

12-1-2011

## The beginning of a beautiful friendship - social sciences research and industry products

Gali Halevi  
*Elsevier*

Follow this and additional works at: <https://www.researchtrends.com/researchtrends>

---

### Recommended Citation

Halevi, Gali (2011) "The beginning of a beautiful friendship - social sciences research and industry products," *Research Trends*: Vol. 1 : Iss. 25 , Article 6.

Available at: <https://www.researchtrends.com/researchtrends/vol1/iss25/6>

This Article is brought to you for free and open access by Research Trends. It has been accepted for inclusion in Research Trends by an authorized editor of Research Trends. For more information, please contact [r.herbert@elsevier.com](mailto:r.herbert@elsevier.com).

## Section 5: Behind the data

The Beginning of a Beautiful Friendship – Social Sciences Research and Industry Products

Gali Halevi

### Co-citation analysis morphs into patents

If you think the term 'co-citation analysis' is strictly related to bibliometrics, and lays within the realm of Library and Information Science, think again!

The term was coined in 1973 when Henry Small described this methodology in his renowned article 'Co-citation in the scientific literature: a new measure of the relationship between two documents', which was cited hundreds of times in the following years. Co-citation analysis has been used for numerous discoveries such as the emergence of author networks in a particular field of investigation, and mapping the emergence of a scientific field<sup>1-3</sup>. Small's technique of using co-citations to analyze a body of research in order to discover scientific trends and collaborations (among other things) was in the most part used in the social sciences; then in the late 1990s the term suddenly started appearing not only in academic discussions but also in patent literature, and the approach has been used by industry pioneers such as Xerox, IBM, Microsoft, and more recently Google.

### Technology catches up

How did co-citation analysis, which was conceived as a way to analyze relationships between articles, morph into a product apparatus? Looking at the time frame when patents using the term appear (see [Figure 1](#)), it seems that computational development at that time might give us an answer. The late 1990s saw a surge in different types of electronic databases which enabled indexing and retrieval of documents. These years also saw an increase in both co-citation analysis research publications and patents alike, including some interesting surges in 2009 and 2010, when semantic information retrieval and web page rankings took center stage in web-based applications. The first patent that uses co-citation analysis as a method is a Xerox patent that uses co-citation analysis to generate clusters of documents in a database<sup>4</sup>.

### References:

1. Culnan, M. J. (1987) Mapping the intellectual structure of MIS, 1980-1985: a co-citation analysis. *MIS Quarterly: Management Information Systems*, Vol. 11, pp. 341-350.
2. Chen, C. (1999) Visualising semantic spaces and author co-citation networks in digital libraries. *Information Processing and Management*, Vol. 35, pp. 401-420.
3. Zhao, D., & Strotmann, A. (2011) Intellectual structure of stem cell research: a comprehensive author co-citation analysis of a highly collaborative and multidisciplinary field. *Scientometrics*, Vol. 87, 115-131.
4. US Patent 6038574 (2000) Method and apparatus for clustering a collection of linked documents using co-citation analysis.

This patent (US Patent 6038574) describes a co-citation analysis method in which hyperlinks in web pages are viewed as references, and the relationships found are used to help create a web page index. Following the Xerox patent application, a series of patents mentioning co-citation analysis in either the method, claims, or references to prior publications appear through the patent literature featuring prominent technological companies (see [Figure 2](#)).

Looking at some of the patents one can see some interesting applications of co-citation analysis: for example Google Patent (WO2008134373A1), which uses "co-citation analysis, or anchor text analysis (e.g., analysis of text in or near links to the multimedia events) to determine related multimedia events" (Description of Embodiment; 0035). IBM applied co-citation analysis to link analysis on the Web (US7792827B2), and Microsoft used co-citation analysis in their development of latent semantic analysis (US20070239431A1).

### Technology? Transferred!

This type of relationship between academic research and product development is a phenomenon that we have learned to expect in areas of medicine or engineering, for example. In these areas, we can more easily find a direct link between theoretical models and their growth into tangible products. What this analysis shows is that social science research can go through a similar evolution, but the cycle through which it develops takes longer to deploy; in this case more than 20 years.

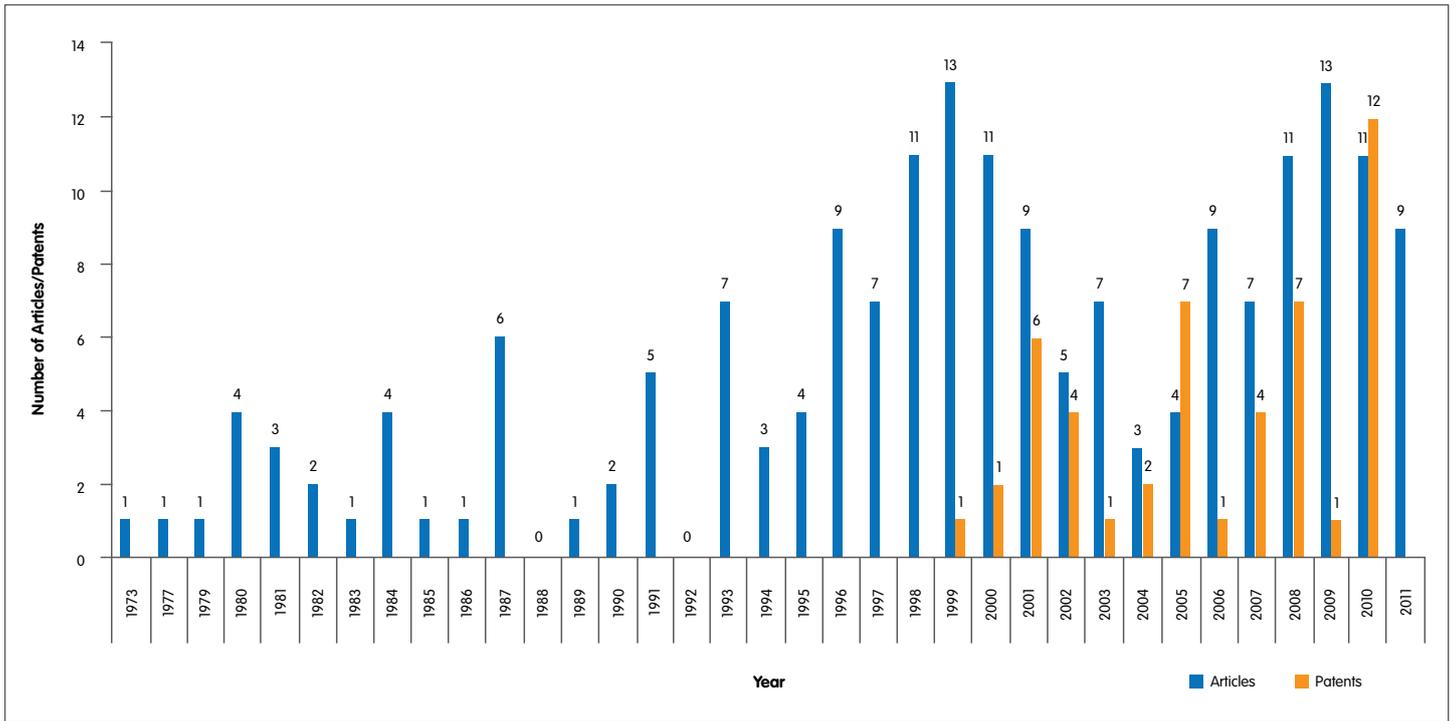


Figure 1 – Occurrence of the term ‘co-citation analysis’ in peer reviewed articles versus patents. Sources: TotalPatent and Scopus.

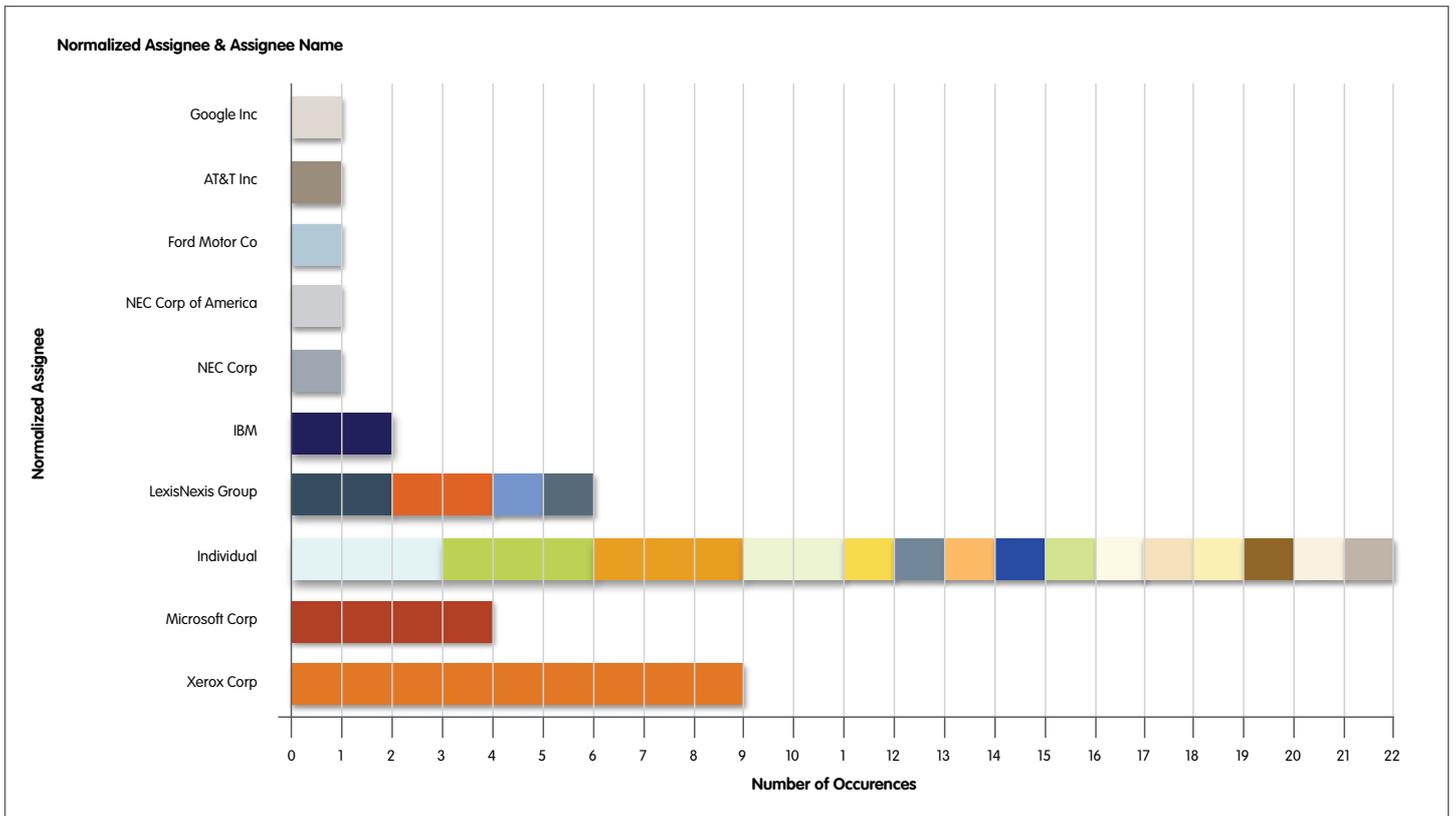


Figure 2 – Co-citation patents assignees. Source: TotalPatent.