A Century of Science: Globalization of Scientific Collaborations, Citations, and Innovations

Yuxiao Dong, Hao Ma, Zhihong Shen, Kuansan Wang
yuxdong|haoma|zhihosh|kuansanw@microsoft.com

Microsoft Research

A century of science

- **[12 years]** The volume of scientific publications doubled every 12 years between 1900 & 2015.
- **[3× author list size]** The size of a publication's author list tripled over the past 116 years, suggesting an increasingly collaborative scientific production process.
- **[4× collaborative innovations]** Science has benefited from the shift from individual work to collaborative effort, with over 90% of the world-leading innovations generated by collaborations in this century, nearly four times higher than they were in the 1900s.
- **[3× less self references]** Modern scientists instead tend to look for literature further back and farther around, rather than the frequent myopic and self-referencing 1900s.
- **[25 × & 7× global collaborations & citations]** Scientific development has globalized over time, including 25-fold and 7-fold increases in international collaborations and citations, respectively.
- **[2× diversified innovations]** Innovations has been more globally diversified, including a dramatic decline in the dominant accumulation of citations by the US, the UK, and Germany, from ~95% to ~50% over the years between 1900 and 2015.

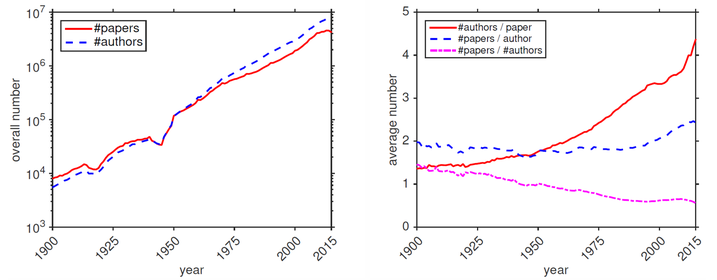
Microsoft Academic Graph (MAG)

A subset of the MAG in summer 2016
- #papers: 89 million
- #authors: 53 million
- #collaboration-links: 1.2 billion
- #citation-links: 795 million
- #years: 116 (1900 -- 2015)

The geographic related data statistics
- #papers: 21 million
- #citation-links: 269 million

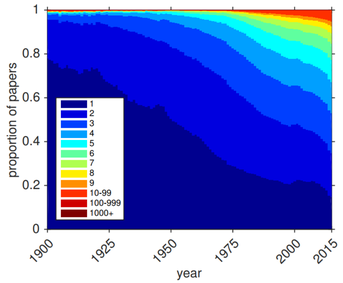
The data is publicly available at Microsoft Academic Knowledge API
https://azure.microsoft.com/en-us/services/cognitive-services/academic-knowledge/

The growth of science



- The volume of scientific publications doubled every 12 years between 1900 & 2015, despite each scientist's individual productivity remaining surprisingly constant.

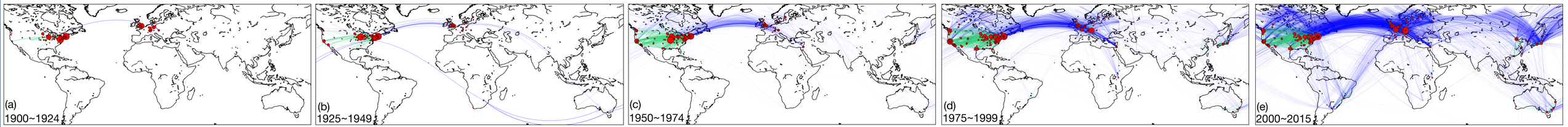
Science became more collaborative



- The share of single-author publications gradually but substantially shank over centuries.

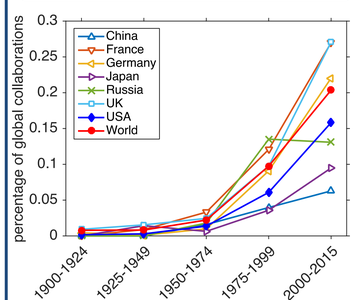
The planetary-scale view of science between 1900 and 2015

Blue and green lines represent the relative collaboration strength between institutions from different countries and from the same country, respectively. The red circles represent the top 200 most-cited institutions in the world.



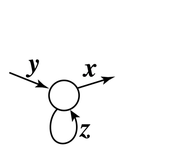
- Only 4% top 200 most cited institutes were located outside of the US, UK, and Germany.
- Collaborations did not break geographical boundary.
- Top institutes in Germany, the UK, and Europe at large drastically shrank during WW II.
- Institutes in the US West Coast started to emerge.
- Top institutions started to emerge in Israel, Japan, Australia, and North Europe.
- Institutes in the US West Coast continued to emerge.
- UK & Continental Europe had recovered from WW II
- Institutes in the US West Coast had been important in science.
- The rise of scientific impact in China, South Korea, & Singapore in Asia and Brazil in South America, as well as many places in Europe.

Globalization of collaborations

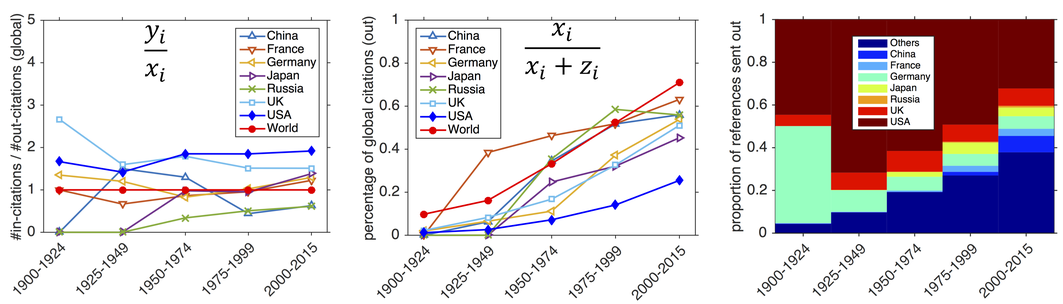


- Global collaborations increased 25 times over the past 116 years

Globalization of citations

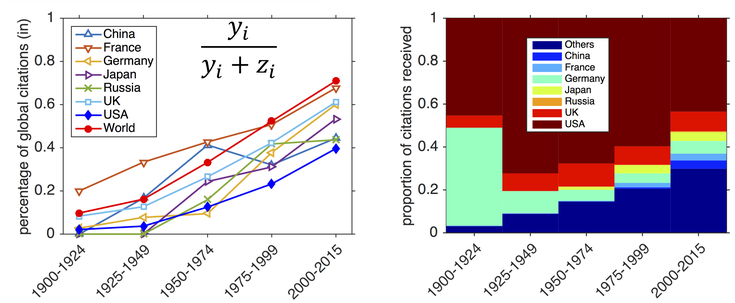


- x: #citations sent
- y: #citations received
- z: #self-citations sent
- z: #self-citations received



- The U.S. and UK consistently received as twice as many citations they cited others.
- Only 10% of citations were sent out to other countries in the early 20th century, while this share has dramatically increased to 70% in the 21st century.

Globalization of innovations



- The U.S., UK, & Germany collected ~95% of the world's citations during the early 20th century, while this share was decreased by about half to 50%.

References

1. A. Sinha et al. An Overview of Microsoft Academic Service (MAS) and Applications. *WWW 2015 Companion*.
2. J. Tang et al. ArnetMiner: Extraction and Mining of Academic Social Networks. *KDD 2008.*
3. R. Sinatra, P. Deville, M. Szell, D. Wang, A-L Barabasi. A Century of Physics. *Nature Physics 2015*.
4. B. F. Jones et al. Multi-University Research Teams: Shifting impact, geography, and Stratification in Science. *Science 2008*.
5. R. Pan et al. World citation & collaboration networks: uncovering the role of geography in science. *Scientific Reports 2012.*

Paper info: (Poster Slot #10)

KDD2017

Yuxiao Dong, Hao Ma, Zhihong Shen, Kuansan Wang.
A Century of Science: Globalization of Scientific Collaborations, Citations, and Innovations.
In *Proc. of 23rd ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (**KDD'17**).*

Microsoft Academic

Research more, search less

Open & big academic data