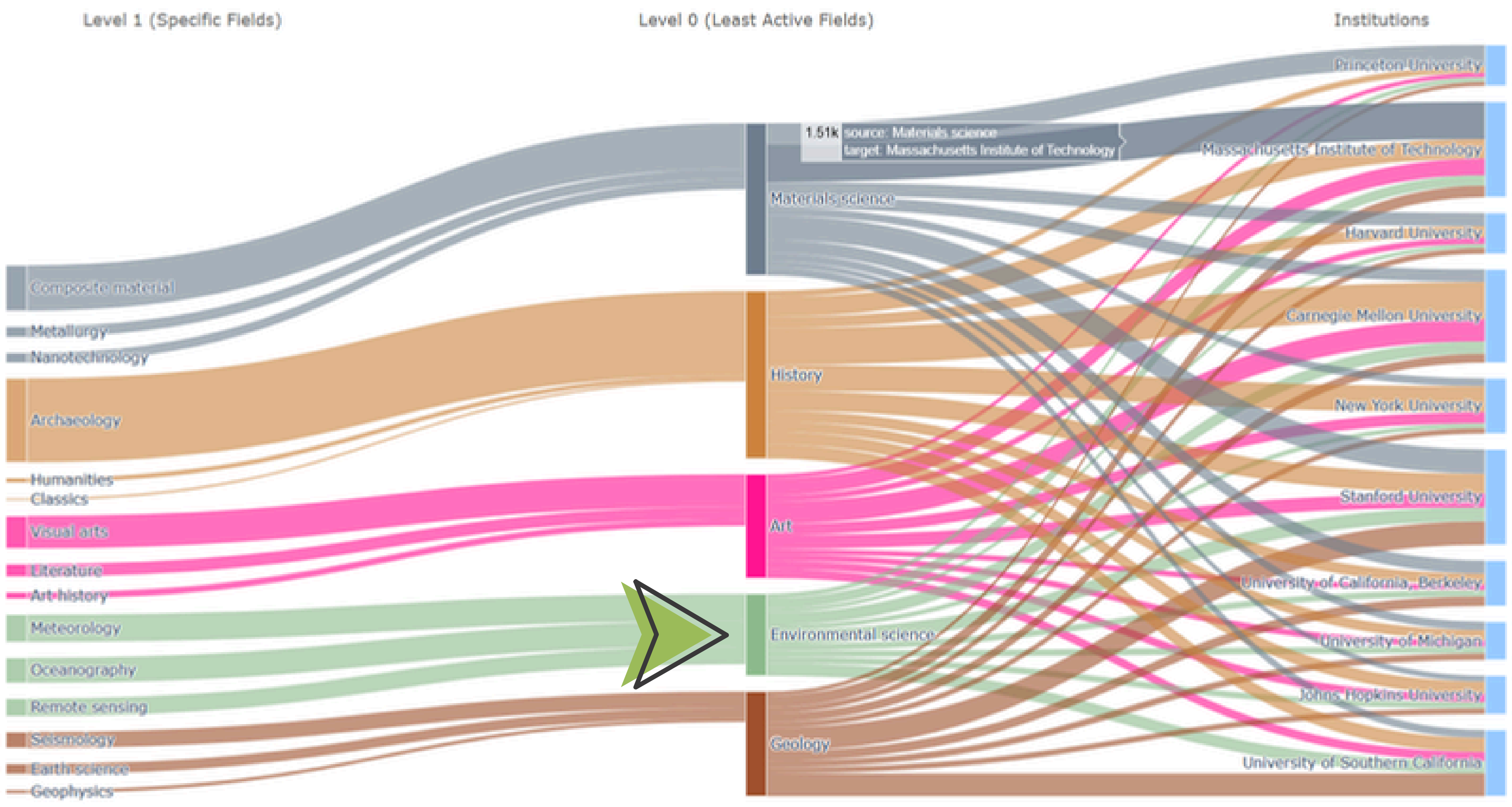


59:46:38:34


USA CS Co-occurrence with Least Active Fields - aff_us_final
Computer Science
Co-occurring Fields
Institutions
24:20:18
Phys
Psyc
Philo
Computer science
Mathematics
Physics
Psychology
Philosophy
Engineering
Biology
Economics
Business
Chemistry
Political science
Medicine
Geography
Environmental science
Sociology
History
Materials science
Geology
Art
Carnegie Mellon University
Massachusetts Institute of Technology
Stanford University
Harvard University
New York University
University of Southern California
Johns Hopkins University
University of California, Berkeley
University of Michigan
Princeton University

Some Blue Ocean Opportunities



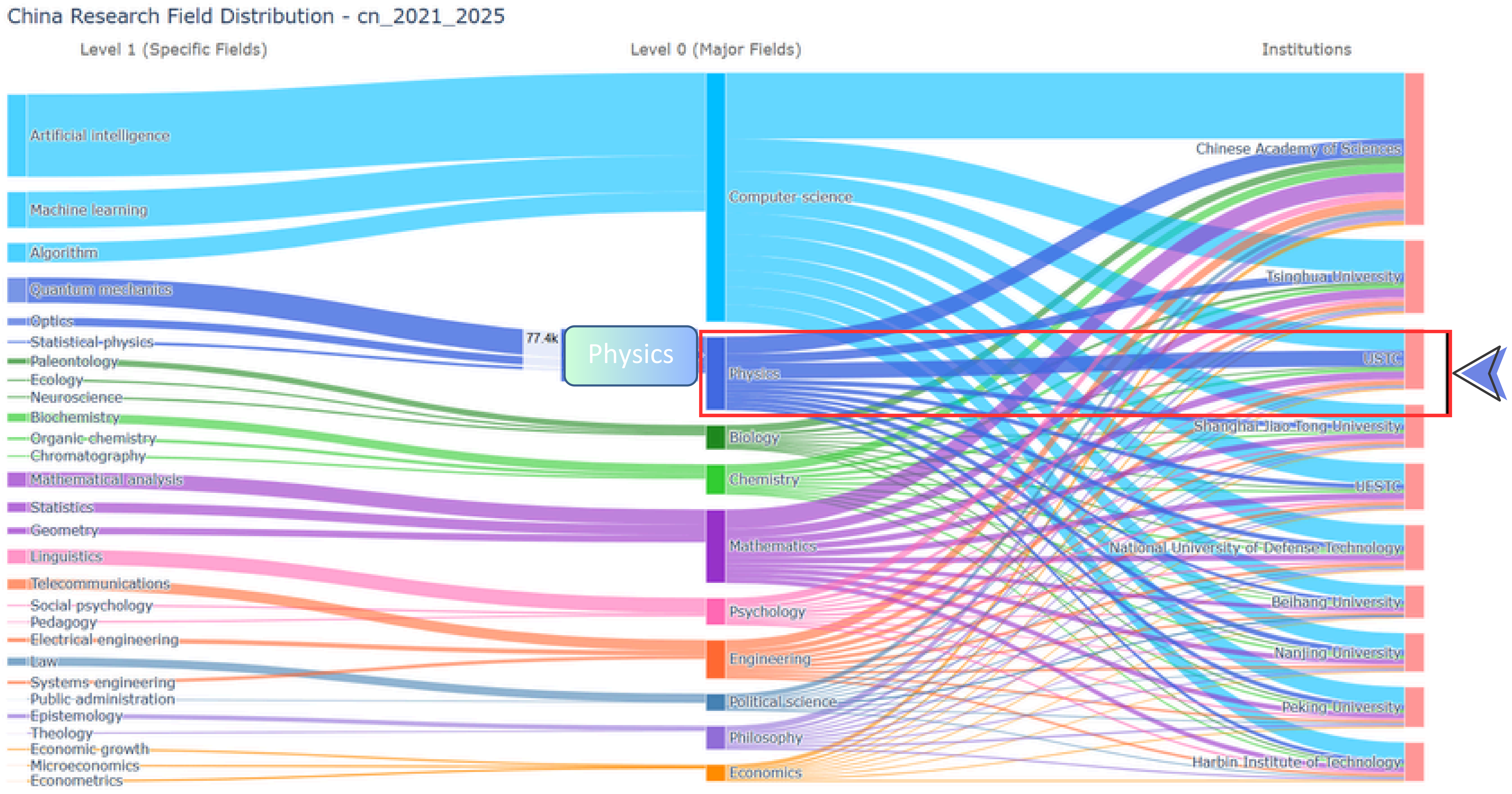


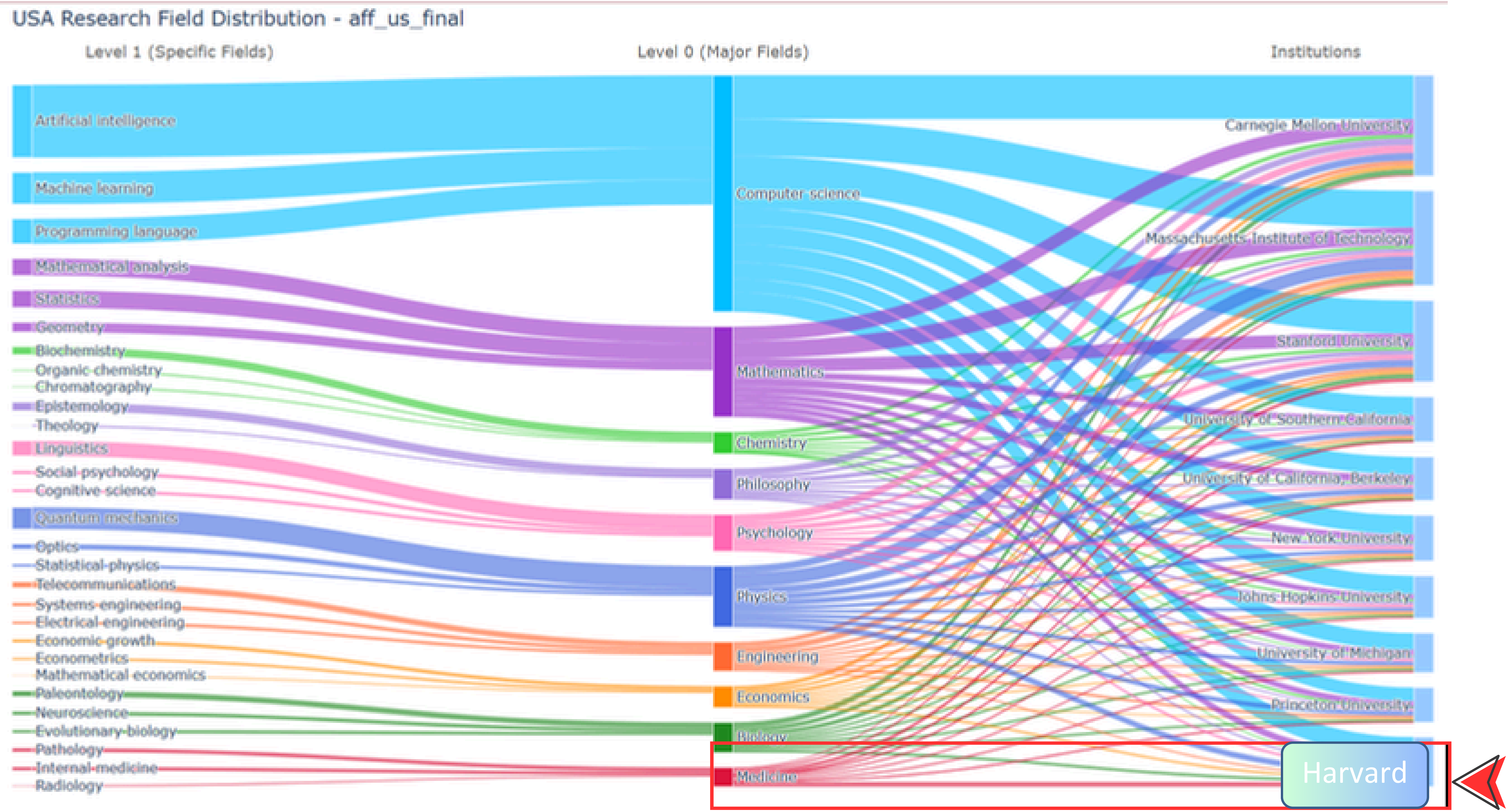
Common “Weak Fields”:

In both China and the US, “Geology, Materials Science, Art, and History” are the fields with the least CS interaction.

China’s Unique “Sociology” Gap vs. US’s Unique “Environmental Science” Gap

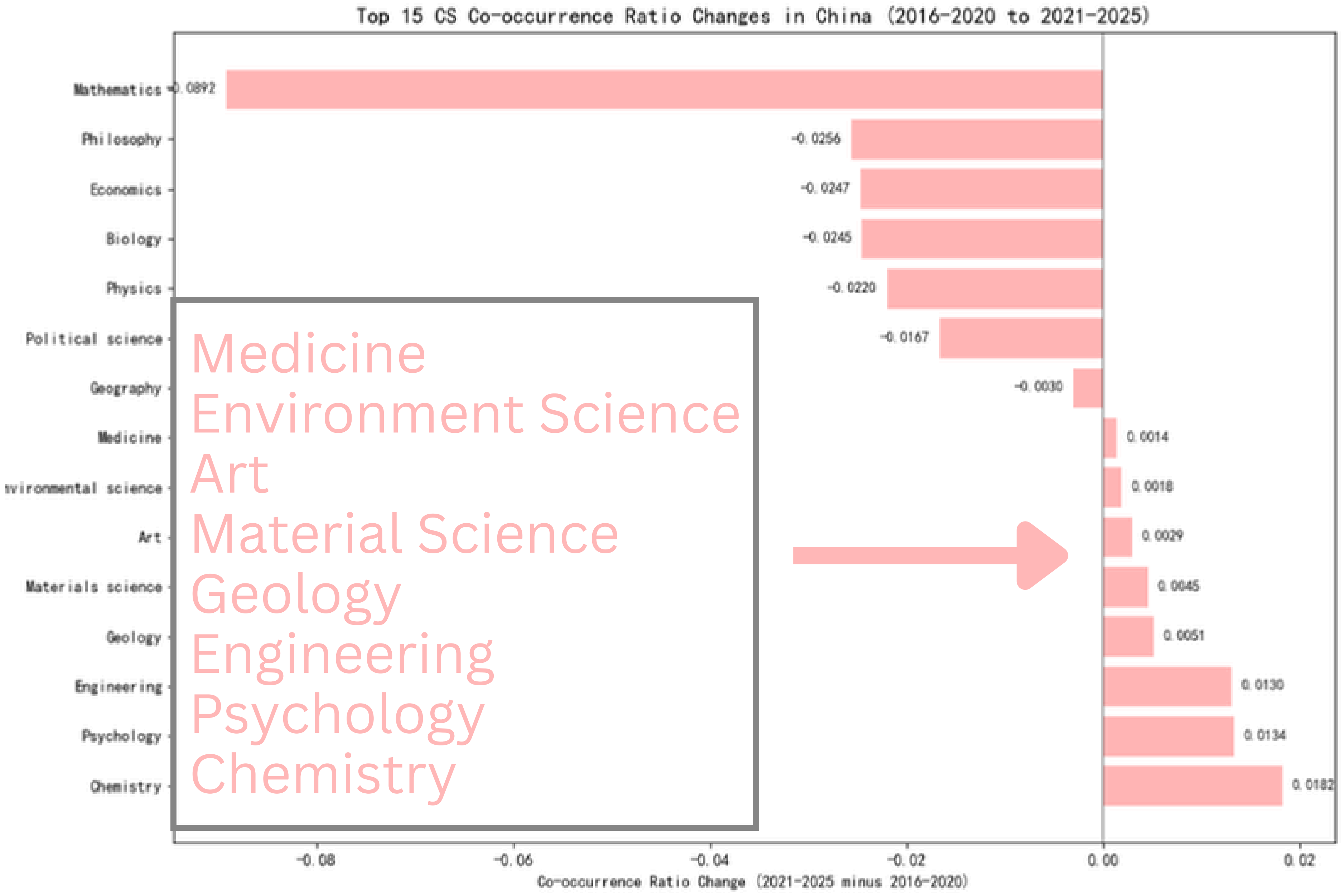
In China, CS-Sociology is almost non-existent, while in the US, CS-Environmental Science is the clear gap.

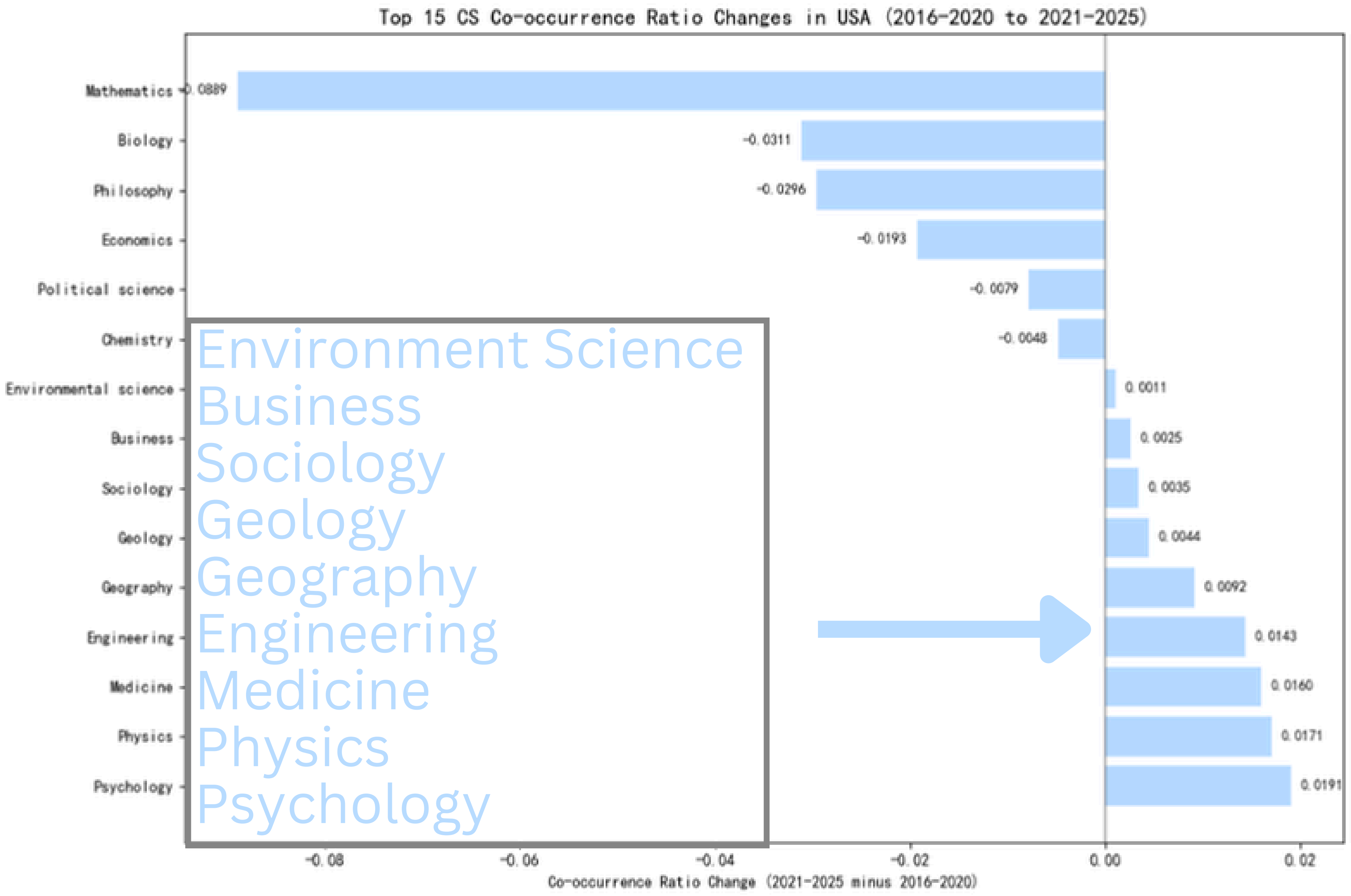

China Research Field Distribution - cn_2021_2025
Level 1 (Specific Fields)
Level 0 (Major Fields)
Institutions
Artificial intelligence
Machine learning
Algorithm
Quantum mechanics
Optics
Statistical physics
Paleontology
Ecology
Neuroscience
Biochemistry
Organic chemistry
Chromatography
Mathematical analysis
Statistics
Geometry
Linguistics
Telecommunications
Social psychology
Pedagogy
Electrical engineering
Law
Systems engineering
Public administration
Epistemology
Theology
Economic growth
Microeconomics
Econometrics
77.4k
Physics
Computer science
Physics
Biology
Chemistry
Mathematics
Psychology
Engineering
Political science
Philosophy
Economics
Chinese Academy of Sciences
Tsinghua University
USTC
Shanghai Jiao Tong University
UESTC
National University of Defense Technology
Beihang University
Nanjing University
Peking University
Harbin Institute of Technology



USA Research Field Distribution - aff_us_final
Level 1 (Specific Fields)
Level 0 (Major Fields)
Institutions
Artificial intelligence
Machine learning
Programming language
Mathematical analysis
Statistics
Geometry
Biochemistry
Organic chemistry
Chromatography
Epistemology
Theology
Linguistics
Social psychology
Cognitive science
Quantum mechanics
Optics
Statistical physics
Telecommunications
Systems engineering
Electrical engineering
Economic growth
Econometrics
Mathematical economics
Paleontology
Neuroscience
Evolutionary biology
Pathology
Internal medicine
Radiology
Computer science
Mathematics
Chemistry
Philosophy
Psychology
Physics
Engineering
Economics
Biology
Medicine
Carnegie Mellon University
Massachusetts Institute of Technology
Stanford University
University of Southern California
University of California, Berkeley
New York University
Johns Hopkins University
University of Michigan
Princeton University
Harvard

Co-occurance Ratio Changes CN : US





In both China and the United States, the links between computer science and traditional neighboring fields have weakened, while new technologies and applications are driving the diffusion of CS into previously distant disciplines and fostering interdisciplinary innovation.

Conclusion

01

Collaboration network

- **Deep collaborations between top-tier US and Chinese universities are declining, while domestic collaboration within each country is intensifying.**
- **After 2020, the co-authorship network becomes denser but more decentralized:**
  → Top 50 scholars are more dispersed
  → More isolated or weakly connected nodes emerge at the edges

- **Geopolitical tensions & national security concerns**
  → Heightened barriers to international collaboration.
- **Strategic shift toward self-reliance**
  → Both countries promote domestic innovation ecosystems.
- **Academic mobility restrictions post-COVID**
  → Limited international exchange programs, visiting scholar opportunities.

Citation Network

02

- **Before 2020**: Citations mainly flowed **from China to the US**.
- **After 2020**: Citation flows become **more mutual**, reflecting increased academic exchange.
- **The growing citation of Chinese papers signals rising research quality**
  → Chinese scholars are gaining **greater international recognition**.
- **However, top producers of high-impact publications remain concentrated in the US**

- **Rising research quality and visibility in China**
  → More Chinese papers published in top journals and indexed globally.
- **The US retains structural advantages**
  → Top journals, global networks, and strong research infrastructure keep the US at the center of high-impact publishing.
- **Faster knowledge diffusion via AI tools**
  → GPT accelerates literature discovery, summarization.

03

Interdisciplinary network

- **The links between CS and traditional neighboring fields have weakened, and links with previously distant disciplines have strengthened.**
- **After 2020:**
  →Total US CS-X collaborations remained roughly constant.
  →China's CS-X collaborations grew by 1.5×

Discipline Distribution

- US: CS collaborations are evenly spread across major associated fields.
- China: CS–Mathematics collaborations far outpace CS links with other disciplines.

- **Institutional strengths drive CS collaboration**
  → Strong fields attract CS resources and joint research.
- **Key Growth Drivers**
  → Likely driven by Chinese policy incentives and the rise of large language models.

Limitation

Algorithmic Bias

These concept names (display_name) and levels (level) are not directly annotated by article authors or human experts. Instead, the OpenAlex team generated them from the original Microsoft Academic Graph (MAG) corpus using an automated classifier—that predicts and assigns labels based on paper titles, abstracts, and publication venue names.
These concept tags and level assignments carry the risk of algorithmic bias, which may introduce systematic distortions in subsequent co-occurrence network construction and analysis.

Challenges in Causal Attribution

While this study highlights shifts in collaboration—like co-authorship density and geographic patterns—it cannot link these changes to specific factors (e.g., policy shifts or COVID-19) without additional data and causal analysis. Therefore, our findings are descriptive, not explanatory.

Data Coverage Limitations

This study's focus on top 30 institutions in China and the US excludes key regions like Europe and omits collaborations involving smaller but influential universities or scholars. Future work should include a wider range of countries and institutions for a more complete view of global academic collaboration.

Discussion

Related Works

[1] Aaron Clauset et al. ,Systematic inequality and hierarchy in faculty hiring networks.Sci.Adv.1,e1400005(2015).DOI:10.1126/sciadv.1400005

[2] Wapman, K.H., Zhang, S., Clauset, A. et al. Quantifying hierarchy and dynamics in US faculty hiring and retention. Nature 610, 120–127 (2022). https://doi.org/10.1038/s41586-022-05222-x

[3] Roberts, H., Cowls, J., Morley, J., Taddeo, M., Wang, V., Floridi, L. (2021). The Chinese Approach to Artificial Intelligence: An Analysis of Policy, Ethics, and Regulation. In: Floridi, L. (eds) Ethics, Governance, and Policies in Artificial Intelligence. Philosophical Studies Series, vol 144. Springer, Cham. https://doi.org/10.1007/978-3-030-81907-1_5

[4] Hine, E., Floridi, L. Artificial intelligence with American values and Chinese characteristics: a comparative analysis of American and Chinese governmental AI policies. AI & Soc 39, 257–278 (2024). https://doi.org/10.1007/s00146-022-01499-8

[5] Knox, J. (2020). Artificial intelligence and education in China. Learning, Media and Technology, 45(3), 298–311. https://doi.org/10.1080/17439884.2020.1754236

[6] Newman, M. E. J. (2001). Scientific collaboration networks. I. Network construction and fundamental results. Physical Review E, 64(1), 016131.

[7] Barabási, A. L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., & Vicsek, T. (2002). Evolution of the social network of scientific collaborations. Physica A, 311(3-4), 590-614.

[8] Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv preprint arXiv:2205.01833.

[9] Abbasi, A., Hossain, L., & Leydesdorff, L. (2012). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. Scientometrics, 94(2), 673-692.

[10] Abbonato, D., et al. (2024). Interdisciplinary research in artificial intelligence: Lessons from COVID-19. Quantitative Science Studies, 5(4), 922-935.

[11] Gomez, C., et al. (2025). Human-AI collaboration is not very collaborative yet: a taxonomy of interaction patterns. Frontiers in Computer Science, 6.

[12] Jia, C., Wang, D., & Xu, S. (2024). The China Initiative and its impact on US-China scientific collaboration. Proceedings of the National Academy of Sciences, 121(15), e2313646121.

[13] PFonseca, B. D. P. F. E., Sampaio, R. B., Fonseca, M. V. D. A., & Zicker, F. (2016). Co-authorship network analysis in health research: Method and potential use. Health Research Policy and Systems, 14(1), 34.

Thank You For Watching