

## Scientific birds of a feather flock together: science communication on social media rarely happens across or beyond disciplinary boundaries



*The success of academic research in reaching out beyond its own scientific community is a perennial concern, even more so following the rapid adoption of social media and the ability to easily transmit information to potentially millions of people. Consequently, many attempts have been made to capture the broad scientific impact beyond academia using social media data. But is increased social media attention really indicative of “broader impact”?* **Qing Ke, Yong-Yeol**

**Ahn and Cassidy R. Sugimoto** have studied how much scientific discourse is happening across and beyond scientific communities on Twitter and found that social media does not broaden scientific communication, but rather replicates and perpetuates pre-established disciplinary boundaries. “Alt” metrics may not be so alternative after all.

Is science trapped in the ivory tower? Are scientists locked in their silos? How scientific knowledge reaches diverse groups beyond its own scientific community is an enduring question, one that is now positioned in a new context because of the rapid adoption of social media. As social media replaces traditional communication channels, it provides a completely new medium via which diverse groups can directly talk to each other, where a single message can potentially reach millions of people within an hour, and provides scientists with revolutionary ways to make detailed quantitative observations on communication at a global scale. As a result, many attempts, often collectively called “altmetrics,” have been made to capture the broad scientific impact beyond academia using social media data. The metrics have been heralded to measure societal impact of research and to complement traditional citation measures for research evaluation.

Progress on this topic has been hampered by a lack of information about the producers of scientific discourse on social media and their networks. For instance, what if all social media sharing of a research paper were by automated bots? Or if all attention came from the author, journal, and publisher of the paper? Are these indicative of “broader impact”? Furthermore, perhaps all the attention is from scientists, but within the same domain: if all tweets about a paper on underwater basket weaving were from a tight clique of underwater basket weaving researchers, is this representative of the broader impact of science?

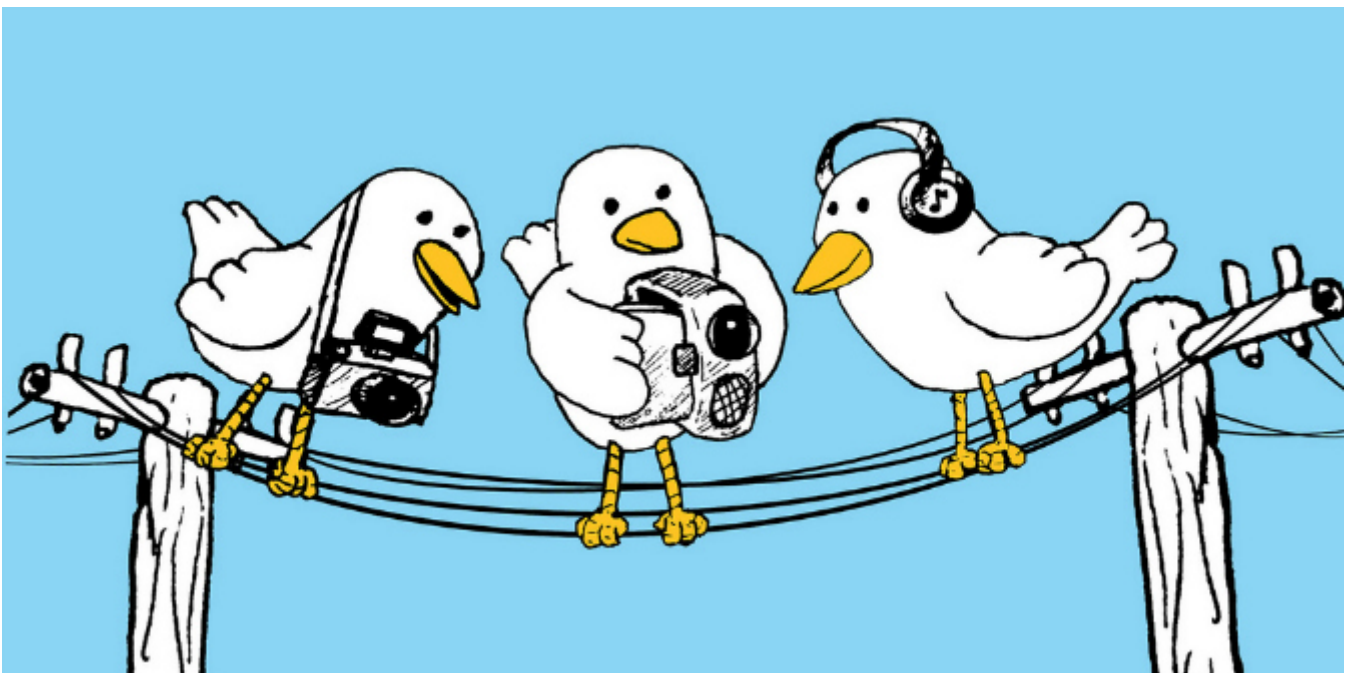


Image credit: **Multimedia Birds of a Feather** by James Nash. This work is licensed under a [CC BY-SA 2.0](https://creativecommons.org/licenses/by-sa/2.0/) license.

It is with these questions of broader impact that we began our research. We were curious to know how much scientific discourse is happening across and beyond scientific communities on social media. To do this, we started with a seemingly simple question: can we generate a list of scientists on social media? This is an inversion from previous research which began with a list of scientists (e.g. from bibliometric data) and then tried to find these individuals on social media. This previous approach led to a host of biases, such as prioritising those who were successful in other metrics (e.g. production and citation), issues with data accessibility, as well as technological complications (e.g. author name disambiguation). Our approach was anchored within the platform, leveraging the wisdom of the crowds in terms of Twitter lists. Our underlying rationale was that we can safely consider a user as a scientist if (1) other users consider this person a scientist and (2) the person identifies as a scientist in their profile.

We were faced with the Herculean task of creating a list of scientific titles. We took a liberal approach, merging the classification from the US Bureau of Labor's **Standard Occupational Classification** and scientific occupations in Wikipedia to prepare a list of "seed" scientists. Our final **list** of titles reveals interesting patterns about self-identification and specialisation of scientists on Twitter. First, our list identifies more practitioner-oriented disciplines than other disciplinary classifications. Secondly, our list demonstrated the role of specialisation in self-identification on Twitter: e.g. historians, by and large, identified as historians; chemists and biologists, on the other hand, identified with a large variety of specialised titles. This differentiation creates problems for identifying disciplinary populations of parallel scale; though this is not an uncommon problem for scientometric research.

Our seed list repeatedly matched the titles with **Twitter list** names and added newly discovered scientists. This process resolved in a sample of 45,867 scientists. Our method has been critiqued on the basis that certain disciplines may be **underrepresented**, using as evidence the comparatively large number of followers of scientific societies and journals. However, as has been demonstrated, a substantial proportion of scientific tweets are generated from bots, and organisational Twitter handles are likely to draw a large number of both bots and organisational followers. We therefore prioritised precision over recall – our objective was to create a replicable and systematic (rather than anecdotal) approach to identifying individuals who were likely to be scientists.

Given a sample of scientists, several questions can be answered that yield additional insight into the composition and behavior of the scientific community on the platform. What are the demographics of scientists on Twitter? What is the distribution across scientific disciplines? How is this population biased compared with the actual population? We automatically inferred gender of the scientists using first names and US Census data. The resulting data suggested that female scientists are overrepresented on Twitter relative to their representation in the scientific workforce. This may suggest greater avenues for participation in scientific discourse for women on this platform, though it would be necessary to control for age and other variables to fully understand this phenomenon. In terms of discipline, social and computer and information scientists are overrepresented, whereas life, physical, and mathematical scientists are underrepresented, compared with the US workforce. **As has been suggested**, it may be useful to replicate this method using other occupational classifications, to examine whether the results hold.

Some approaches to identifying scientists rely on the content of tweets. Therefore, using our verified list of scientists, we wanted to know the degree to which they tweeted about science and what other sources frequented their tweets. It turns out that scientists are people, too: the vast majority of what they share is the same as the general population. Social sites such as Instagram, Facebook, and YouTube, and major news sites such as *The Guardian*, *The New York Times*, and *The Huffington Post* are common sources. At the same time, it is clear that they share content relevant to their disciplines: the **arXiv preprint server** and the **American Physical Society website** are popular among physicists, the **Association for Computing Machinery website** among computer scientists, and the **London School of Economics and Political Science blogs** among social scientists.

This leads us to the final and perhaps most important question of our analysis. Do scientists form strong cliques based on their disciplines? We looked at how the scientists followed, retweeted, and mentioned each other. Our results showed high degrees of disciplinary assortativity—that is, scientific birds of a feather do indeed flock together. This has critical implications for the interpretation of social media metrics as metrics of broader or social impact. Our results suggest that social media does not broaden scientific communication, but rather replicates and perpetuates pre-established disciplinary boundaries. “Alt”-metrics may not be so alternative after all.

*This blog post is based on the authors’ article, “[A systematic identification and analysis of scientists on Twitter](#)”, published in PLoS ONE (DOI: 10.1371/journal.pone.0175368).*

*Note: This article gives the views of the authors, and not the position of the LSE Impact Blog, nor of the London School of Economics. Please review our [comments policy](#) if you have any concerns on posting a comment below.*

## About the authors



**Qing Ke** is a PhD Candidate at Indiana University School of Informatics and Computing, interested in complex systems.



**Yong-Yeol Ahn** is an assistant professor at [Indiana University School of Informatics and Computing](#) and a co-founder of [Janys Analytics](#). He develops and leverages mathematical and computational methods to study complex systems such as cells, the brain, society, and culture. His recent contribution includes [a new framework to identify pervasively overlapping modules in networks](#), [network-based algorithms to predict viral memes](#), and [a new computational approach to study food culture](#). He is a recipient of several awards including [Microsoft Research Faculty Fellowship](#). He worked as a postdoctoral research associate at the [Center for Complex Network Research at Northeastern University](#) and as a visiting researcher at the [Center for Cancer Systems Biology at Dana-Farber Cancer Institute](#) for three years after earning his PhD in Statistical Physics from [KAIST](#) in 2008..



**Cassidy R. Sugimoto** researches within the domain of scholarly communication and scientometrics. She examines the ways in which knowledge is produced, disseminated, and rewarded. She has presented her work at numerous conferences and has received research funding from the National Science Foundation, Institute for Museum and Library Services, and the Sloan Foundation, among other agencies. Sugimoto is actively involved in teaching and service and has been rewarded in these areas with an [Indiana University Trustees Teaching award \(2014\)](#) and a [national service award from the Association for Information Science and Technology \(2009\)](#). She is currently [President of the International Society for Scientometrics and Informetrics](#). Sugimoto has an undergraduate degree in music performance, an M.S. in library science, and a Ph.D. in information and library science from the [University of North Carolina at Chapel Hill](#).



Share This Story, Choose Your Platform!

This work by [LSE Impact of Social Sciences blog](#) is licensed under a [Creative Commons Attribution 3.0 Unported](#).

☺