Introduction of “Kriging” to Scientometrics for Representing Quality Indicators in Maps of Science

Masashi Shirabe[1]

[1]*shirabe.m.aa@m.titech.ac.jp*
Tokyo Institute of Technology, Oookayama 2-12-1 S6-5, Meguro
152-8550, Tokyo (Japan)

Introduction

Maps of science are an effective technique, especially for non-experts, to facilitate intuitive understanding of science activities, even though they could be cut both ways. Among such maps, science overlay maps have received adequate attention from scientometrics researchers (Perianes-Rodríguez et al., 2011; Grauwin & Jensen, 2011; Klaine et al., 2012; Leydesdorff, Rotlo, & Rafols, 2012; Boyack & Klavans, 2013; Gorjiara & Baldock, 2014). Actually they are an attractive approach “to visually locate bodies of research within the sciences, both at each moment of time and dynamically.” (Rafols, Porter, & Leydesdorff, 2010)

To produce science overlay maps, (1) we draw a basemap, which contains positional information of nodes from bibliographical data, then (2) we overlay other information on the basemap by assigning the information (i.e., indicators like publications and citations) to the nodes with such factors as colors and/or size of circles representing the nodes.

To think more abstractly, an essence of science overlay maps is “sharing” of positional information of nodes by different science maps, which are similar in concept to thematic maps in geography. What makes such “sharing” possible is the stability of global maps (Rafols, Porter, & Leydesdorff, 2010). This perspective could broaden choices of expressions in science overlay maps to improve our understandings. For example, VOSviewer (Van Eck & Waltman 2010) provides five different views, i.e., label view, density view, scatter view, cluster view, and cluster density view, for a fixed set of positional information of nodes. By switching these views, we can understand phenomena behind the maps deeply and multidimensionally. Therefore, introducing a new way to project bibliographical information on given maps is expected to expand availability of science overlay maps, just as a new method to produce thematic maps does in geography.

From this perspective, the author first pays attention to density view provided by VOSviewer. By mapping journals in the fields of Business, Business-Finance, Economics, Management, and Operations Research & Management Science, Van Eck and Waltman (2010, p. 529) explain functionality of the density view as follows: “The density view immediately reveals the general structure of the map. Especially the economics and management areas turn out to be important. These areas are very dense, which indicates that overall the journals in these areas receive a lot of citations.” As they pointed out, this view is helpful to outline the macro structures of maps and to show which areas in the maps are important. Basically, however, density view can be used only for representing quantitative indicators, because “the item density of a point in a map depends both on the number of neighboring items and on the weights of these items.” (p. 533) If citations were used as weights of items, the density map might be seen to show “quality” of areas. Actually, citation densities are only a representation of quantities. That is particularly evident in assuming to represent quality (impact) indicators like proportion of top 10% publications in the density view.

Judging from many scientometrics studies rely on density or heat maps (e.g., Pinto, Pulgarin, & Escalona, 2014), it would be reasonable to assume that graphical representations like the density view to represent quality indicators on science maps is very helpful to outline the structures of bibliographical data and to show which areas in maps of science are efficient, superior, or highly shared. Then, this paper introduces “kriging” to scientometrics for representing quality indicators.

Data

The author uses a data platform that consists of datasets from SCI Expanded, PubMed, and USPTO patent databases. By adopting matching methods developed in Shirabe (2014), records in PubMed are linked to those in SCI expanded, and non-patent references in the face sheets of US utility patents are also matched to records in SCI Expanded. As a result, three databases can be analyzed in an integrated fashion by using this platform.

This platform contains the product set (number of items is 8.5 millions) of SCI expanded (articles, reviews, letters, notes, and articles & proceedings papers; their database years are between 1992 and 2011) and PubMed (their publication years are between 1991 and 2012) as well as science citations of US utility patents registered between 1991 and 2012.

Method

First "macro and micro" basemaps are constructed by co-occurrence analysis of MeSH terms (Leydesdorff & Opthof 2013), where VOSviewer is used for mapping and clustering. For making the macro map, all the items of the product set are included in the analysis, and only third layer descriptors are treated as subjects of co-occurrence analysis. For that, lower layers' MeSH terms are replaced by their higher taxon. For making the micro map, only items containing mesenchymal stromal cells, mesenchymal stromal cell transplantation, totipotent stem cells, multipotent stem cell, induced pluripotent stem cells, pluripotent stem cells, and embryonic stem cells as their MeSH terms are included in analysis. Top 150 MeSH terms (except highly shared terms) are used in co-keyword analysis. Thus, this micro map is a map of pluripotent stem cell research.

Secondly, sets of data overlaying on the basemaps are produced. For that, positional data (i.e., two-dimensional position coordinate) of nodes produced by VOSviewer are transmitted to SAGA (Böhner, McCloy, & Strobl, 2006). Then, overlaying data for density maps (by Gaussian kernel function) or those for isograms (by kriging) are calculated from bibliographic indicators and overlaid on the basemaps.

Results

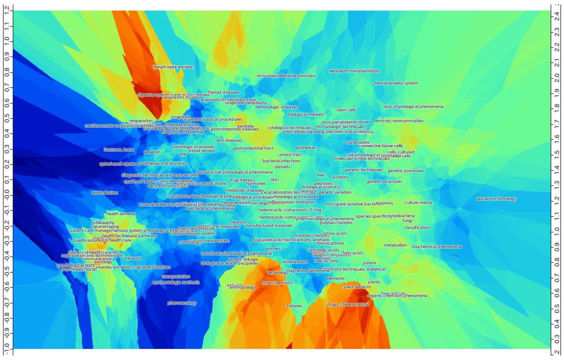

Figure 1. Japanese Share of Life-Science Papers cited by US Patents Registered between 2001-11.

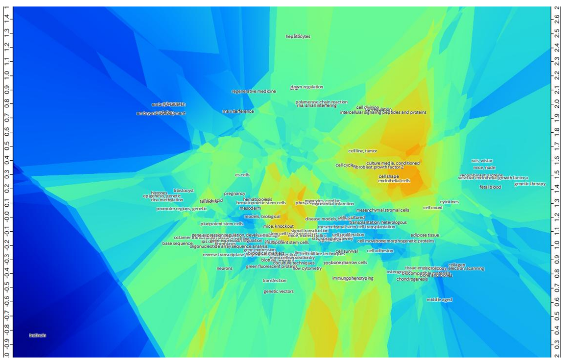

Figure 2. Japan's Relative Frequencies of Top 10% Cited Papers in Stem Cell Research.

The above figures show examples of overlay maps to represent quality indicators. They make it easier to understand the quality of Japanese research outputs intuitively and multidimensionally either at macro or micro level.

Acknowledgments

This research is partly supported by JST/RISTEX research funding program "Science of Science, Technology and Innovation Policy".

References

Böhner, J., McCloy, K.R., Strobl, J. (Eds.) (2006). *SAGA – Analysis and Modelling Applications*. Göttinger Geographische Abhandlungen, Vol. 115.

Boyack, K.W., & Klavans, R. (2013). Creation of a Highly Detailed, Dynamic, Global Model and Map of Science. *JASIS&T*, *65*(4), 670–685.

Gorjiara, T., & Baldock, C. (2014). Nanoscience and nanotechnology research publications: a comparison between Australia and the rest of the world. *Scientometrics*, *100*(1), 121-148.

Grauwin, S. & Jensen, P. (2011). Mapping scientific institutions. *Scientometrics*, *89*(3), 943-954.

Klaine, S. J., Koelmans, A. A., Horne, N., Carley, S., Handy, R. D., Kapustka, L., Nowack, B., & von der Kammer, F. (2012). Paradigms to Assess the Environmental Impact of Manufactured Nanomaterials. *Environmental Toxicology and Chemistry*, *31*(1), 3-14.

Leydesdorff, L., & Opthof, T. (2013). Citation analysis with Medical Subject Headings (MeSH) using the Web of Knowledge: A new routine. *JASIS&T*, *64*(5), 1076–1080.

Leydesdorff, L., Rotlo, D., & Rafols, I. (2012). Bibliometric Perspectives on Medical Innovation Using the Medical Subject Headings of PubMed. *JASIS&T*, *63*(11), 2239–2253.

Perianes-Rodríguez, A., O'Hare, A., Hopkins, M. M., Nightingale, P., & Rafols, I. (2011). Benchmarking and visualising the knowledge base of pharmaceutical firms (1995-2009), *Proceedings of ISSI 2011*, Vol. II, 656-661.

Pinto, M., Pulgarin, A., & Escalona, M. I. (2014). Viewing information literacy concepts: a comparison of two branches of knowledge. *Scientometrics*, *98*(3), 2311-2329.

Rafols, I., Porter, A. L. & Leydesdorff, L. (2010). Science Overlay Maps: A New Tool for Research Policy and Library Management. *JASIS&T*, *61*(9), 1871–1887.

Shirabe, M. (2014). Identifying SCI covered publications within non-patent references in U.S. utility patents. *Scientometrics*, *101*(2), 999-1014.

Van Eck, N.J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523-538.