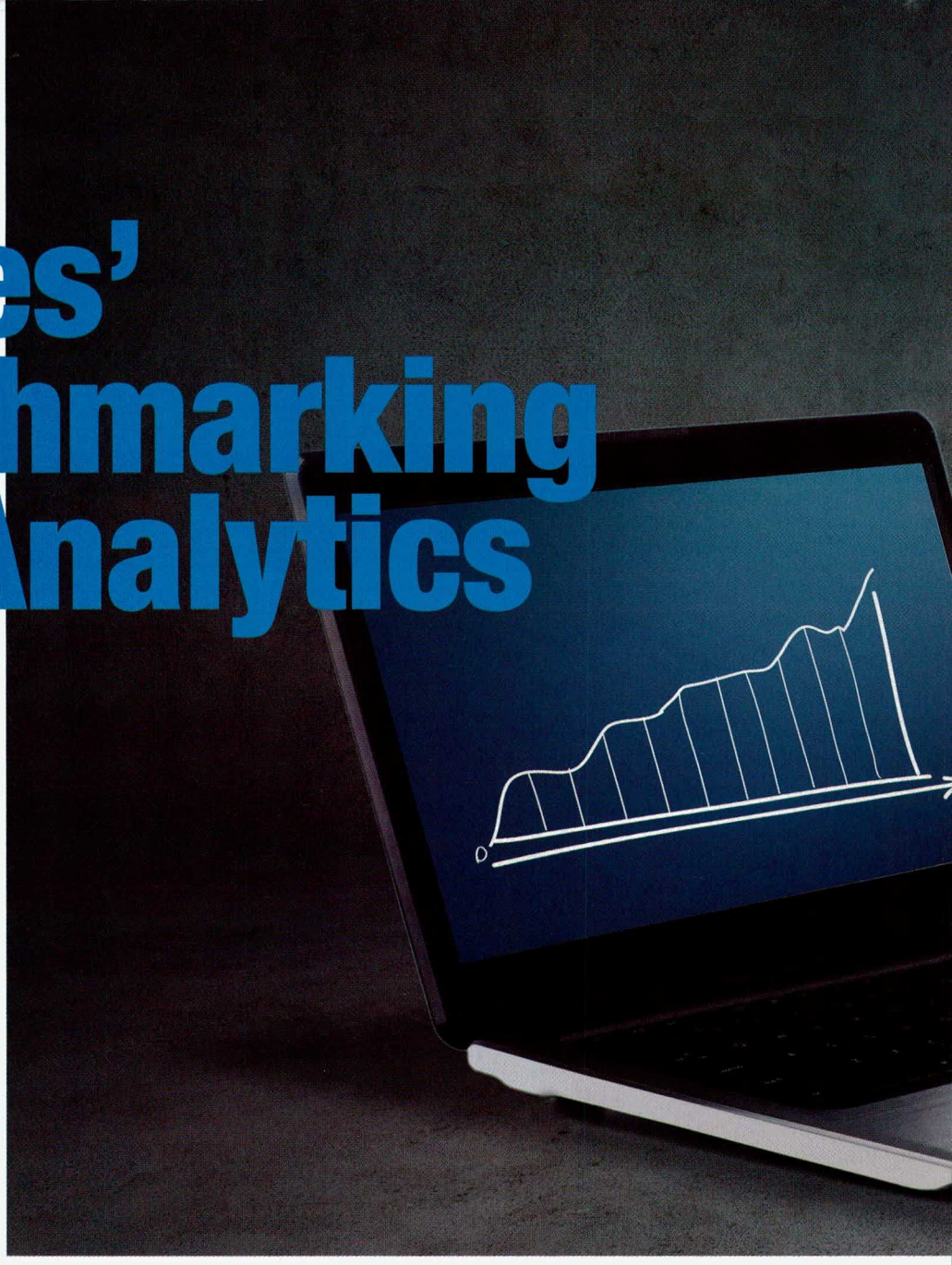
InCites' Benchmarking and Analytics



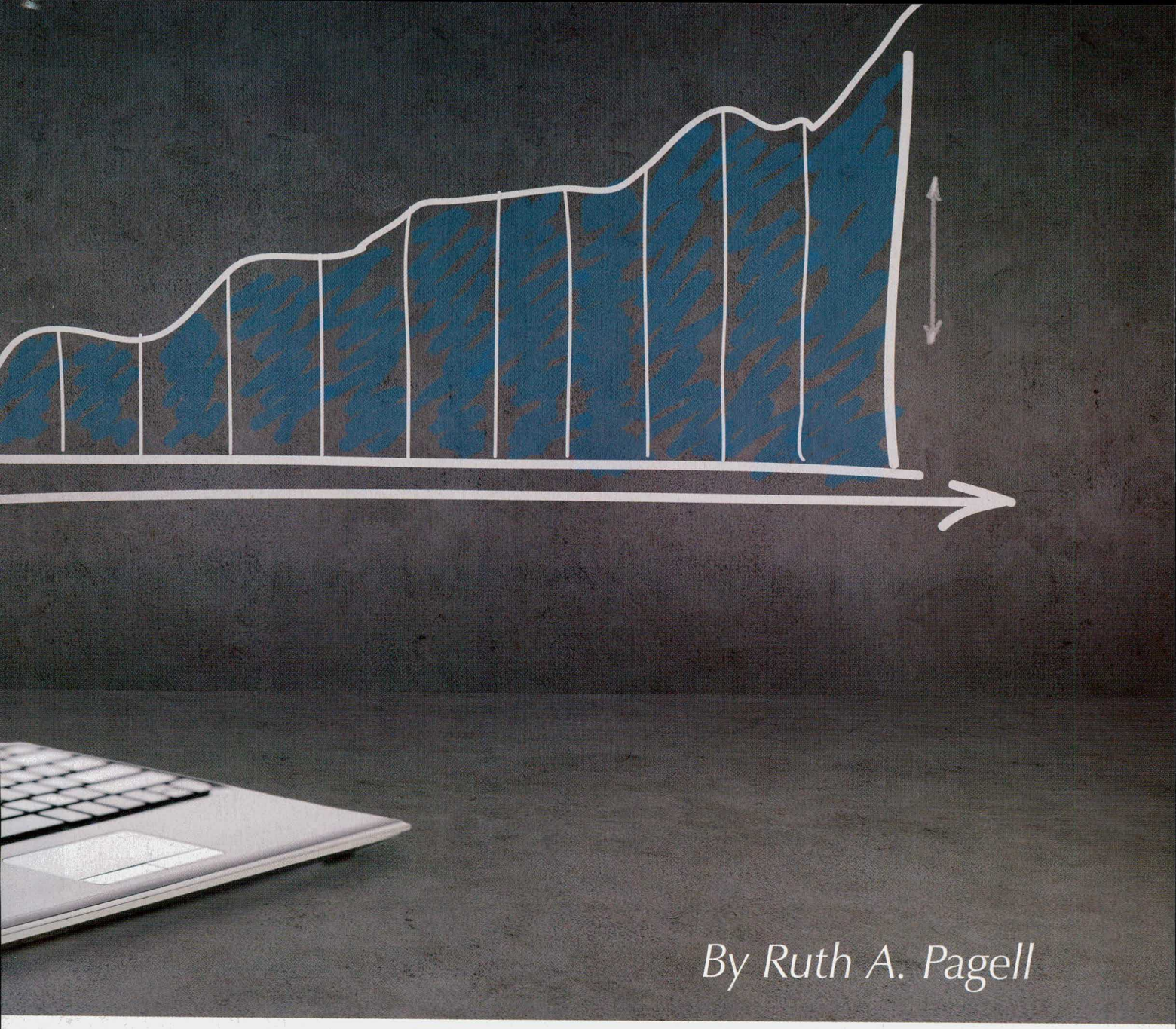
The new Thomson Reuters InCites platform for Journal Citation Reports (JCR) and Essential Science Indicators (ESI) frustrates me at times. I'm unable to combine a research area with an institution; find out the journals in which those researchers working within that area were publishing; identify the authors in that research area and list the other authors with whom they collaborate; or link to the Web of Science (WoS) records for the individual authors. However, with InCites Benchmarking and Analytics (B&A) tools, I can reach that level of detail. With more than 20 million documents, 100,000 publications, 4,500 institutions,

As with JCR and ESI, these capabilities come with the price of an additional subscription, but these B&A tools add greatly to the filtering options I discussed in my November/December 2014 article, "Insights Into InCites: Journal Citation Reports and Essential Science Indicators" in *Online Searcher* (Vol. 38, No. 6, pp.16–19).

MICRO-LEVEL BIBLIOMETRIC DATA

The InCites platform provides micro-level bibliometric data for a macro-level target audience of organizational senior administrators, policy makers, and funders. Subscribers have access to their institutions' research performance, as represented by WoS data, as well as to similar data for all institutions in the data file. The recently established bibliometrics services program at the National Institutes of Health (NIH) Library, led by informationists Chris Belter and Ya-Ling Lu, uses InCites to track the scientific output of NIH researchers, with emphasis on article production, research topics, and collaboration.

According to Thomson Reuters, the platform allows for "holistic research evaluation" down to the article level, covering 10 years of data for all editions and document types in the WoS Core Collection. This expands the number of documents for analysis but creates some pitfalls that are both common and uncommon for WoS users.



By Ruth A. Pagell

JCR and ESI contain, as source editions, Science Citation Index–Expanded and Social Science Citation Index; B&A adds Arts & Humanities Citation Index, Conferences for both Science and Social Science, and Books for Science and Social Science and Humanities. When it comes to document types, JCR and ESI include only articles and reviews, but B&A expands to proceedings, letters, editorials, books, and book chapters.

What kind of data can be gleaned from B&A? My searches revealed that Wuhan University has the most WoS publications in the research area of "Information Science & Library Science" of all listed Chinese Universities, and the Chinese Academy of Sciences has the most citations. Wang, Jing-feng has the most citations of the Chinese Academy authors; his paper with the most citations is a collaboration with Robert Haining at the University of Cambridge that was published in 2010 in the *International Journal of Geographical Information Science.*

ORGANIZATION OF INCITES

Subscribers to InCites access a Dashboard (for individually created reports), InCites preformatted "system reports," and five modules for individual analysis.

InCites System Reports include Research Performance, Collaborations, Trending Technology (which are news stories sourced from Recorded Future), and personal Tiles.

The Research Performance and Collaborations preformatted reports are at an institutional level and display standardized graphic representations and the underlying performance data for any institution listed in InCites. Research Performance includes Web of Science documents for the past 10 years, times cited per year, and normalized citation impact per year. It also has visuals for open access articles and leading research areas.

Collaborations among authors, institutions, and countries have become an integral part of 21st-century scholarship as faculty members partner with international colleagues on

>

research projects. Thomson Reuters captures this data in InCites and incorporates the results into THE (Times Higher Education) World University Rankings (timeshighereducation.co.uk/world-university-rankings/2014-15/world-ranking).

Reports include a map and a matrix of collaborating countries by subject area. Figure 1 below shows a reformatted collaboration matrix for a university in Singapore. Researchers in 10 scientific and technical disciplines have collaborated in varying degrees with researchers in nine other countries.

Tiles are personalized reports of results from analytics. The most important can also go on the Dashboard. The personal tiles shown in Figure 2 below are the result of two of my analyses—a map of leading authors in the Research Area of Meteorology and Atmospherics and a bubble graph of the top five universities represented in the research area of Information Science & Library Science in Essential Science Indicators.

SCORDCARDS FOR BENCHMARKING AND ANALYTICS

When learning about the capabilities of InCites' B&A, I spent most of my time elbow deep in data, exploring its five modules: People, Organizations, Regions (Countries), Research Areas, and Publications (Journals, Books, and Conference Proceedings). Analytics include special filters, analysis categories, and standard and additional output indicators for all modules. Although this isn't a tutorial on how to use InCites, in the spirit of this baseball fan who is watching the World Series as she writes, I have created scorecards to illustrate what is available where. Readers wishing to know more about the functionality of the different modules can watch a series of informal YouTube videos (youtube.com/watch?v=cJ20dP0dd0c&list=PLM1kuGdwRdGkEZ_bBSsQ0_18oNJh2GcrP&index=9), which includes links to other training videos).

My first scorecard is Indicators by Module. (See Table 1 on page 19.) Despite the large number of records, a small number of countries and institutions dominate the output:

- More than half of the 20 million documents are from five countries: the U.S., China, England, Germany, and Japan.
- More than 40% of the 147 million citations are from the U.S., followed by England, Germany, Mainland China, and Japan.
- 15% of the institutions have at least 10,000 documents and citations over the 10-year period.

	SINGAPORE	USA	CHINA MAINLAND	AUSTRALIA	JAPAN	ENGLAND	GERMANY (FED REP GER)	SOUTH KOREA	INDIA	HONG KONG
PHYSICS, APPLIED	2,877	406	560	95	100	96	79	89	71	51
MATERIALS SCIENCE, MULTIDISCIPLINARY	2,712	404	636	124	92	76	70	68	112	44
ENGINEERING, ELECTRICAL & ELECTRONIC	2,278	446	581	99	46	78	35	71	48	75
BIOCHEMISTRY & MOLECULAR BIOLOGY	2,002	651	300	160	156	145	115	55	51	57
CHEMISTRY, MULTIDISCIPLINARY	1,612	269	411	81	60	49	56	85	42	33
CHEMISTRY, PHYSICAL	1,579	219	438	97	71	52	51	31	42	25
CELL BIOLOGY	1,410	472	151	109	114	92	77	38	31	34
NANOSCIENCE & NANOTECHNOLOGY	1,383	222	295	46	55	47	44	38	49	24
ONCOLOGY	1,378	465	170	146	110	93	86	83	34	98
PHYSICS, CONDENSED MATTER	1,196	198	254	41	54	66	65	32	30	14

FIGURE 1. The collaboration matrix for a university in Singapore

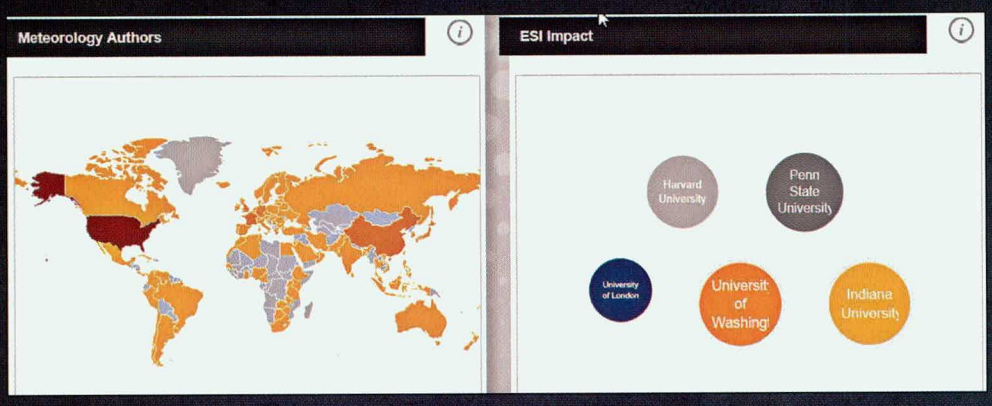


FIGURE 2. Personal tiles showing global reach and related universities

TABLE 1. Indicators by Module

	PEOPLE (Names)	ORGANIZATIONS	REGIONS	RESEARCH AREAS	PUBLICATIONS
NUMBER in MODULE	24,427,512	4525	230	251	101.952
TIME PERIOD	2004–2014	2004–2014	2004–2014	2004–2014	2004–2014
DOCUMENTS	20,084,205	20,407,479	20,407,479	20,407,479	N/A
STANDARD OUTPUTS					
• Name of	X	X	X	X	X
• Rank	X	X	X	X	X
• Web of Science Documents	X	X	X	X	X
• Times Cited	X	X	X	X	X
• Normalized Citation Impact [1]	X	X	X	X	
• % Documents Cited	X	X	X	X	
• Affiliation	X				
STANDARD FILTERS [2]					
By Attributes	X	X	X	X	X
By Research Network	X	X	X	X	X
By Research Output			X	X	X
THRESHOLDS: Minimum and maximum; WoS documents; Times cited; Time period	X	X		X	
INDICATORS FOR DISPLAY AND DOWNLOAD	21	17	18	19	5

1. Per paper normalized for subject, year and document type
2. See Table 2

- 13% of journals, books, and conference proceedings have at least 100 documents and citations over the 10-year period.

To navigate through InCites, start searching in any one of the categories and then use the Filters or Refocusing options to create customized lists. Highlight one or more items by "pinning" them to the benchmark list for easy comparisons.

My second scorecard, Customizing Results by Filters and Refocusing (see Table 2 on page 20), lists ways in which the data can be sliced and diced. The new download feature, with up to 21 different indicators, further expands the ability to manipulate and analyze the data, as shown in scorecard 3, Downloadable Indicators. (See Table 3 on pages 20.)

What are the research areas in which your institution has strengths? Starting in Institution, I selected University of Hawaii–Manoa, and refocused on research areas. Based on WoS documents and times cited, the top research area is Meteorology and Atmospheric Sciences. I continued drilling down (or Refocusing), step by step to reveal the top journal for UH, *Journal of Climate*, and the top UH author in that journal, Prof. Xie, Shang-Ping. The highest number of outside citations in that journal comes from Xie's collaboration with the Chinese Academy of Sciences, where his top collaborator for articles in this journal is Du, Yan. At any step, I can connect to the WoS article records.

VISUALIZING INCITES

Each set of results has visualization tools, including bar graph, geographic map, bubble graph, tree map, pie graph, trend graph, and the newest, radar chart. Choose to display any available indicator for the numbers of benchmarks you wish to include.

Different tools lend themselves better to different output. I do not see the visualization tools in B&A to be as useful as they are in JCR. However, the new radar chart is interesting. It shows comparisons across multiple indicators, as in Figure 3 on page 21, for the top Asian universities in the THE rankings.

PITFALLS

JCR and ESI are highly selective while InCites is all-inclusive. InCite's training encourages researchers "to use Threshold limiters to avoid meaningless reports based on too few documents/citations." This method is effective in all areas except people. Citation impact and percentage of documents cited for individual authors are the indicators most affected by "meaningless reports." Authors who have their names on one cited WoS document automatically get a 100%

TABLE 2. Customizing Results by Filtering and Refocusing

	PEOPLE	ORGANIZATIONS	REGIONS	RESEARCH AREAS	PUBLICATIONS
Collaborating Organizations	Filter Refocus	Refocus	Filter Refocus	Filter	Filter
Collaborating Organizational Associations		Filter			
Collaborating People	Filter Refocus	Filter Refocus	Refocus Filter	Filter	Filter
Collaborating Countries	Refocus	Filter Refocus			
Collaborating Regions			Refocus		
Collaborating Locations	Filter	Filter	Filter	Filter	Filter
Research Areas	Filter Refocus	Filter Refocus	Filter Refocus	Filter	Filter Refocus
Journals	Refocus Filter	Refocus Filter	Refocus Filter	Refocus Filter	Filter
Journal ISSN					Filter
Open Access	Filter	Filter	Filter	Filter	Filter
Affiliated Organizations	Filter	Refocus			
Affiliated Countries	Refocus	Refocus			
Affiliated People	Refocus	Refocus			
Associated Organizations			Refocus		
Associated Countries		Refocus	Refocus		
Organizations Name		Filter	Filter	Filter Refocus	Filter Refocus
Countries				Refocus	
People by Name or Researcher ID	Filter		Filter	Filter	Filter Refocus
Regions			Filter		Refocus
Location	Attribute	Filter		Filter	Filter
Document Type	Filter	Filter	Filter	Filter	Filter
Rank		Filter			

TABLE 3. Downloadable Indicators

		People	Organizations	Regions	Research Areas	Publications
1	Name	X	X	X	X	X
2	Rank	X	X	X	X	X
3	WoS Documents •	X	X	X	X	X
4	Normalized Citation Impact	X	X	X	X	
5	Times Cited •	X	X	X	X	X
6	% Documents Cited	X	X	X	X	X
7	ESI Most Cited (Yes/No)		X			
8	International Collaborations	X	X	X	X	
9	Average Percentile	X	X	X	X	
10	% Docs in Top 1%	X	X	X	X	
11	% Doc in Top 10%	X	X	X	X	
12	Journal Normalized Citation Impact	X	X		X	
13	% Highly Cited Papers	X	X	X	X	
14	Highly Cited Papers	X	X	X	X	
15	% Hot Papers	X	X	X	X	
16	% International Collaborations	X	X	X	X	
17	Impact Relative to the World	X	X	X	X	
18	Affiliation	X				
19	H Index	X				
20	% Industry Collaboration	X		X	X	
21	Researcher ID	X				
22	Citation Impact	X				

• *The only two indicators that have Threshold limiters*

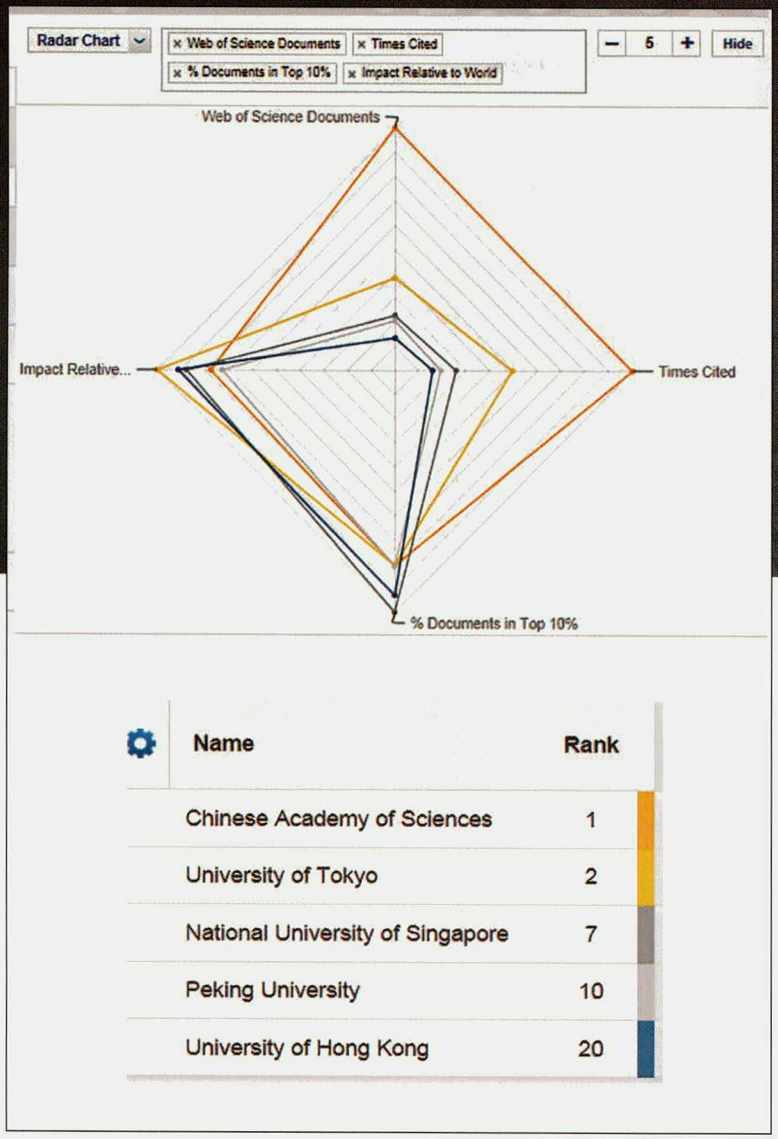


FIGURE 3: Radar chart of THE's top five Asian universities

citation rate. Many of these documents are highly cited, so their impact factors are in double digits.

Take our own profession as an example. *Library Journal* has the most documents in the IS & LS field, with more than 52,500 records and more than 6,000 authors. It also has a 1% rate of documents cited. By raising the Threshold to 10 citations, the number of authors drops to under 30.

However, this method also has its problems because of the longtime WoS issue of name disambiguation. Until 2008, WoS only used first and middle initials. Authors can voluntarily get a Researcher ID to unite their publications in one record, but many do not. Authors whose names appear in multiple formats but have more than 10 citations fall under the Threshold.

The author name in InCites with the most documents and citations is Wang, Y, who is listed as publishing more than 4,000 documents in 200 research areas. To set a lower limit on documents or citations is problematical.

In my meteorology example, Wang, Y ranks 19th in number of articles and 110th in times cited. Raising the limit to 10 WoS documents moves him up to 68. Setting the indicators to those of his peers, with at least 50 documents and 500 citations, he ranks 19th. Organizations reporting to their policy makers can improve their stories by playing with InCites Thresholds to create benchmarks based on comparable size.

WORK IN PROGRESS

The platform is still a work in progress, with a major update at the end of October 2014 that added export capabilities (in .csv format only) and enhanced renaming and saving of results shown in tiles. A video outlining the changes is narrated by Don Sechler (youtube.com/watch?v=yESFjCw7PeE&feature=youtu.be).

From being frustrated by too few combinations using JCR or ESI, I am overwhelmed by the number of available combinations in B&A, especially now that I can download data. In addition to the potential this product offers to decision makers, this tool is a gold mine for researchers who want to drill down to study the impact of different relationships on scholarly performance among the five categories.

Thanks to Don Sechler, who answered my many questions and included me in the InCites Group.

Ruth A. Pagell *(nellanoel@yahoo.com) is adjunct faculty library and information science program, University of Hawaii.*

Comments? Email the editor-in-chief (marydee@xmission.com)