



Grand Challenges in Measuring and Characterizing Scholarly Impact

Chaomei Chen*

College of Computing and Informatics, Drexel University, Philadelphia, PA, USA

Keywords: scholarly metrics, research assessment, scientometrics, scholarly communication, science and technology studies

INTRODUCTION

Scientists, policy makers, and the general public need to access, understand, and communicate scientific knowledge. As *Heilmeier's Catechism* advocated, researchers should be able to communicate the value of their research to the public regardless whether it is a mission to Mars or a search for a cure for cancer.

The constantly growing body of scholarly knowledge of science, technology, and humanities is an asset to mankind. While new discoveries expand the existing knowledge, they may simultaneously render some of it obsolete. It is crucial for scientists and other stakeholders to keep their knowledge up to date. Policy makers, decision makers, and the general public also need an efficient communication of scientific knowledge.

Research metrics and analytics aims to provide an open forum to address a diverse range of issues concerning the creation, adaptation, diffusion of scholarly knowledge, and advance quantitative and qualitative approaches to the study of scholarly knowledge. The following grand challenges illustrate some of the major issues concerning the interdisciplinary community.

GRAND CHALLENGE 1: ACCESSIBILITY

Scientific literature is increasingly volatile. *PLoS One* alone published 30,000 articles in 2014, an average of 85 articles per day¹. The Web of Science has accumulated over one billion cited references². The scale of retraction has stepped up – in one incidence, publishers retracted 120 gibberish papers simultaneously (Noorden, 2014). While it is easy to locate a paper that we are looking for, keeping abreast of the advances of scholarly work is a constant challenge.

In addition to the common focus on documents, more efficient and incrementally maintainable approaches should enable researchers to recognize and match information of interest beyond the constraints of the form or the language. The appropriate scope of a subject should be naturally and automatically expanded to attract documents through multiple types of intellectual linkage, such as semantic, linguistic, social, citation and usage, just as an experienced expert would do to grow his/her own oeuvre of domain expertise. In addition, the self-organized and updated oeuvre of knowledge should help us understand the significance of research at the same level of clarity as *Heilmeier's Catechism*. It will be fundamentally valuable to researchers and decision makers if new techniques can help us identify the state of the art of a topic more efficiently and effectively. For example, a reader can choose any topic of interest, and an intelligent system can generate a systematic review of the topic of the same quality as a panel of domain experts would produce.

OPEN ACCESS

Edited and Reviewed by:

Neil R. Smalheiser,
University of Illinois at Chicago, USA

*Correspondence:

Chaomei Chen
chaomei.chen@drexel.edu

Received: 08 January 2016

Accepted: 07 July 2016

Published: 21 July 2016

Citation:

Chen C (2016) Grand Challenges in Measuring and Characterizing Scholarly Impact. *Front. Res. Metr. Anal* 1:4. doi: 10.3389/frma.2016.00004

¹<http://blogs.plos.org/everyone/2014/01/06/thanking-peer-reviewers/>

²<http://stateofinnovation.thomsonreuters.com/web-of-science-1-billion-cited-references-and-counting>

GRAND CHALLENGE 2: CLARITY ON UNCERTAINTY

Scientific knowledge is never free of uncertainty. It is difficult to communicate uncertainty clearly, especially on issues with widespread concerns, such as climate change (Heffernan, 2007) and Ebola (Johnson and Slovic, 2015). The way in which the uncertainty of scientific knowledge is communicated to the public can influence the perceived level of risk and the trust (Johnson and Slovic, 2015).

A good understanding of the underlying landscape of uncertainty is essential, especially in areas where information is incomplete, contradictory, or completely missing. For instance, there is no information on how long Ebola virus can survive in the water environment (Bibby et al., 2015). If surrogates with similar physiological characteristics can be found, then any knowledge of such surrogates would be valuable. Currently, finding such surrogates in the literature presents a real challenge (Bibby et al., 2015).

Another form of uncertainty rises when new inputs alter the existing structure of scholarly knowledge. A new discovery may strengthen a previously weak or missing link as well as undermine or eliminate previously strong dependencies. Distortions may be introduced by citations and reinterpretations (Greenberg, 2009) or false claims made by retracted studies (Chen et al., 2013). In many areas, damages may remain unnoticed for a long time due to the lack of efficient and systematic mechanisms.

Active researchers are aware of such uncertainties in their areas of expertise. They choose words carefully and use hedging and other rhetorical mechanisms to convey their findings in the context of uncertainty. These common practices in scholarly communication have further increased the complexity of understanding science, especially for those without relevant expertise and for computational approaches. Future developments should enable stakeholders to access scholarly knowledge with a great degree of clarity on uncertainty as well as the knowledge itself.

GRAND CHALLENGE 3: CONNECTING DIVERSE PERSPECTIVES

The vast body of scholarly knowledge is a gold mine for making new discoveries. Pioneering efforts in literature-based discovery have demonstrated the value of connecting disparate bodies of knowledge discovery (Swanson, 1986; Smalheiser and Swanson, 1994; Cameron et al., 2013). The idea of a recombinant search in technology landscapes has a great impact (Fleming and Sorenson, 2001). An array of attempts have been made more recently to enhance the process of scientific discovery with publicly available knowledge, including detecting potentially transformative ideas and emerging trends based on structural variations (Chen, 2012), atypical combinations (Uzzi et al., 2013), diversity in interdisciplinary research (Rafols and Meyer, 2010), systematically generating and representing hypotheses (Soldatova and Rzhetsky, 2011; Malhotra et al., 2013), and the role of analogy in connecting different scientific domains (Small, 2010).

Research reveals that the influential ideas share a fundamental property – they tend to be richly interlinked with other ideas

(Goldschmidt and Tatsa, 2005). A profound theme shared by many of the attempts is the role of divergent thinking in scientific discovery, decision making, and creative problem solving, including the assessment of research excellence and impact. The value of reconciling multiple perspectives has been long recognized and advocated (Linstone, 1981). The point is not so much to enlist multiple perspectives in an interdisciplinary research team; rather, the key is to expose conflicting views on the same issue and resolve seemingly contradictory evidence at a new level (Chen, 2014).

To meet this challenge, new computational and analytic tools should enable researchers and evaluators to work with multiple perspectives directly. The unit of operation and analysis should focus on perspectives and paradigms as well as their premises, evidence, and chains of reasoning.

GRAND CHALLENGE 4: BENCHMARKS AND GOLD STANDARDS

Repositories of well-documented exemplar cases analyzed from multiple perspectives should be created, maintained, and shared with the research community so as to enable researchers to test and calibrate their metrics and analytic tools as well as reflect on lessons learned from these cases. Such repositories should include the most representative examples of high-impact scientific breakthroughs, the most complex cases of retracted studies, and the most extensive scientific debates in the history of science so that researchers can reproduce findings of previous studies. In particular, original datasets or queries that generate such datasets, metadata at various levels of granularity, narratives, and analytic procedures that have been applied by various studies should be preserved and made accessible. As shared resources, they will be valuable for the development and evaluation of new metrics and analytic capabilities as well as for preserving the provenance of scientific discoveries.

The role of readily available benchmarks and gold standards is crucial for a wide variety of scholarly activities. For example, Swanson's pioneering study of the possible linkage between fish oil and Raynaud's syndrome has become an exemplar case in literature-based discovery. Many subsequent studies validate newly introduced techniques with reference to the classic case. However, despite the fact that it is widely known as a classic case in literature-based discovery, the lack of essential benchmarks and gold standards makes it difficult to perform a systematic and comprehensive validation of scholarly metrics and analytic paths without spending a considerable amount of time and effort on reconstructing the vehicle for evaluation.

We can envisage how a shared repository would enable researchers to check out a snapshot of scientific knowledge exactly as what was available to Swanson when he conducted his classic study. The snapshot would include all the information that Swanson had used in his study and discoveries he made in his original study. In addition, the repository should register and preserve similar snapshots associated with subsequent studies inspired by Swanson's original work. Although subsequent studies may introduce new sources of data, different types of information, or a wider range of levels of granularity in comparison

with previous studies, gold standards should provide a consistent framework of reference such that one can systematically assess the efficiency and effectiveness of the application of a new approach to the same problem.

As research in research metrics and analytics advances, we can expect that new approaches will be able to reach scientific knowledge with a greater degree of depth and breadth than before. Consequently, the evaluation of new metrics and techniques requires gold standards at comparable levels of granularity. For instance, the novelty of a hypothesis can be established at different levels of abstraction, ranging from a simple link derived from co-occurrences of keywords, a semantic path that connects two concepts separated by many other concepts to an even broader context that, for instance, contains information reachable with a k -degree of separation. These levels of detail should be made readily accessible as part of the shared benchmark and gold standard repository.

GRAND CHALLENGE 5: INTEGRATING RESEARCH METRICS AND ANALYTICS

Scholarly metrics and qualitative studies of scientific discoveries and long-range foresights need to work together. The value of experts' opinions has been widely recognized. The challenge is in soliciting and synthesizing a wide variety of views from a diverse range of experts (Linstone and Turoff, 1975, 2011; Cozzens et al., 2010). As strongly advocated in the Leiden manifesto, scholarly metrics should serve the supporting role to qualitative and in-depth analytics of scholarly content and activities (Hicks et al., 2015).

Numerous scholarly metrics have been proposed, ranging from the widely known h -index, citation counts with or without field normalization, to altmetrics. Scholarly metrics are meant to be universal, quantifiable, field invariant, and easy to communicate (King, 2004; Bollen et al., 2009; Moed, 2010; Leydesdorff et al., 2011; Kaur et al., 2013). They convey extrinsic characteristics of research.

In contrast, scholars have examined prominent scientific discoveries in great detail from historical, sociological, and

philosophical viewpoints. Studies in this category aim to reveal intrinsic patterns that convey insights into critical paths leading to a breakthrough (Kuhn, 1962) or foresights into future developments (Martin, 2010). We will not be able to appreciate the significance of scholarly work until we learn about the perspective of the scholar, the focus of the attention, and the context of its origin.

A profound challenge to integrate the indicative power of research metrics and the insight-seeking analytic approaches is the difficulty in linking two perspectives that differ in so many ways at so many levels. A single perspective is not capable of characterizing and conveying the breadth and the depth of scholarly activities. Aggregation is often necessary but important details may be lost.

A problem of great challenge in one perspective may become resolvable in another. Field normalization, for example, has been intensively studied for improving the universality of research metrics. Drawing the boundary of a field or a discipline is notoriously hard. A more effective method may require a holistic view of interconnected disciplines. Many research questions may benefit from reconciling seemingly contradictory information. Until we are able to move back and forth between distinct perspectives efficiently and effectively, our ability to fully utilize the value of the scholarly knowledge that so many have spent so much effort to obtain would be rather limited.

In summary, the challenges outlined above illustrate the diverse range of theoretical and practical questions that may stimulate not only the study of research metrics and analytics but also the practice of research assessment, science policy, and many other aspects of our society. There are many more challenges ahead. Setting the study of research metrics and analytics on a holistic and integrative stage is a step toward fostering creative and impactful interactions between distinct perspectives and viewpoints.

AUTHOR CONTRIBUTIONS

The author is responsible for the entire article.

REFERENCES

- Bibby, K., Casson, L. W., Stachler, E., and Haas, C. N. (2015). Ebola virus persistence in the environment: state of the knowledge and research needs. *Environ. Sci. Technol. Lett.* 2, 2–6. doi:10.1021/acs.estlett.5b00193
- Bollen, J., Van de Sompel, H., Hagberg, A., and Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS ONE* 4:e6022. doi:10.1371/journal.pone.0006022
- Cameron, D., Bodenreider, O., Yalamanchili, H., Danh, T., Vallabhaneni, S., Thirunarayan, K., et al. (2013). A graph-based recovery and decomposition of Swanson's hypothesis using semantic predications. *J. Biomed. Inform.* 46, 238–251. doi:10.1016/j.jbi.2013.07.007
- Chen, C. (2012). Predictive effects of structural variation on citation counts. *J. Am. Soc. Inf. Sci. Technol.* 63, 431–449. doi:10.1002/asi.21694
- Chen, C. (2014). *The Fitness of Information: Quantitative Assessments of Critical Evidence*. Hoboken, NJ: Wiley.
- Chen, C., Hu, Z., Milbank, J., and Schultz, T. (2013). A visual analytic study of retracted articles in scientific literature. *J. Am. Soc. Inf. Sci. Technol.* 64, 234–253. doi:10.1002/asi.22755
- Cozzens, S., Gatchair, S., Kim, K.-S., Ordóñez, G., Porter, A., Lee, H. J., et al. (2010). Emerging technologies: quantitative identification and measurement. *Technol. Anal. Strateg. Manag.* 22, 361–376. doi:10.1080/09537321003647396
- Fleming, L., and Sorenson, O. (2001). Technology as a complex adaptive system: evidence from patent data. *Res. Policy* 30, 1019–1039. doi:10.1016/S0048-7333(00)00135-9
- Goldschmidt, G., and Tatsa, D. (2005). How good are good ideas? Correlates of design creativity. *Des. Stud.* 26, 593–611. doi:10.1016/j.destud.2005.02.004
- Greenberg, S. A. (2009). How citation distortions create unfounded authority: analysis of a citation network. *BMJ* 339:b2680. doi:10.1136/bmj.b2680
- Haas, C. N. (2014). On the quarantine period for Ebola virus. *PLOS Currents Outbreaks*. doi:10.1371/currents.outbreaks.2ab4b76ba7263ff0f084766e43abbd89
- Heffernan, O. (2007). Clarity on uncertainty. *Nat. Rep. Clim. Change*. doi:10.1038/climate.2007.57
- Hicks, D., Wouters, P., Waltman, L., Rijcke, S. D., and Rafols, I. (2015). Bibliometrics: the Leiden Manifesto for research metrics. *Nature* 520, 429–431. doi:10.1038/520429a
- Johnson, B. B., and Slovic, P. (2015). Fearing or fearsome Ebola communication? Keeping the public in the dark about possible post-21-day symptoms and

- infectiousness could backfire. *Health Risk Soc.* 17, 458–471. doi:10.1080/13698575.2015.1113237
- Kaur, J., Radicchi, F., and Menczer, F. (2013). Universality of scholarly impact metrics. *J. Informetr.* 7, 924–932. doi:10.1016/j.joi.2013.09.002
- King, D. A. (2004). The scientific impact of nations. *Nature* 430, 311–316. doi:10.1038/430311a
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Leydesdorff, L., Bornmann, L., Mutz, R., and Opthof, T. (2011). Turning the tables on citation analysis one more time: principles for comparing sets of documents. *J. Am. Soc. Inf. Sci. Technol.* 62, 1370–1381. doi:10.1002/asi.21636
- Linstone, H. A. (1981). The multiple perspective concept: with applications to technology assessment and other decision areas. *Technol. Forecast. Soc. Change* 20, 275–325. doi:10.1016/0040-1625(81)90062-7
- Linstone, H. A., and Turoff, M. (eds) (1975). *The Delphi Method*. Reading, MA: Addison-Wesley Publishing Co.
- Linstone, H. A., and Turoff, M. (2011). Delphi: a brief look backward and forward. *Technol. Forecast. Soc. Change* 78, 1712–1719. doi:10.1016/j.techfore.2010.09.003
- Malhotra, A., Younesi, E., Gurulingappa, H., and Hofmann-Apitius, M. (2013). ‘HypothesisFinder’: a strategy for the detection of speculative statements in scientific text. *PLoS Comput. Biol.* 9:e1003117. doi:10.1371/journal.pcbi.1003117
- Martin, B. R. (2010). The origins of the concept of ‘foresight’ in science and technology: an insider’s perspective. *Technol. Forecast. Soc. Change* 77, 1438–1447. doi:10.1016/j.techfore.2010.06.009
- Moed, H. F. (2010). Measuring contextual citation impact of scientific journals. *J. Informetr.* 4, 265–277. doi:10.1016/j.joi.2010.03.009
- Noorden, R. V. (2014). Publishers withdraw more than 120 gibberish papers. *Nature*. Rafols, I., and Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: case studies in bionanoscience. *Scientometrics* 82, 263–287. doi:10.1007/s11192-009-0041-y
- Smalheiser, N. R., and Swanson, D. R. (1994). Assessing a gap in the biomedical literature: magnesium deficiency and neurologic disease. *Neurosci. Res. Commun.* 15, 1–9.
- Small, H. (2010). Maps of science as interdisciplinary discourse: co-citation contexts and the role of analogy. *Scientometrics* 83, 835–849. doi:10.1007/s11192-009-0121-z
- Soldatova, L., and Rzhetsky, A. (2011). Representation of research hypotheses. *J. Biomed. Semantics* 2(Suppl. 2), S9. doi:10.1186/2041-1480-2-9
- Swanson, D. R. (1986). Fish oil, Raynaud’s syndrome, and undiscovered public knowledge. *Perspect. Biol. Med.* 30, 7–18. doi:10.1353/pbm.1986.0087
- Uzzi, B., Mukherjee, S., Stringer, M., and Jones, B. (2013). Atypical combinations and scientific impact. *Science* 342, 468–472. doi:10.1126/science.1240474
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.