How A New Company Uses Machine Learning To Measure Academic Impact

Michael T. Nietzel Senior Contributor ⓘ
I am a former university president who writes about higher education.

Follow

🔖   💬 0   Mar 14, 2024, 06:06am EDT

A new company is using machine learning to rank colleges and departments and measure scholarly ... [+] GETTY

Higher education is awash in data. Whether it's the indicators the college rankings industry uses to compare institutions, the metrics faculty and administrators employ to gauge the importance of scholarly work, or the outcomes selected to measure the value of a college education, quantification is the name of the game when evaluating much of academic work.

But how best to harness all that data? How should they be combined and analyzed to yield accurate and meaningful estimates of institutional quality, scholarly influence and social impact?

A relatively new company – AcademicInfluence.com, co-founded in 2020 by Jed Macosko, a Berkeley PhD, professor of physics at Wake Forest University and president of the fledgling startup — believes it has the answer with a novel use of machine learning to measure various aspects of higher education, including college and department rankings, scholarly reputations and publications with the greatest impact. (In full disclosure, I have written two invited articles for the company's website.)

Its approach is based on an algorithm that measures how often a person's work is mentioned in various databases. For example, how often is Richard Phillips Feynman, the Nobel-winning theoretical physicist, mentioned in Wikidata, the text associated with Wikipedia, and in the abstracts, titles, references, keywords in publications found at Crossref and Semantic Scholar. It then extrapolates from those numbers to arrive at cumulative measures of different kinds of impact.

College Rankings

One of Academic Influences' first products was a college ranking system using a methodology based on the premise that the people affiliated with an institution determine its quality.

Using machine-learning technology, originally developed with funding from the Defense Advanced Research Projects Agency, Academic Influence searches open-source data in three massive sources - Wikipedia, Semantic Scholar and CrossRef - for papers, chapters, books, and citations to individuals worldwide. Collectively, these databases contain billions of continually updated data points about millions of individuals' achievements.

Then, an institution's influence score is calculated by combining all "mentions" of the individuals who've been associated with it as faculty, administrators or alums. That score is then divided by the school's total number of students so that small and mid-sized schools have an equal chance of competing with larger colleges. A small school with proportionately more influential faculty than a large school, whose absolute influence may be bigger, will nonetheless score higher using what's termed "Concentrated Influence."

To control for various confounds, the data are restricted to the past ten years, and the names of famous politicians, artists, performers, and athletes are suppressed so they don't exert an inordinate, misleading influence on a school's ranking. The University of North Carolina doesn't get a boost from Michael Jordan, and the University of Kentucky receives no advantage from alum and star actress Ashley Judd.

Article by Kris Putnam-Walkerly, Contributor
Published Apr 29, 2024

Guiding Generosity: Overcoming Family Office Challenges In

Read More

Using its influence calculations, Academic Influence has churned out rankings for all kinds of schools and programs - research universities, global institutions, online MBAs, liberal arts colleges, online institutions, HBCUs, religious institutions and community colleges. It also ranks the best college in each state.

All those lists drive students, parents and other consumers to its website, attracting a growing audience of readers through organic searches. That traffic is critical to the for-profit company's business model — most of its revenue comes from generating applications by prospective students to colleges, which pay the company a fee for each student who expresses interest in applying. Licensing its proprietary ranking algorithm to other sites is another revenue stream for the company.

Since launching its website in August 2020, Academic Influence's traffic has grown from 50,000 to more than 1,000,000 visitors per year, as measured by Clicky.

Scholarly Impact

Academic Influence also ranks the individuals who've had the greatest influence in their respective academic fields using the same logic to that employed for ranking colleges and programs. Think of it as sizing up an academic's intellectual footprint on the web.

Currently, it provides data on scholars in 24 disciplines and approximately 300 sub-disciplines. So if you're interested in finding history's most influential philososophers, or the world's most influential mathematicians currently, you can find them at its website. Those lists are often broken down further by demographic categories — for example, who are the most influential women engineers, or the most influential Black scholars?

Evaluating scholarly impact has become a high-stakes exercise in academia. Universities often contract with services such as Academic Analytics to measure the scholarly impact of faculty members' research products — their grants, papers, journal articles and books - and then include that analysis as part of faculty promotion-and-tenure dossiers.

Counting publications and citations by others to those publications are common measures of scholarly impact. One particularly influential metric is the h-index, named after its originator, physicist Jorge E. Hirsch. It combines the volume of an author's publications with the extent of their citations into a single measure of impact.

Macosko, who serves as the President and Research Director of Academic Influence, believes his company's method of computing influence based on thousands of inputs is an improvement over combining just the h-index's two inputs of publication volume and citation extensiveness.

He points to the fact that persons at the top of Academic Influence's lists often don't have an ultra-high h-index, a fact that points to some of the index's well-known limitations. For example, it does not distinguish between review articles, which are highly cited, and empirical studies, which report original data but get fewer citations. An author who

produced more of the former might earn a higher h-index than the scholar who published an equal number of empirical articles containing important new discoveries.

In addition, the h-index is largely insensitive to publications that receive a large number of citations, thereby possibly understating the impact of a particularly influential contribution.

The quality of academic departments and graduate programs, especially in the sciences, are also often judged by big data-driven impact assessments. Those data are then used by universities to inform strategic decisions such as which academic departments should be strengthened and which should be diminished or which research areas represent an institution's best prospects for more research funding.

As a result of these kinds of consequential applications, faculty are cautious — often even suspicious – of relying on big data to judge the quality of their work. Their critiques need to be taken seriously. No measure of impact is flawless, including those generated by the algorithms at Academic Influence.

As one example, a person could receive many mentions for his or her work largely because it's publicly controversial rather than fundamental or impactful in an academic field. Take the psychologist Jordan Peterson as an example. Currently, he ranks 31st among the world's most influential clinical psychologists, according to Academic Influence. However, among many academic psychologists, Peterson would be viewed critically as more of a cultural figure or a divisive voice, rather than a credible academic.

When I recently asked Macosko about this objection, he told me the example of Peterson shows how his company's measure of influence differs from typical metrics. "Someone like Peterson, or even better, Claude Shannon, the creator of the field of information science, highlights why we think our metric adds value. Claude Shannon, when measured with traditional metrics, rises only to the middle of the pack with an h-index of 66. However, using our metric, he rightly climbs to the top, with a #14 overall ranking and #2 in computer science. Thus, we are giving higher education a new tool that gets at the heart of an academic's influence, and it's a tool that draws from orders of magnitude more datapoints than merely what gets fed into h-index and similar measures."

Academic Influence aims to be a disrupter in the college ranking business and other academic impact analyses. Its founders believe its computational methodology eliminates much of what they call the "gameability" of other approaches – the thumbs-on-the-scale that allow subjective factors to get mixed in with numeric indicators of institutional and scholarly impact. But whether removing most human judgment from those processes eliminates only noise and not also some signal remains a key question.

As with all rankings, it's best for consumers to view those from Academic Influence as one piece in the puzzle rather than a final answer. As Macosko said, right after he mentioned Claude Shannon's ranking, "If an academic department wants to use our algorithm to help with a faculty search or in the tenure and promotion process, they don't have to worry about getting tricked into hiring someone who is merely a culture figure or only a lightening rod for controversy."

only a lightening rod for controversy.

"Obviously, an academic department will want to use our algorithm to help them make a wise decision about candidates they've already vetted," he added. "Our information provides another tool in their toolbox when it comes to doing what departments have always needed to do: find the best people. That's why the algorithm at AcademicInfluence.com always starts with people. It's 'people power' that makes academia the force of good that it has been and can continue to be."

Follow me on *Twitter*.

Michael T. Nietzel

Editorial Standards   Print   Reprints & Permissions

Join Our Conversation

One Community. Many Voices. Create a free account to share your thoughts. Read our community guidelines here

🔔 Log in

Be the first to comment...

No one seems to have shared their thoughts on this topic yet

Leave a comment so your voice will be heard first.