

SciSciPeople - Identify new people for SciSci

Yifan Qian (Northwestern University)
Sirag Erkol (Northwestern University)
Henry Xu (Carnegie Mellon University)

In this work, we introduce a pipeline – SciSciPeople – for identifying people, both producers and consumers, who are not yet a part of the science of science community, but exhibit potential interests in the science of science. In order to find the new people, we leverage data from ICSSI website, SciSciNet, and Google Scholar.

First, we use the contributors of ICSSI as a proxy to identify the current community of the science of science. We parse the names of contributors from the website of ICSSI, and create a network from these names, in which nodes represent contributors, and edges are drawn between nodes if the names of two people appear in the same talk or poster presentation (see Figure 1).

After identifying the current community of the science of science, we use SciSciNet and Google Scholar to identify people potentially interested in the science of science. First, from SciSciNet, we gather people who have cited the science of science review papers by Fortunato *et al.* [1] and Azoulay *et al.* [2], then compare this list of citers to the list of contributors, and identify all people who have cited these two review papers, but did not have a contribution at ICSSI. Furthermore, we parse the list of people who have selected science of science as a research interest in their Google Scholar profile, and find the ones who have not had a contribution to ICSSI 2023.

As a result, we identify 707 people who have cited the science of science review papers but did not have a contribution to ICSSI. Similarly, we identify 117 people from the Google Scholar dataset. Interestingly, the third most represented country in this list is Ukraine. The list of these people can be leveraged and used together with different ranking metrics to identify potential new people for the science of science community.

To summarize, we are embarking on an essential initial phase in the development of a pipeline designed to identify new individuals for SciSci. While our current focus lies primarily on identifying potential producers, it is equally crucial to consider identifying potential consumers as a future direction. The pipeline we propose has the potential to be applicable to various fields that aim to discover new people to grow the field.

[1] Fortunato, Santo, Carl T. Bergstrom, Katy Börner, James A. Evans, Dirk Helbing, Staša Milojević, Alexander M. Petersen et al. "Science of science." *Science* 359, no. 6379 (2018): eaao0185.

[2] Azoulay, Pierre, Joshua Graff-Zivin, Brian Uzzi, Dashun Wang, Heidi Williams, James A. Evans, Ginger Zhe Jin et al. "Toward a more scientific science." Science 361, no. 6408 (2018): 1194-1197.

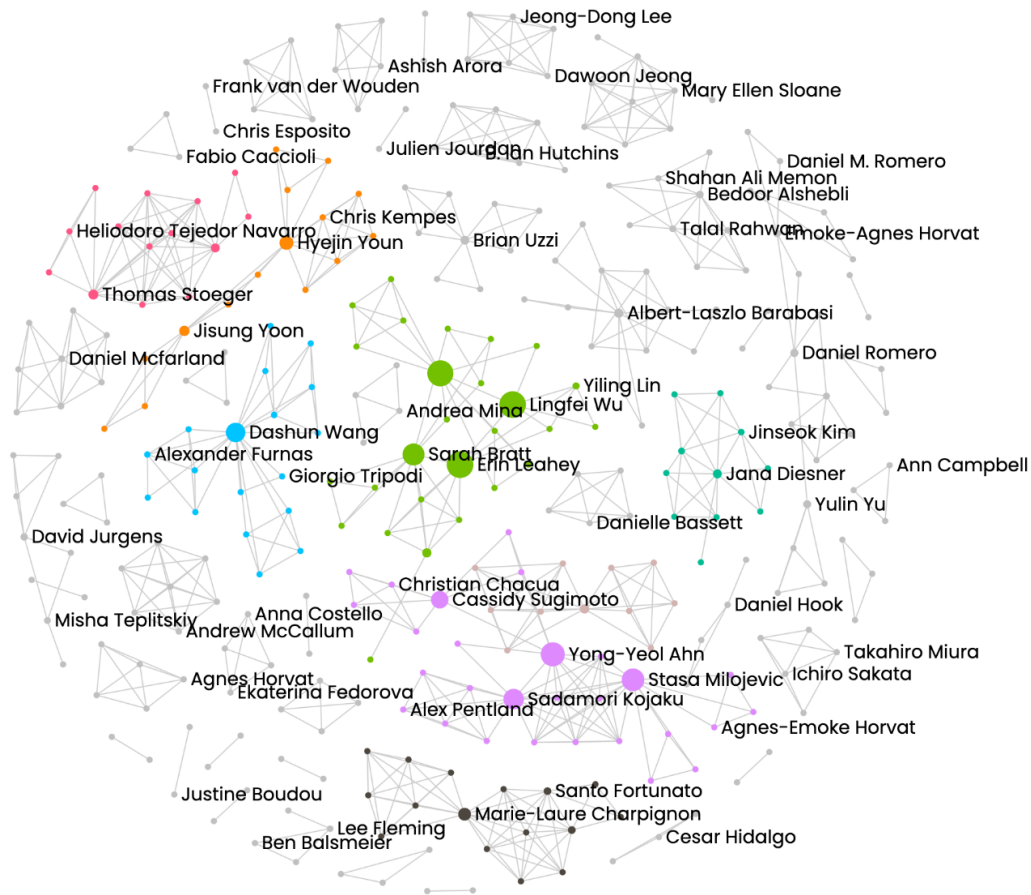


Figure 1: A network representation of ICSSI 2023 contributors. Nodes represent the contributors, while edges connect two nodes if they have contributed to the same talk or poster. Nodes are colored based on their assigned communities using the Louvain method. Node sizes are proportional to their betweenness centrality values.

Appendix:

Presentation slides:

<https://docs.google.com/presentation/d/1KtTGQ4L3vKpU8UEEQXdSownraaEGHvv93guXPZsMHWY/edit?usp=sharing>