

Citations within the GI literature – Theoretical works or good examples?

Gustaf Nelhans

Department of History of Ideas and Theory of Science, Göteborg University

SUMMARY

The aim of this study is to discuss the role of the term 'Geographic Information Science' in citations within the GI literature. Citation data from 12.144 articles in Geographic information systems and science were collected from the Science Citation Index (SCI). These data were analyzed by means of visual data mining based on citation patterns in the collected data.

The study shows that there exists a bias in the patterns of citations in GI literature, where the top three GIS papers outnumber the top three GIScience papers by an almost 3 to 1 ratio, a total of 225 versus 78 citations. Furthermore, it is shown that not more than two papers containing the term GIScience make it into the top 200 cited GIS papers in terms of acquired citations from the collection of 12.144 papers.

It is suggested that researchers in GIS rather turn to good examples of research than to normative theories of research when they refer to work by others. Furthermore, the term GIScience, which has been put forward as a concept linked to the core of GIS research, is shown to have low visibility in the data analyzed.

In conclusion a set of questions regarding the role of GIScience in the research community are posed.

INTRODUCTION

The term 'geographic information science' has been a buzz-word in the research community for more than a decade since it was coined by Michael Goodchild in the early 1990:s (Goodchild, 1992). It has been associated with a tension in the GI literature between GIScience as the theoretical 'underpinnings', and GISystems as the technology of geographic information. Although this tension has been questioned by critics of GIS as well as by GI researchers, the term still maintains a high status as evident from its influence in the naming of the research discipline, key journals (e.g. IJGIS and CaGIS), professional organizations (UCGIS and, in the description of AGILE), research agendas (e.g. McMaster et al., 2004), and educational materials (e.g. Longley et al., 2005).

In this study, the ambiguous interpretations of GIS as *GISystems* or *GIScience* are tested by an analysis of the citation patterns between papers in the literature on GIS. My goal is not to question GIScience as a legitimate research area, but to engage in an ongoing discussion (e.g. Chrisman, 2003; Fisher et al., 2005) about the contents and settings of GIScience in general and to present one way to identify empirical data about how it is mirrored in the scientific literature on geographic information.

The use of the SCI in sociology and history of science was suggested by Eugene Garfield (1971). Recently, the computer software HistCite™¹ was released, which "...will be used to help the searcher quickly identify the most significant work on a topic and enable the searcher to trace its year-by-year historical development" (Garfield et al., 2003, p. 183).

¹ <http://www.histcite.com>

The critics of the SCI claim that it does not hold complete coverage of published articles within a given research area and, further, that publication patterns in different research areas makes it impossible to compare research by use of the SCI (Monastersky, 2005).

Furthermore, citation data analyses might not show what their users expect. Aant Elzinga, writing in a science policy context, suggests that the SCI exhibits inequality in who counts as a speaker:

Considering that citation counts only indicate *visibility* and not *quality* in science (something users of these citations sometimes confuse), we can only say that Garfield's map constructs a picture of who counts as a speaker and who does not. Societal relevance does not enter into the picture either. (Elzinga, 1995)

Despite these objections, I assume that citation analyses could provide interesting empirical material of a semi quantitative sort, about the *visibility* of certain aspects of scientific writings.

DATA AND METHODS

To do a citation study of journal papers within the SCI, two discernable approaches are possible. One is to start with a given set of sources, (i.e. a specific journal) trying to identify the citation patterns of papers published within that journal. This approach was used by Fisher (2001), who evaluated the ranking of a specific journal (International Journal Geographical Information Systems/Science), together with an analysis of the citation frequency of papers published therein.

The second possible approach for citation analysis, which is used in the present study is to follow the path taken by (Dorn, 2002), in which papers containing terms associated with the concept 'geomorphology' cited more than 20 times where analyzed.

The difference between the two approaches could be boiled down to a distinction between how data to be analyzed should be collected. On the one hand the set of data to be analyzed could be defined by an a priori judgment of what is to be defined to be "within GIScience" as was done in the study by Fisher. With the second approach instead, I take the tension of definitions of 'GIS' and 'GIScience' as used by researchers in the field as a starting point to try to elicit in what way GIScience papers are referred to by the GIS community.

In this study, the following method was used to collect two sets of data from the SCI:

First, a collection of 12.144 papers were acquired by use of the following search string:

TS²=geograph* information system* OR TS=geograph* information science* OR TS=(GIS)

The time span was 1992 - 2006 for the three available databases provided by the SCI³.

Second, an additional 88 citations were retrieved:

TS=(gisci) OR TS=(geograph* information science) OR TS=(GIScience)

The citation patterns of the most cited papers in the GIS collection (Figure 1) were plotted by use of HistCite™ and visually analyzed. Thereafter the second (GIScience) collection of data was plotted in the same way (Figure 2). Finally, matches between these two "citation maps" were identified.

² The 'TS' field tag indicates that all subject-related fields including title, keywords, extended keywords and abstract (if present in SCI for that paper) are searched.

³ The search was conducted on on April 5, 2006.

Some methodological considerations about this approach should be noted:

First, the two collected sets of papers are related in the way that all GIScience papers (the 88 papers) are available in the GIS collection (12.144 papers), but not the other way around.

Second, a test of completeness for the collection GIS papers shows that not all papers referring to the 88 GIScience papers are present among the GIS collection. As an example, Table 1 presents the ten most cited GIScience papers in the set listed together with their number of local citations (LCS, references from the GIS papers) versus their global citations (GCS, references from all papers included in the SCI). The local citation frequency rate varies between 50 and 100 per cent, with a median of 70 per cent of the top ten GIScience citations. Although a full inclusion of referring papers would have been preferable, it would only make minor differences on the order of the top cited GIScience papers. As indicated in Table 2, a similar kind of bias is present for the GIS articles.

Table 1: The ten most cited GIScience papers (n=88) in the SCI ordered by local citation score (LCS), together with their corresponding global citation score (GCS). The node number refers to an arbitrary identity number which corresponds to the numbers in Figures 1 and 2.

#	Node / Date / Journal / Author	LCS	GCS
1	187 1992 <i>International Journal of Geographical Information Systems</i> 6 (1): 31-45 Goodchild MF Geographical Information Science	43	53
2	2425 1997 <i>Annals of the Association of American Geographers</i> 87 (2): 346-362 Wright DJ; Goodchild MF; Proctor JD Demystifying the Persistent Ambiguity of GIS as "Tool" versus "Science"	22	25
3	3428 1998 <i>International Journal of Geographical Information Science</i> 12 (7): 651-671 Sui DZ GIS-Based Urban Modelling: Practices, Problems, And Prospects	13	14
4	2701 1997 <i>International Journal of Geographical Information Science</i> 11 (6): 523-528 Wright DJ; Goodchild MF Data from the Deep: Implications for the GIS Community	9	12
5	4217 1999 <i>International Journal of Geographical Information Science</i> 13 (8): 747-774 Mark DM; Freksa C; Hirtle SC; Lloyd R; Tversky B Cognitive Models of Geographical Space	8	14
6	4112 1999 <i>Geographical Analysis</i> 31 (4): 373-399 Miller HJ Potential Contributions of Spatial Analysis to Geographic Information Systems for Transportation (GIS-T)	7	7
7	5101 2000 <i>International Journal of Geographical Information Science</i> 14 (5): 475-496 Ahlqvist O; Keukelaar J; Oukbir K Rough Classification and Accuracy Assessment	7	14
8	6109 2001 <i>International Journal of Geographical Information Science</i> 15 (7): 591-612 Smith B; Mark DM Geographical Categories: an Ontological Investigation	7	14
9	4218 1999 <i>International Journal of Geographical Information Science</i> 13 (8): 775-796	6	9

10	Egenhofer MJ; Glasgow J; Gunther O; Herring JR; Peuquet DJ Progress in Computational Methods for Representing Geographical Concepts 6094 2001 <i>International Journal of Geographical Information Science</i> 15 (3): 201-220 Chen J; Li CM; Li ZL; Gold C A Voronoi-based 9-intersection Model for Spatial Relations	5	8
----	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---	---

Table 2: The top 3 GIS papers.

#	Node / Date / Journal / Author	LCS	GCS
1	4842 2000 <i>Ecological Modelling</i> 135 (2-3): 147-186 Guisan A; Zimmermann NE Predictive Habitat Distribution Models in Ecology	96	300
2	1676 1995 <i>Progress in Physical Geography</i> 19 (4): 474-499 Franklin J Predictive Vegetation Mapping: Geographic Modelling of Biospatial Patterns in Relation to Environmental Gradients	65	135
3	1030 1994 <i>Journal of Applied Meteorology</i> 33 (2): 140-158 Daly C; Neilson RP; Phillips DL A Statistical Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain	64	419

RESULTS

According to ISI classification, the share of research articles among the 50 most cited GIS articles are very high (46 out of 50). Despite this, the two top cited articles are reviews (three in total). Lastly, one article is an editorial piece. Among the research papers, the citation share of methodological or application papers is very high in the collected data. These papers typically describe a specific model or method to measure some variable and apply this measure to a test area. Table 2 shows the top three cited papers in the collection of 12.144 GIS papers. Only at place 12, the first GIScience paper, Michael Goodchild's (1992) *Geographical Information Science* (no. 187 in Figures 1 and 2.) is found. This paper was influential for the renaming of the field, and introduced the term Geographic Information Science as the foundational theoretical and essential discipline underscoring GISystems. Next, at place 74, Dawn J. Wright et al's (1997) paper on GIS as 'Tool or Science' (not shown in Figure 1, no 2425 in Figure 2) is found. This paper set the task of discussing whether GIS was a tool or a science by posting a question to subscribers of the GIS-L e-mail list. The third of the 88 papers including the term GIScience is found at place 212 with a citation score of 13 citations (no. 3428 in Figure 2).

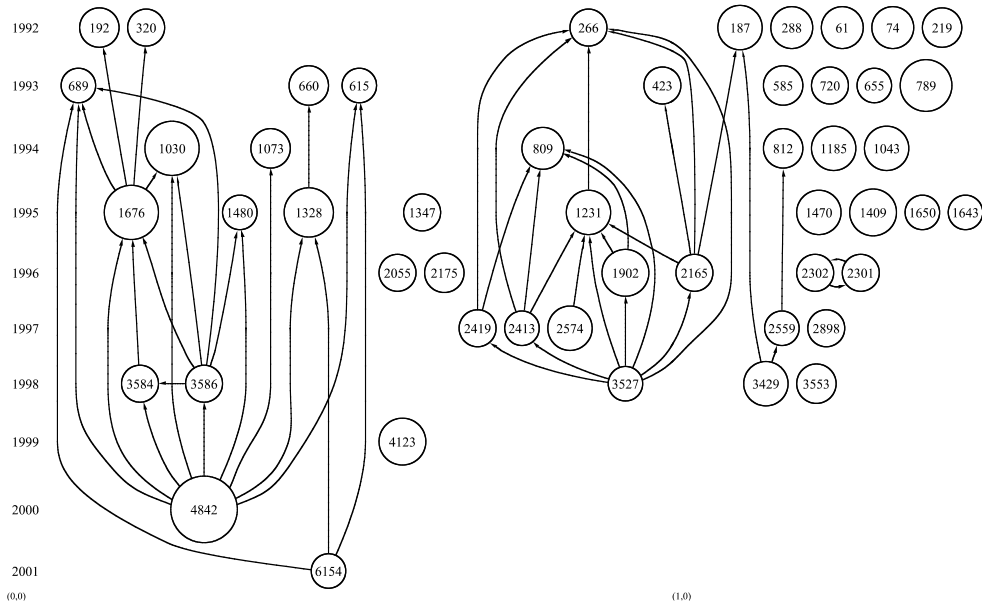


Figure 1. Citation map of the 50 most cited articles containing the term GIS in the collection of 12,144 articles from SCI.

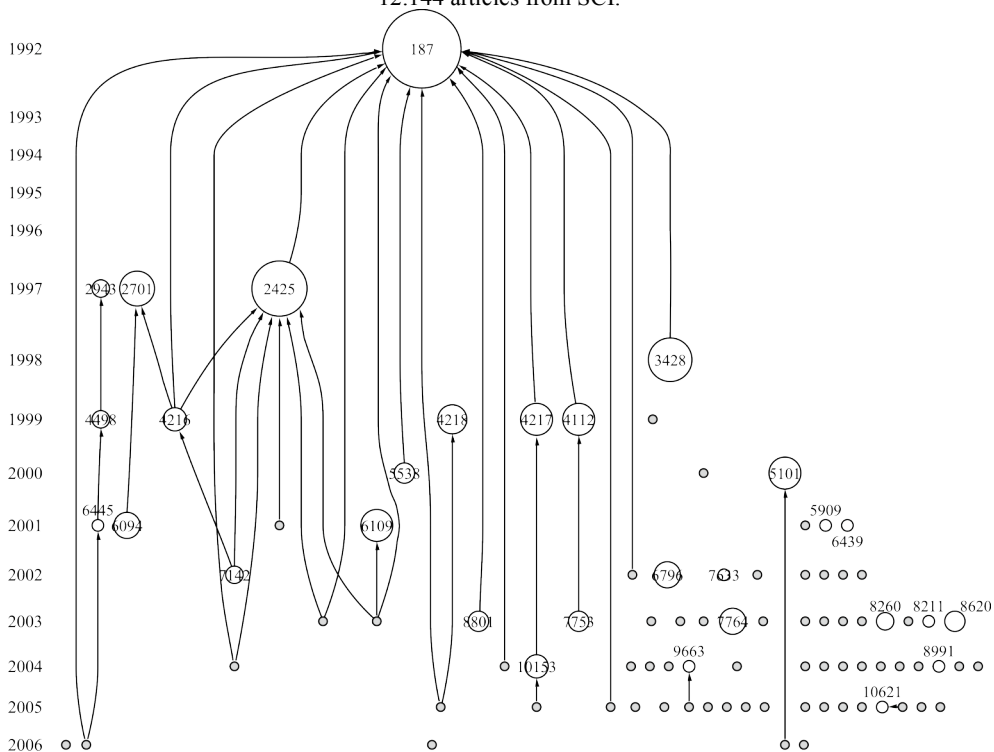


Figure 2: Citation map for the 88 'GIScience' papers found in SCI. Note: node numbers for uncited documents are omitted to improve readability.

In trying to describe the most cited papers in this collection, the citation map is used to analyse the citation patterns between the papers. Each circled node in Figures 1 and 2 corresponds to a specific paper in the database, and the size of the circle corresponds to the number of citations the paper has attracted. Arrows indicate that a reference to a node is given in the originating node. In Figure 1, which shows the citation relations between the 50 most cited GIS papers (n=12.144), there are two distinct clusters of papers. The papers in the left cluster, which contains all the three most cited papers in the collection, are concerned about modeling, remote sensing and GIS in environmental, ecological or wildlife research. The right cluster of papers corresponds to papers in medicine and biology, linking research in epidemiology and tropical medicine to entomology, and in one case in this excerpt of the 50 most cited GIS papers, one reference to Goodchild's 1992 paper. A third group of papers further to the right, beginning with Goodchild's 1992 paper (no. 187) could be described as focused on GIS. While referring to each other only to a very low degree, at least half of these papers contain subjects that could be described as "pure GIS" or for that matter as papers "underscoring" GIS albeit not having the term GIScience in its title/abstract.

Among the GIScience papers, one single article takes the center position. The paper by Goodchild (1992) is referred to by 15 of the GIScience papers, while the second most cited paper (no. 2425, co-authored by Goodchild) is referred to by 6 other GIScience papers.

DISCUSSION

We need to elaborate on the choice of interpreting 'GIScience' as a term synonymous with theoretical content. Primarily, this stems from statements made by GI researchers maintaining a tension between GIS considered as a tool or as a science (Wright et al., 1997).

This tension is expressed in the following way in a fundraising workshop report:

GIScience provides the basic intellectual underpinnings for geographic information technologies, and GIScience research should be supported at levels appropriate to the importance of these technologies and their application (Mark, 1999).

In an undergraduate textbook it is even more normatively stated:

GIScience *should* study the fundamental issues arising from geographic information (Longley et al., 2005) [my italics].

If these statements could be taken to justify that the subject matter of GIScience could be linked to the concept 'theory', I argue that the corresponding relationship between GISystems that is described as a 'tool' or 'technology', could be substituted with 'application' or 'methodological'. This is verified by reference to the Call for papers for the latest GIScience conference: "Since GIScience 2006 focuses on advances in the fundamentals of Geographic Information Science, *submission of GIS application papers is discouraged*" [my italics]⁴.

In this study I have observed a bias in the patterns of citations in GI literature, where the top three GIS papers outnumber the top three GIScience papers by a ratio of 2.88 to 1 (a total of 225 versus 78) citations. Furthermore, I have found that not more than two papers containing the term GIScience make it into the top 200 cited GIS papers in terms of acquired citations from the collection of 12.144 papers.

⁴ Retrieved, 31/8 2006 from <http://www.giscience.org/submissions.php>

The observations found here are in line with other research on citations in scientific literature. In a study by Eugene Garfield, who found a similar pattern of a high share of methodological citations in an analysis of the 50 most cited publications in SCI in 1967 it is noted that:

(...) I do want to point out that many of these particular papers are methodological. In retrospect, one expects that such method papers will be frequently cited, but it comes as a surprise that they predominate so strongly. (Garfield, 1971, pp. 161-162)

CONCLUSION

In this paper it is suggested that researchers in GIS rather turn to good examples of research than to normative theories of research when they refer to work by others. Furthermore, the term GIScience, which has been put forward as a concept linked to the core of GIS research, is shown to have low visibility in the data analyzed.

The title of this study suggests that there is a distinct break between theoretical works and good examples within the study of geographic information. I have also presented examples from the GI literature that maintains that this is the case. Still, one way to evaluate the results of the present study is that this dividing line is a blurred one.

Further questions that need to be pursued are what needs a scientific terminology raises and what role it plays for the research community. From the outset, the term GIScience was used to demarcate the research in GI from mere 'handling' of spatial data (Goodchild, 1992). At the same time it was maintained that "Disciplines are like tribes, with their own totems, symbols and membership rules, languages and social networks" (Ibid., p. 43).

What role does GIScience play for GI researchers?

REFERENCES

- Chrisman, N. (2003). Revisiting fundamental principles of GIS. In D. Kidner, G. Higgs & S. White (Eds.), *Socio-Economic Applications of Geographic Information Science*. London: Taylor & Francis.
- Dorn, R. I. (2002). Analysis of Geomorphology Citations in the Last Quarter of the 20th Century. *Earth Surface Processes and Landforms*, 27, 667-672.
- Elzinga, A. (1995). Traces of Eurocentrism in Current Representations of Science. *VEST Tidskrift för vetenskapsstudier*, 8(4), 85-95.
- Fisher, P. (2001). Citations to the International Journal of Geographical Information Systems and Science: the first 10 years. *International Journal of Geographical Information Science*, 15(1), 1-6.
- Fisher, P., & Unwin, D. J. (Eds.). (2005). *Re-presenting GIS*. Chichester: Wiley.
- Garfield, E. (1971). Citation indexing, historio-bibliography, and the sociology of science. *Current Contents*(6), 158-174.
- Garfield, E., Pudovkin, A. I., & Istomin, V. I. (2003). Mapping the Output of Topical Searches in the *Web of Knowledge* and the Case of Watson-Crick. *Information Technology and Libraries*, 22(4), 183-187.
- Goodchild, M. F. (1992). Geographical Information Science. *International Journal of Geographical Information Systems*, 6(1), 31-45.
- Longley, P. A., Maguire, D. J., Rhind, D. W., & Goodchild, M. F. (2005). *Geographic information systems and science* (2. ed.). Chichester: Wiley.

- Mark, D. M. (1999). Geographic Information Science: Critical Issues in an Emerging Cross-Disciplinary Research Domain [Electronic Version]. *Workshop on Geographic Information Science and Geospatial Activities at NSF, January 14-15, 1999*, 18. Retrieved January 18, 2007 from <http://www.geog.buffalo.edu/ncgia/GIScienceReport.pdf>.
- McMaster, R. B., & Usery, E. L. (2004). *A research agenda for geographic information science*. Boca Raton, FL: CRC Press.
- Monastersky, R. (2005). The Number that's Devouring Science. *Chronicle of Higher Education*, 52(8), 12-17.
- Wright, D. J., Goodchild, M. F., & Proctor, J. D. (1997). Demystifying the Persistent Ambiguity of GIS as "Tool" Versus "Science". *Annals of the Association of American Geographers*, 87(2), 346-362.