

# Diagnosing Emerging Science: The Cases of the "New Science of Networks" and Scientometrics

Haiko H. Lietz<sup>1</sup>

Institute for Research Information and Quality Assurance, Godesberger Allee 90, 53175 Bonn, Germany

Competence Centre for Bibliometrics, Godesberger Allee 90, 53175 Bonn, Germany

University of Duisburg-Essen, Faculty of Social Sciences, Institute of Sociology, Lotharstr. 65, 47057 Duisburg, Germany

Hochschule Mittweida – University of Applied Sciences, Faculty of Mathematics, Physics and Computer Science, Technikumplatz 17, 09648 Mittweida, Germany

lietz@forschungsinfo.de

## Introduction

---

How and why do some scientific fields thrive and establish themselves while others do not? What is "emerging science"? Contributions to answers can be found in the classics, most visibly in Solla Price (1963) who studied growth dynamics of science, but also in Kuhn (1962) who described the modus operandi of normal science and occasional paradigm shifts where a new order emerges. Recently, Bettencourt et al. have shown for various scientific fields that they can well be described by logistic growth models (2008) and that they undergo a topological transition in their collaboration structure on their way to maturity (2009). This transition can be measured in terms of the densification and the emergence of a giant component of a field's co-authorship network. According to their theory, this transition in the social structure mirrors the establishment of shared concepts or paradigms in the cultural structure of the field. Bettencourt et al. apply scaling analysis (2007) to measure network densification and field productivity.

Two hypotheses for two scientific fields will be tested. The end of the nineties brought breakthroughs in the field that studies complex networks. Since much of the work deals with social networks, this science entered the sociological canon to such an extent that it came to be labelled as the "invasion of the physicists" (Bonacich 2004). H1: The "New Science of Networks" (NSoN) has undergone a topological transition, mirroring the diffusion of shared scientific concepts.

On the other hand, Scientometrics, the field that studies science using bibliometric and informetric data, has been criticized for a lack of standards (Larsen 2008). H2: The Field of Scientometrics (FoS) has not undergone a topological transition, mirroring the lack of diffusion of shared scientific concepts.

Data was extracted from the Web of Science® databases SCI-EXPANDED and SSCI in the timespan 1900-2008, using keyword searches. Author names were hand-cleaned. Only the following document types were selected: Article, Bibliography, Discussion, Editorial Material, Letter, Meeting Abstract, Note, Proceedings Paper, and Review. All selections made resulted in 6,321 publications for the NSoN, and 4,382 publications for the FoS.

Complete counting is used for authors, i.e. every author in a publication receives 1 score. Co-authorship networks are produced by linking authors who have co-authored a publication. Because the real collaboration structure is unknown, per publication all authors are commonly connected in a complete graph (Newman 2001). In this case edges obviously can't represent communication loads because then the load would be the same whether two scientists are the sole authors of a publication or members of a 100-author-team. Such networks have been labelled catnets (category-networks) because they imply a collaboration potential due to shared attributes (White 2008: pp.52) – in this case a co-authored publication. To study social networks that describe collaboration in a probabilistic sense fractional counting (Gauffriau et al. 2007) is used. Here, per publication each co-authorship is given a score of  $1/(n-1)$ , where  $n$  is the number of authors per publication, resulting in an edge sum of 1 per author and publication.

---

<sup>1</sup> Help with data processing by Andre Müller is gratefully acknowledged.

Co-authorship networks are visualized using the CASOS Network Visualizer. The underlying TouchGraph spring embedder causes connected nodes to attract each other, the more so the denser clusters are. On the other hand, nodes generally repel each other, causing stringy networks to form lines and unconnected components to separate. Large components have a large inertia, so they will gather in the center of the network.

Co-authorship networks are tested for the hypothesized topological transition, measurable in terms of network densification and the emergence of a giant component. Networks densify over time if the exponent  $\alpha$  of the scaling law  $\text{edges} = A \cdot (\text{nodes})^\alpha$ , where  $A$  is a constant, is larger than 1. The size of the largest component is simply the number of nodes in the largest network component. In addition, field productivity over time is measured. A field is productive if the exponent  $\beta$  of the scaling law  $\Delta \text{ publications} = B \cdot (\Delta \text{ authors})^\beta$ , where  $B$  is a constant, is larger than 1. Exponents are estimated using least-squares fitting.

Neither hypothesis needs to be rejected. The NSoN started to emerge in the early nineties and growth accelerated following the publication of two key papers at the end of the century. During these growth periods the field was quite productive, it densified according to both counting methods, and a giant component emerged in 2006. The FoS showed earlier signs of sustainable growth which go back to the middle of the seventies. As in the NSoN growth accelerated in the 90s which is probably related to the increasing availability of data and computing power. Yet the field was never productive according to the indicator used. The field seems to densify, at least starting in the nineties, even though whole and fractional counting differ much more than for the NSoN. Most importantly, no giant component has emerged until 2008.

The scaling indicators, in combination with giant component analysis and network visualization, are powerful tools to diagnose emerging science. Cautiousness in interpreting results is required because the densification measure is not robust when different counting methods are used, confirming earlier results (Gaufriau & Larsen 2005). More data for different fields is desirable to gain insights into how the indicator behaves under normalization. In general, scale-independent indicators are the indicators of choice because they don't attempt to describe scale-invariant systems by linear means (Katz 1999).

Integrating classical work by Solla Price and Kuhn, the case studies presented here have shown that social and cultural analysis can plausibly be linked using network concepts (White 2008). In light of the reasons for formulating the hypotheses, the results cohere with the explanation that those scientific fields thrive and establish whose scientists have a common style of practice.

Policy Implications will be discussed.

## References

---

- Bettencourt, Luís M. A., David I. Kaiser, and Jasleen Kaur. 2009. "Scientific discovery and topological transitions in collaboration networks." *Journal of Informetrics* 3:210-221.
- Bettencourt, Luís M. A., David I. Kaiser, Jasleen Kaur, Carlos Castillo-Chávez, and David E. Wojick. 2008. "Population Modeling of the Emergence and Development of Scientific Fields." *Scientometrics* 75:495-518.
- Bettencourt, Luís M. A., José Lobo, Dirk Helbing, Christian Kühnert, and Geoffrey B. West. 2007. "Growth, innovation, scaling, and the pace of life in cities." *Proceedings of the National Academy of Sciences* 104:7301-7306.
- Bonacich, Phillip. 2004. "The Invasion of the Physicists." *Social Networks* 26:285-288.
- Gaufriau, Marianne, and Peder O. Larsen. 2005. "Counting methods are decisive for rankings based on publication and citation studies." *Scientometrics* 64:85-93.
- Gaufriau, Marianne, Peder O. Larsen, Isabelle Maye, Anne Roulin-Perriard, and Markus von Ins. 2007. "Publication, cooperation and productivity measures in scientific research." *Scientometrics* 73:175-214.
- Katz, J. Sylvan. 1999. "The self-similar science system." *Research Policy* 28:501-517.
- Kuhn, Thomas S. 1962. *The Structure of Scientific Revolutions*. University of Chicago Press.
- Larsen, Peder O. 2008. "The state of the art in publication counting." *Scientometrics* 77:235-251.
- Newman, M. E. J. 2001. "Scientific collaboration networks. I. Network construction and fundamental results." *Physical Review E* 64:016131.
- Solla Price, Derek J. de. 1963. *Litte Science, Big Science*. Columbia University Press.
- White, Harrison C. 2008. *Identity and Control: How Social Formations Emerge*, 2nd ed. Princeton University Press.