Mapping academic migration of German-affiliated researchers across countries using 8 million Scopus publications from 1996 to 2020

Xinyi Zhao [a]*, Samin Aref [a], Emilio Zagheni [a], Guy Stecklov [b]

[a] *Laboratory of Digital and Computational Demography, Max Planck Institute for Demographic Research, Germany,*
  Xinyi Zhao - zhao@demogr.mpg.de, Samin Aref - aref@demogr.mpg.de, Emilio Zagheni - zagheni@demogr.mpg.de
[b] *Department of Sociology, University of British Columbia, Canada, Guy Stecklov - guy.stecklov@ubc.ca*

* Corresponding author

Keywords: Academic migration, Density equalizing cartograms, Bibliometric data, Germany

Abstract:

Academic mobility plays a predominant role in the context of increasing the internationalization of science and research production, thus creating a global market of qualified professionals (Franzoni et al., 2015, Schiller and Cordes, 2016, Netz and Jaksztat, 2017). Germany has launched a series of programs that have the aim to attract researchers with high citation performance and link German-affiliated scientists abroad with each other (Schiller and Cordes, 2016). Thus, mapping the German academic in- and out-flow patterns across countries and disciplines can help clarify the position and attraction of Germany in the global science system. Taking all fields of scholarship, this analysis speaks directly to policy development. Using large-scale bibliometric data from over 8 million Scopus publications during the period 1996-2020, we analyze and visualize the international migration to and from Germany among published researchers by taking into account their citation information and semantically identified disciplines. This reveals the spatial hotspots of migration and in- and out-flows for all fields of scholarship.

The large-scale bibliometric data in our study includes an exhaustive set of over one million researchers who have published with an affiliation address in Germany at some point during the 1996-2020 period in Scopus-indexed publication venues. The raw bibliometric data from over 8 million Scopus publications are firstly pre-processed to infer missing country information using a neural network algorithm, and to disambiguate authors using an unsupervised machine learning algorithm. We assign the country of academic origin and the country of academic destination to each author according to their most frequent (mode) country(ies) during the first year of publishing and that of the most recent year of publishing. Accordingly, immigrant researchers to Germany (origin: not Germany; destination: Germany) and emigrant researchers from Germany (origin: Germany; destination: not Germany) can be detected. We illustrate the academic migration in- and out-flow pattern (figure 1, 2) using density equalizing cartograms (Hennig et al., 2012), which transforms the shape of map polygons proportional to the total migration inflows (outflow) of researchers moving to (from) Germany. The total citations for each researcher are determined by the total number of citations (as of April 2020) for all publications authored by a researcher. We further calculate the average annual citation rate for each researcher by dividing their total citations by their academic age (the number of years since first publication until 2020). The divisions of the average annual citation rates of mobile researchers to Germany and from Germany are represented by different colors in the two maps and explained in the legends.

A Latent Dirichlet allocation (LDA) model is used to determine the main discipline for each German-affiliated migrant researcher, by identifying the semantic topic of bibliometric information composed of the publication titles, journal titles, and keywords authored by a researcher. We further evaluate and visualize the migration preferences of mobile researchers across different disciplines, and use kernel density estimation to detect the respective hotspots of the origins of migration inflows to Germany and the destinations for migration outflows at different scales.

In the next steps of this work, we intend to gain further insights into the role of migration both into and out of Germany in each field and discipline from a micro perspective. Methodologically, our computational approach of processing, analyzing, and visualizing large-scale bibliometric data can be adopted as a framework of analysis for studying scholarly migration in other countries.
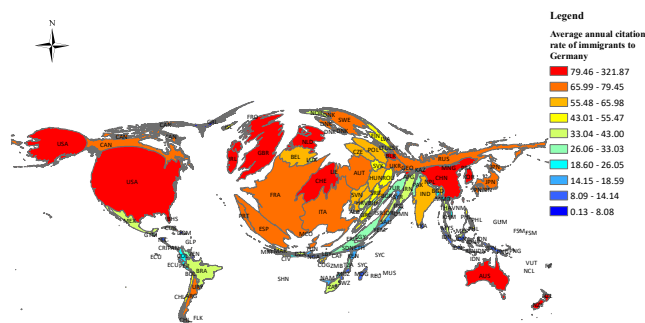
Acknowledgements

Figure 1. Migration inflows to Germany (proportional to the area) and average annual citation rates of immigrant researchers to Germany.
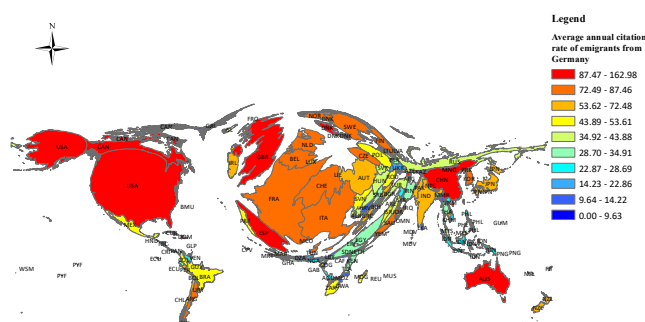

Figure 2. Migration outflows from Germany (proportional to the area) and average annual citation rates of emigrant researchers from Germany.

References

Franzoni, C., Scellato, G. and Stephan, P., 2015. Chapter 2 - international mobility of research scientists: Lessons from GlobSci. In: A. Geuna (ed.), *Global Mobility of Research Scientists*, Academic Press, pp. 35–65.

Hennig, B. D., Pritchard, J., Ramsden, M. and Dorling, D., 2012. Remapping the world's population: Visualizing data using cartograms. In: *Fair Play: A Daniel Dorling reader on social justice*, Policy Press.

Netz, N. and Jaksztat, S., 2017. Explaining scientists' plans for international mobility from a life course perspective. *Research in Higher Education* 58(5), pp. 497–519.

Schiller, D. and Cordes, A., 2016. Measuring researcher mobility. In: *OECD Blue Sky Forum*, p. 24.