

Journal Maps on the Basis of *Scopus* Data: A Comparison with the *Journal Citation Reports* of the ISI

Loet Leydesdorff

Amsterdam School of Communications Research (ASCoR), University of Amsterdam, Kloveniersburgwal 48, 1012 CX Amsterdam, The Netherlands. E-mail: loet@leydesdorff.net; <http://www.leydesdorff.net>

Félix de Moya-Anegón

CSIC, CCHS, IPP, SCImago Research Group, Spain. E-mail: felix.moya@scimago.es

Vicente P. Guerrero-Bote

Universidad de Extremadura, Information and Communication Science Department, Spain. E-mail: guerrero@unex.es

Using the *Scopus* dataset (1996–2007) a grand matrix of aggregated journal-journal citations was constructed. This matrix can be compared in terms of the network structures with the matrix contained in the *Journal Citation Reports (JCR)* of the Institute of Scientific Information (ISI). Because the *Scopus* database contains a larger number of journals and covers the humanities, one would expect richer maps. However, the matrix is in this case sparser than in the case of the ISI data. This is because of (a) the larger number of journals covered by *Scopus* and (b) the historical record of citations older than 10 years contained in the ISI database. When the data is highly structured, as in the case of large journals, the maps are comparable, although one may have to vary a threshold (because of the differences in densities). In the case of interdisciplinary journals and journals in the social sciences and humanities, the new database does not add a lot to what is possible with the ISI databases.

Introduction

Derek de Solla Price's (1961) fascination with the growth of journal literature during more than three centuries was followed by both Garfield and Narin as early pioneers of the bibliometric enterprise. Narin et al. (1972) elaborated on Price's (1965) suggestion to use aggregated journal-journal citations for multivariate analysis to distinguish groups of journals representing specialties, and Garfield (1972) developed citation analysis as a tool for journal evaluation. Ever since, a large number of indicators for journals has been developed, and science mapping on the basis of aggregated journal-journal citations has become an industry (e.g., Rosvall & Bergstrom,

2008; Klavans & Boyack, 2009a; Leydesdorff & Rafols, 2009; De Moya-Anegón et al., 2008).

Hitherto, maps and indicators have been largely confined to the domain of the journals included in the *Science Citation Index (SCI)* and the *Social Science Citation Index (SSCI)* of the Institute of Scientific Information (ISI of Thomson-Reuters). In 2004, both *Scopus* and *Google Scholar* were launched, providing alternatives to the ISI databases for citation analysis. Although one could, in principle, construct an aggregated journal-journal citation matrix from the *Google* database, this may not be an easy task (Leydesdorff, 1994). Unlike the *SCI* and *Scopus*, *Google Scholar* is not delimited in terms of journals processed at the backend, but on searching the Web for scholarly contributions. Journal names in this case are only attributes to documents and are not controlled for variants and misspellings.

Both the *SCI* and the *Scopus* are based on a delineation in terms of scholarly journals. The contents of these so-called "source journals" are processed bibliographically in terms of publications and the references provided by the authors of these publications. These references include citations to "nonsource" journals, books, reports, patents, etc. Both Elsevier and the ISI identify a number of journals as additional source journals when they are cited without including them in terms of the citing journals. In the *Scopus* database, journals are identified with a unique item identifier which is kept constant over the years, whereas in the (*Social*) *SCI*, the titles of the journals (and their corresponding abbreviations) function as identifiers.

The ISI aggregates the aggregated citation data among journals in the yearly *Journal Citation Reports (JCR)*. *JCRs* have been available since 1974 and 1977 for the *SCI* and the *SSCI*, respectively. Electronic versions on CD-Rom have been available for these two databases since 1994; since

Received June 15, 2009; revised August 29, 2009; accepted September 14, 2009

© 2009 ASIS&T • Published online 19 November 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.21250

1997, the data are also available online. At the time of this project, the latest version (2007) of the two ISI databases contained 7,940 journals. Note that there is no *JCR* available for the *Arts & Humanities Citation Index (A&HCI)*.

Hitherto, Elsevier has not published the aggregated journal-journal citations among journals contained in the *Scopus* database. In 2007, this database contained 23,503 journals with an item identifier; that is, approximately three times as many as the source journals in the two *JCRs* combined (7,940 in 2007).¹ However, only 14,964 (63.7%) of these journals were processed by *Scopus* in 2007. The ISI processed 7,925 (99.8%) of the source journals. In other words, the *Scopus* references draw upon a much larger pool than the ones included in the *JCRs*.

The *Scimago* group in Spain (at <http://www.scimago.es>) has made the *Scopus* (1996–2007) data available for scientometric analysis. These data contain the full set provided with an item identifier since 1996 (that is, 23,503 journals). In principle, a set of the cited or citing journals for each year can be derived from this database by using a dedicated routine. We used this set containing aggregated citations among all journals with item identifiers during the period 1996–2007.

The unique identification in the *Scopus* database is limited backwards to 1996, while the ISI database includes all references provided in a certain year. This includes both references to nonsource journals and to sources of older dates in the case of the ISI set, while the *Scopus* dataset used in this study contains only references to the source journals identified as items during the period 1996–2007.² However, the ISI database allows us additionally to distinguish between the citations to sources in the last 10 years (1998–2007) and the rest. We shall test below for the effects of this disaggregation of the ISI database and compare the results with those obtained using *Scopus* data covering the period 1996–2007.

The *Scopus* data allow us to ask questions different from those addressed to the ISI data. For example, the arts and humanities are included, and the extension may include many more journals in languages other than English. The two databases can also be compared with other bibliometric characteristics such as citation patterns or the *h*-indices of authors (Bar-Ilan, 2008; Jacso, 2005, 2008; Meho & Rogers, 2008; Meho & Yang, 2007). Focusing on a subset of oncology journals, López-Illescas, de Moya-Anegón, and Moed (2008) found that with reference to this set, the ISI database could be considered as a subset of the *Scopus* database. Although the coverage of the latter database was 90% larger, the average impact factor in the latter set was only 2.6% higher than in the ISI set.

Klavans & Boyack (2007) compared maps based on citation analysis at the document level in both databases. Klavans & Boyack (2009a) argue that a consensus map of science seems possible based on communalities between the two databases (cf. Rafols & Leydesdorff, 2009). Recently (July 2009), Elsevier launched a facility *SciVal* (at <http://info.funding.scival.com/>) based on a further extension of Klavans and Boyack's (2009b) cocitation maps of the *Scopus* data.

We pursue the analysis at the level of journals. Our original research question was to investigate comparatively the informative quality of journal maps on the basis of data in the two databases from a user perspective. Raising questions about the differences, however, brought us to a (second) analytical question about the different qualities of the data for this purpose. Can the differences also be explained? As noted, a major difference is the cutoff in the *Scopus* data in 1996: data older than 1996 are included, but not systematically marked with item identifiers and, therefore, not usable for our purpose without considerable data cleaning. The expectation is that the data of *Scopus* (1996–2007) are richer than the *SCI* (1998–2007) because of the larger coverage, but potentially less informative than the full set of the *SCI* because of the cutoff of the historical tails of the distributions.

The two databases are also different in size. There has been continuous debate about how the ISI selects journals for inclusion as source journals in the database (Garfield, 1979, 1990; Testa, 1997). Sivertsen (2003) found no bias in the coverage of Scandinavian publications by the *SCI*. However, the ISI concedes that the selection system for inclusion in the *SCI* may be biased against journals written in non-Latin alphabets. Special care has been taken in the past to include Russian and Japanese journals into the database (Garfield, 1979, 1998). *Scopus* brings a document online specifying its criteria for inclusion of journals and other content into the database at http://info.scopus.com/docs/content_coverage.pdf.

As reported by Thomson-Reuters Scientific in May 2008,³ the *SCI* was recently extended with another 700 journals to improve its “regional” coverage. Leydesdorff and Wagner (2009) found a notable effect of this extension on the coverage of Eastern European contributions, but less so for other world regions. This recent expansion seems to correspond to the larger coverage by the *Scopus* database in these linguistic domains. In other words, the two databases are competing with each other on the market of academic libraries. In June 2009, the *Scopus* team announced to increase the coverage of journals in the arts and humanities to 3,500 by adding all (approximately 1,450) top global journals using the European Reference Index for the Humanities (ERIH) of the European Science Foundation.⁴ These recent extensions are not included in our data; the analyses are based on the tape years 2007 for both datasets.

Language Coverage of the Journals in the two Databases

A major advantage of the larger coverage of the *Scopus* data in terms of journals included could be the coverage of journals in languages other than English. *Scopus* advertizes on its Web site to “cover titles from all geographical regions including non-English titles as long as English abstracts can be provided.”⁵ As noted, the lack of coverage in the non-English domain has been a major complaint about the ISI databases (Maricic, 1997; Van Leeuwen, Moed, Tijssen, Visser, & Van Raan, 2001). This can particularly be

TABLE 1. Distribution of journals in the ISI and *Scopus* databases in terms of languages.

	ISI		<i>Scopus</i> 1996–2007		<i>Scopus</i> Aug. 2009	
		%		%		%
English	6,912	87.1	14,109	81.1	14,193	80.7
Multilanguage	802	10.1	2,993	17.2	1,227	7.6
German	88	1.1	44	0.3	363	2.1
French	45	0.6	67	0.4	292	1.7
Spanish	33	0.4	70	0.4	360	2.1
Chinese	17	0.2	0	0.0	323	1.8
Japanese	13	0.2	0	0.0	126	0.7
Russian	13	0.2	1	0.0	36	0.2
Portuguese	5	0.1	19	0.1	110	0.6
Italian	2	0.0	23	0.1	134	0.8
Rumanian	2	0.0	1	0.0	7	0.0
Croatian	1	0.0	4	0.0	33	0.2
Dutch	1	0.0	1	0.0	30	0.2
Eskimo	1	0.0	0	0.0	0	0.0
Korean	1	0.0	0	0.0	15	0.1
Polish	1	0.0	21	0.1	69	0.4
Slovene	1	0.0	0	0.0	6	0.0
Swedish	1	0.0	0	0.0	6	0.0
Turkish	1	0.0	13	0.1	44	0.3
Hungarian	0	0	10	0.1	24	0.1
Czech	0	0	2	0.0	30	0.2
Slovakian	0	0	2	0.0	6	0.0
Lituanian	0	0	1	0.0	1	0.0
Danish	0	0	1	0.01	1	0.0
Iceland	0	0	1	0.01	0	0.0
other	0	0	0	0	158	1.0
Total	7,940		17,383		17,594	

Note. The ISI distribution is based on the set in 2007; the *Scopus* database on the journals included from 1996 to 2007.

considered as a problem for the evaluation of research performance in the social sciences and the humanities (Glänzel & Schöpfli, 1999; Nederhof, Zwaan, Bruin, & Dekker, 2001; Nederhof, 2006).

For example, the French Centre National de la Recherche Scientifique (National Center for Scientific Research) subsidizes approximately 225 journals that are only partially covered by the United States-based *Science Citation Indices* (De Looze, Cronini, Legentil, Jeannin, & Magri, 1996; Legentil, personal communication). Among the 4000+ journals published in Chinese, only 17 were included in the ISI database in 2007 (Ren & Rousseau, 2002; Jin, Jiangong, Chen, & Zhu, 2002; Leydesdorff & Jin, 2005; Moed, 2002). Eighty journals in the *JCRs 2007* were published in China (but not necessarily in Chinese). Of the 7,940 journals included in the ISI database, 6,912 (87.1%) are published in English. Another 10.1% of the journals are multilingual, but mostly include English as a main language. Journals in other languages are marginal to the database (Table 1).

Table 1 informs us that this promise of a wider coverage of languages by the *Scopus* database can be nuanced. According to Table 1, the *Scopus* 1996–2007 database would not contain Chinese, Japanese, or Korean journals in 2007. Because we aggregated the journal data from the documents, some

journals in these Asian languages are included in the multilanguage category. The extension in the French and German domain seems less substantial than in some of the other (mainly European) languages. For this reason, we checked with Elsevier (Ove Kähler at Scopus, personal communication, August 28, 2009) who was so kind to provide us with data for the fifth and sixth column, including the breakdown of the multilanguage category as registered by *Scopus* at the journal level. Because the totals are approximately the same, one may also assume that the language distribution in terms of percentages is approximately the same in our data.

Methods and Materials

The aggregated journal-journal citation data of the ISI were harvested from the CD-Rom version of the *Journal Citation Reports 2007*. These data were identical to the online version of the *JCR*. However, we combined the two databases for the *SCI* and the *SSCI* into a single *JCR*. A similar matrix was constructed using the *Scopus* database on the basis of the so-called item identities in both the cited and the citing items. Because these item identities are unique, the *Scopus* database does not contain errors. Misspellings of journal names, however, are included in the *SCI* as references to non-source journals. For example, Leydesdorff (2008, Table 4, at p. 285) found 14 so-called “nonsource journals” in the *SCI* based on common misspellings of journal names with values of more than 10,000 citations in 2005.

Both datasets were brought under the control of a relational database management system and exported to *Pajek*, a free-ware program for the visualization and analysis of networks.⁶ The resulting files are asymmetrical matrices which can be analyzed in both the cited and the citing dimensions in terms of journal citation patterns. We use SPSS version 15 for the factor analysis and focus on the cited dimension unless stated otherwise. For the comparisons among the three sets (*Scopus* 1996–2007, *SCI* 1998–2007, and the complete *SCI*), we extract ego-networks of seed journals using dedicated software. All journals contributing 1% or more to the being-cited totals of citations of the seed journal in 2007 are included in these ego-networks. We use the same sets of journals when comparing the original ISI data with the truncated set *SCI* (1998–2007) to facilitate the comparison and for the technical reason that we have no easy hold in the data for adjusting the chosen thresholds of one percent in the reduced sets. In all cases, the cosine among the citation patterns is used for normalization (Ahlgren, Jarneving, & Rousseau, 2003) with a threshold of cosine > 0.2. We keep the procedures as constant as possible across the different journal comparisons between the databases.

In the visualizations of the cosine matrices (using *Pajek*), the sizes of the nodes are vertically proportionate to the logarithm of the number of citations, and horizontally to the logarithm of this number after correction for “within journal self-citations.” The width of the links is proportionate to the cosine value (above the threshold of cosine > 0.2). Colors were automatically generated using the k-core algorithm in

TABLE 2. Summary statistics of the three databases under study.

	<i>Scopus</i> 2007 (1996–2007)	<i>SCI</i>	<i>SCI</i> (1998–2007)
a. Number of source journals	23,503 ^a	7,940	7,900
b. Number of source journals (citing)	14,964	7,925	7,793
c. unique journal-journal relations among source journals	3,570,020	1,460,847	1,318,180 (= 90.23%)
d. Total “cited”	22,370,409	27,477,499	
e. Total “citing” (source journals)	28,492,874 ^b	24,979,391	16,549,053
g. Sum journal-journal relations	22,370,409	24,979,391	16,549,053 (= 66.25%)
h. Average non-zero cell value (= g/c)	6.27	17.10	12.55 (= 73.39%)
i. Density	0.0065	0.0232	0.0209
j. Largest component	13,376 (89.4%)	7,694 (96.9%)	7,674 (97.1%)
k. Betweenness centralization	0.05198	0.08354	0.08628
l. Average Clustering Coefficient	0.2700429 ± 0.1719204	0.3478451 ± 0.1255273	0.3460075 ± 0.1289489

Note. ^aIn the case of the *Scopus* data, this is the number of journals with an item identity in the period 1996–2007. Of these journals, 22,238 were processed in the cited direction. Journals may be included in the database in one or more years without necessarily presence during the whole period under study.

^bThis number was kindly provided to us by one of the referees. Our analysis of the *Scopus* data aggregates only identified item-item relations without further attempt to resolve the cited journal name to keys which were not already provided in the data.

Pajek. The factor structures are indicated in the figures using ellipses with dotted lines; factor designations were provided by the analysts. The number of factors is reasoned on the basis of screeplots in the respective analyses.

Descriptive Statistics

Table 2 provides some descriptive statistics of the three citation matrices (*Scopus* 1996–2007, *SCI*, and *SCI* 1998–2007). In addition to the citation indicators, we computed betweenness centralization and the average clustering coefficients as network measures because these two indicators teach us about the coherence and the decomposability of the databases, respectively.⁷

More detailed analysis of the data revealed that only 14,964 journals of the 23,503 journals with item identifiers included in *Scopus* (1996–2007) are processed as source journals (citing). (Only 13,699 were processed as both cited and citing in 2007.) In the case of the ISI set, the “all others” category leads to discrepancies between “total cited” (d.) and “total citing” (e.). The ISI database also contains a large number (292,105) of non-ISI source journal names among the references, which are, in principle, available for further analysis. The huge number of nonsource items in *Scopus* (2,664,407)⁸ was not available for analysis. This study focuses on the 24,979,391 unique journal-journal relations among the 7,940 journals in the case of the ISI and 22,370,409 such relations in the case of *Scopus* (boldfaced as line g. in Table 2).

The descriptive statistics for the *SCI* (1998–2007) subset are added in the third column of Table 2. The aggregated number of references to ISI source publications of 1997 or older is 8,430,338.⁹ This amounts to 33.75% of the total of 24,979,391 references included in the analysis. However, 90.23% of the unique journal-journal relations (c.) in the full set are also covered by the reduced set. The density (i.) of the reduced set consequently is not much lower than that

of the full set, but the average nonzero cell value (h.) is. This average may be low in the *Scopus* database because of the inclusion of journals with lower citation rates.

The numbers of journal-journal relations are of the same order in both databases, but the number of journals is larger in the *Scopus* database than in the two *Science Citation Indices* combined. Accordingly, the density is larger in the ISI database. In other words, the aggregated journal-journal citation matrix based on the *Scopus* database contains many more sparse areas than the ISI database. One would expect that this difference is caused by the limitation of the *Scopus* database—reaching back only to 1996—and the larger number of journals covered. Cutting off the historical tails of the distributions in the *SCI*, however, does not have a large effect on the density which remains in the *SCI* (1998–2007). This latter density is higher than in the *Scopus* (1996–2007) database.

One should keep in mind that the matrix of citations among journals in the case of *Scopus* (1996–2007) is based on aggregation of citations among identified documents. While the ISI database covers the source journals cover-to-cover—although it uses only the so-called citable issues (articles, reviews, letters, and proceedings papers) as the denominator of the impact factor—*Scopus*, for example, does not process cited references in book reviews (Ove Kähler at *Scopus*, personal communication, August 28, 2009). In other words, the different organization of the *Scopus* database may have a large effect that can be expected to vary across fields of science (Price, 1970).

For example, for the *SCI*, the percentage of citations in the historical tail (1997 or older) is 33.16%, while it is 44.41% for the *SSCI*.¹⁰ All network measures are lower in the *Scopus* database than in the ISI database, but for the latter database, it makes not so much difference in terms of the values of the network measures whether one includes the historical tails or not. The structural properties of the database are preserved when focusing only on the later years.

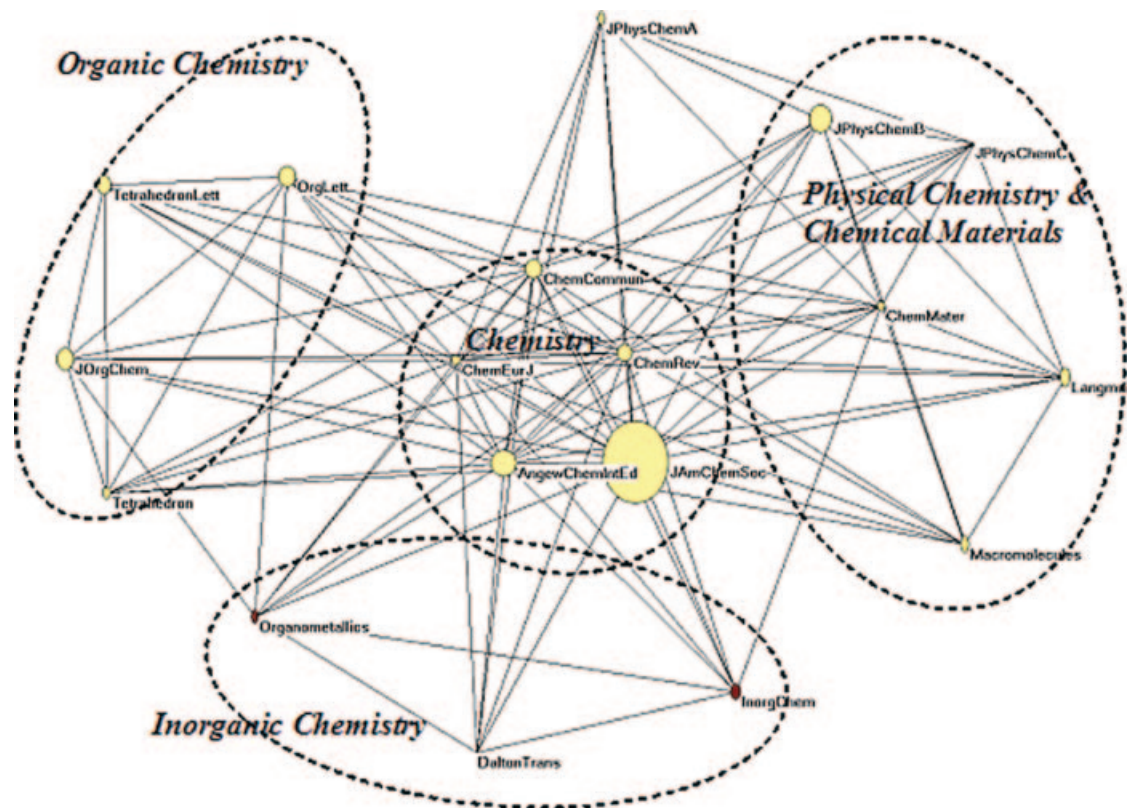


FIG. 1. The citation impact environment of the *Journal of the American Chemical Society* in Scopus; $N = 18$; threshold 1%; cosine > 0.2 .

Mapping of Journal Structures

One would not expect the two databases to be very different when the focus is on established fields of science and major journals. To test for this baseline, we used two major journals: the *Journal of the American Chemical Society* (JACS) and this journal itself—that is, the *Journal of the American Society for Information Science and Technology* JASIST—and compare the resulting visualizations using the parameters as specified above. Both these journals are major journals in their respective disciplines and rich sources of citations. JACS had an impact factor of 7.885 in 2007. One of us has analyzed the structure of this journal's citation environment in other studies (Leydesdorff, 1991; Leydesdorff & Bensman, 2006). JASIST had an impact factor of 1.436 in 2007. The resulting maps of the citation environments of this journal may be easier to compare (and validate) for the readership.

The Citation Impact Environment of the JACS

Figures 1 and 2 show the citation impact environments of the JACS in the Scopus database and the ISI database, respectively. In both cases, 18 journals are drawn into the analysis. However, the two lists of journals are not identical.

Some differences between the two maps are generated because journals are sometimes included in the one database but not in the other. For example, the *Journal of Physical Chemistry C* is not included in the ISI database; the *Journal*

of *Organometallic Chemistry* does not pass the threshold (of 1,831 citations) in *Scopus* by contributing only 1,512 to the citations of the JACS.

In both Figures 1 and 2, the main structures can easily be recognized, but these representations differ significantly in the details. Factor analysis of both citation matrices provides a similar structure in terms of four factors explaining 72.9% of the variance in the case of *Scopus* and 79.1% in the case of the ISI. In the solution for the ISI database, *Langmuir* and *Macromolecules* are more clearly distinguished from the *Journal of Physical Chemistry A and B*, while the cases of *Scopus*, the *Journal of Physical Chemistry C*, and *Chemical Materials* were drawn into this component.

Because other journals are included in both ego-networks, 1% of the total number of citations will also be different in the two databases. One would expect this number to be higher in the journal-journal citation data based on *Scopus* than in the one based on ISI database because more journals are included. However, the total number of citations to JACS is lower (183,186) in the journal-journal citation matrix based on the *Scopus* database, and the total number of references in JACS is 77,778, including 14,985 self-citations. These numbers are 295,465 for citations to JACS (+61.3%), 120,506 references from articles in JACS (+54.9%), and 20,706 self-citations (+38.2%) in the journal-journal citation data based on the ISI database (Table 3). The citation frequencies among journals are also different between the two journal-journal citation matrices. The network in Figure 1 is accordingly considerably less dense than in Figure 2.

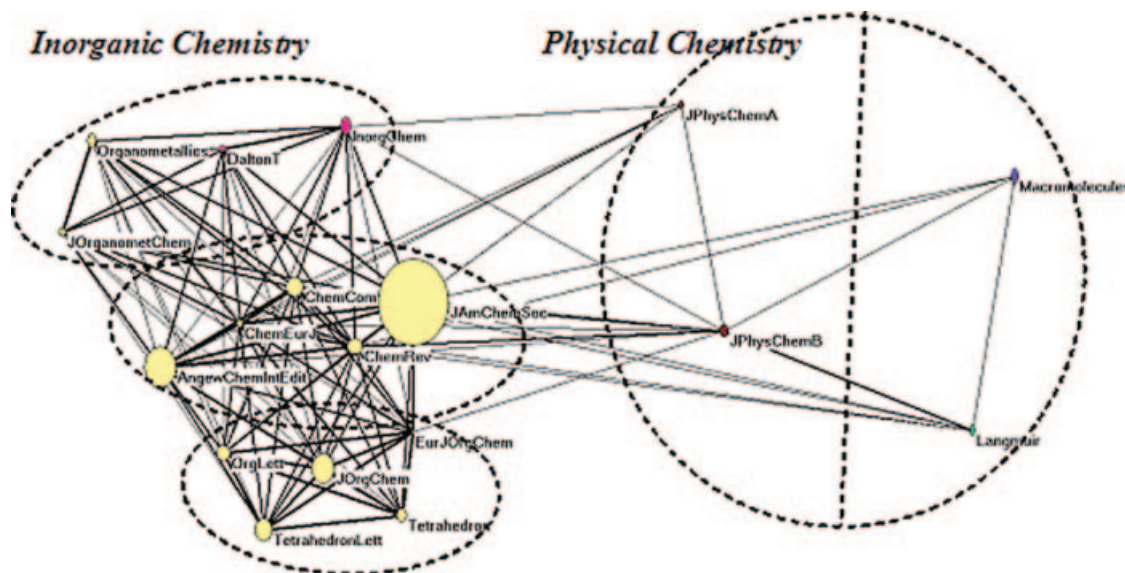


FIG. 2. The citation impact environment of the *Journal of the American Chemical Society* in the *Science Citation Index*; $N = 18$; threshold 1%; cosine > 0.2 .

TABLE 3. Comparison of cited and citing patterns of *JACS* in the three sets.

<i>JACS</i>	<i>Scopus</i> (1996–2007)	<i>SCI</i> (1998–2007)	<i>SCI</i>
Total cited in ego network	183,186	73,414	295,465
Total citing in ego network	77,778	38,192	130,506
Self-citations	14,985	14,725	20,706

Note. *JACS* = *Journal of the American Chemical Society*; *SCI* = *Science Citation Index*.

As noted, the ISI database allows us to select data only for the last 10 years. The number of within journal self-citations for *JACS* decreases to 14,725 of the 20,706 citations in the unrestricted case (that is, 71.1%). This is very close to the 14,985 self-citations found in the *Scopus* database. Comparison among the columns of Table 3 positions the *Scopus* (1996–2007) set between the *SCI* set and the reduced *SCI* (1998–2007) as was to be expected.

If we apply the same set of 18 journals used in the full analysis to the ISI data for 1998–2007, the structure of the data remains virtually the same (Figure 3). The two citation matrices (with and without the references to publications older than 10 years) correlate with $r = 0.978$ ($p < 0.001$). Structure is also maintained: The correlation between the four-factor matrices with and without the older data is $r = 0.993$ ($p < 0.001$). Thus, these ISI data are not structurally affected by the older data, although they provide a substantial part (29.8% in this case) of the citations. However, the organization of the factor structure is different in the case of using the *Scopus* database (Tables 4 and 5, respectively). Both the order of the factors and the rank-orders of the variable loadings within the components are different.

Both the figures and the structural components are rather different between the two databases, while they are not significantly different between the *SCI* and the reduced set of *SCI* (1998–2007). The comparison with *Scopus* (1996–2007), however, does not allow us to draw any inferences about which one of the two representations to choose.

JASIST

In the context of this study, *JASIST* provides us with an interesting case because one would expect this journal to have suffered from its name change from *JASIS* to *JASIST* in 2001, and perhaps from the previous name change from *American Documentation* into *JASIS* in 1969. In the case of a name change, the historical record may not be carried over to the new name (Bensman & Leydesdorff, 2009).

Table 6 shows first that the name change of *JASIS* into *JASIST* affects the *Scopus* data, but not the ISI data.¹¹ The older name of *American Documentation* can still be retrieved in the ISI database as a nonsource item with 65 references in 2007. Although the difference between the two totals for “cited” can be explained in terms of the cutoff in the *Scopus* database before 1996, the difference in “total citing” is huge. These citations are provided by authors of articles published in *JASIST* during 2007: 315 unique journal-journal relations contain 3,663 references in the *SCI* 1998–2007 set against 575 unique journal-journal relations which carry only 2,843 references in the *Scopus* 1996–2007 set. The lower number may be partly due to the relative large share of book reviews in the *JASIS* previously (Ove Kähler, personal communication, August 28, 2009).

The number of self-citations may vary because of differences in the respective definitions of citable issues and

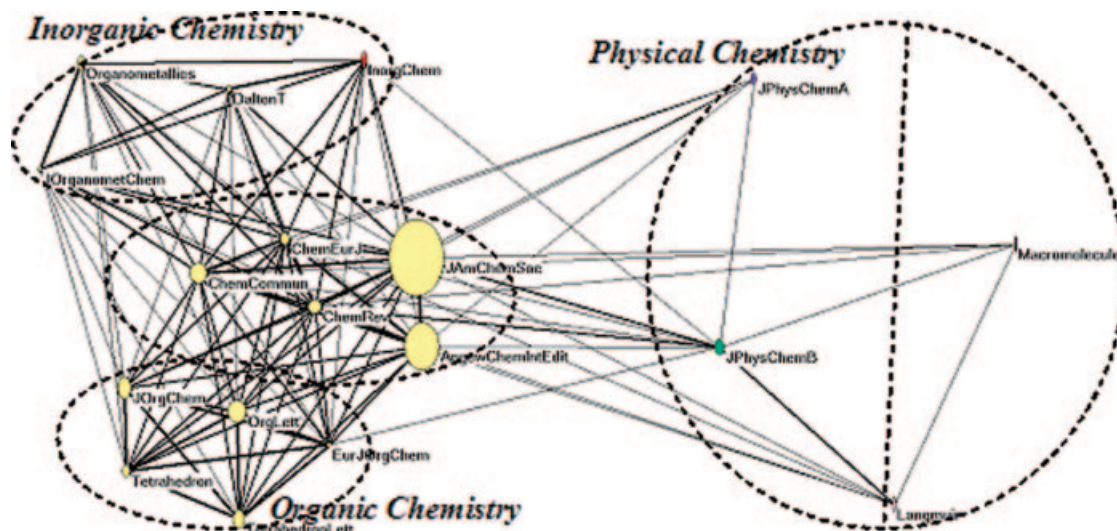


FIG. 3. The citation impact environment of the *Journal of the American Chemical Society* in the *Science Citation Index* (1998–2007); $N = 18$; cosine > 0.2 .

TABLE 4. Factor matrix for the citation patterns among 18 chemistry journals in the citation impact environment of *JACS* in 2007 using the *SCI* (1998–2007).

	Component			
	1	2	3	4
<i>Tetrahedron</i>	0.938			
<i>Eur J Org Chem</i>	0.925	0.179	−0.117	
<i>J Org Chem</i>	0.922	0.192	−0.122	
<i>Tetrahedron Lett</i>	0.916			
<i>Org Lett</i>	0.891	0.302	−0.135	
<i>J Am Chem Soc</i>		0.898	−0.135	
<i>Chem Rev</i>	0.202	0.879		0.169
<i>Chem-Eur J</i>	0.150	0.864	0.273	
<i>Angew Chem Int Edit</i>	0.137	0.787		−0.108
<i>Chem Commun</i>	0.154	0.785	0.363	−0.199
<i>J Organomet Chem</i>			0.832	
<i>Dalton T</i>	−0.377	0.238	0.807	
<i>Organometallics</i>	−0.204		0.761	
<i>Inorg Chem</i>	−0.360	0.388	0.582	
<i>J Phys Chem A</i>	−0.244		−0.225	0.892
<i>J Phys Chem B</i>	−0.514		−0.595	0.227
<i>Langmuir</i>	−0.441		−0.543	−0.342
<i>Macromolecules</i>	−0.351	−0.213	−0.369	−0.461

Note. Four factors explain 77.6% of the variance.

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

JACS = *Journal of the American Chemical Society*; *SCI* = *Science Citation Index*.

^aRotation converged in 6 iterations.

the time spans under consideration. The number of self-citations in the ISI database during the last 10 years was 588 (of the 868 self-citations). The number of citations in *JASIST* 2007 to *JASIS* in the *Scopus* database is 133. Thus, we have to compare the 588 self-citations during the period 1998–2007 in the ISI database with 566 (= 430 + 133) self-citations to the set of both *JASIST* and *JASIS* during the period 1996–2007 in the *Scopus* database. Given this match, the lower number of references (“citing”) in the *Scopus*

TABLE 5. Factor matrix for the citation patterns among 18 chemistry journals in the citation impact environment of *JACS* in 2007 using the *Scopus* dataset, 1996–2007.

	Component			
	1	2	3	4
<i>Chem Rev</i>	0.915	0.237	−0.103	−0.152
<i>J Am Chem Soc</i>	0.876		0.228	−0.187
<i>Chem Eur J</i>	0.800		−0.320	
<i>Chem Commun</i>	0.782		−0.189	0.468
<i>Angew Chem Int Ed</i>	0.752	0.277		0.126
<i>Tetrahedron</i>	0.148	0.912	−0.235	
<i>Tetrahedron Lett</i>	0.114	0.900	−0.236	
<i>J Org Chem</i>	0.271	0.896	−0.235	
<i>Org Lett</i>	0.346	0.872	−0.221	
<i>J Phys Chem B</i>		−0.196	0.853	−0.241
<i>J Phys Chem C</i>		−0.141	0.849	−0.117
<i>Chem Mater</i>	−0.125	−0.263	0.807	0.148
<i>Langmuir</i>	−0.269	−0.137	0.664	
<i>Macromolecules</i>	−0.406	−0.132	0.122	
<i>J Phys Chem A</i>		−0.237		−0.877
<i>Dalton Trans</i>		−0.290	−0.303	0.500
<i>Inorg Chem</i>	0.486	−0.515	−0.231	0.301
<i>Organometallics</i>	0.345	−0.377	−0.372	0.118

Note. Four factors explain 72.9% of the variance.

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

^aRotation converged in 6 iterations.

database, must find its origin in the differences between the journal sets in the relevant citation environments of *JASIST*.

The two maps using the full sets (Figures 4 and 5) are very different. Note that we have adjusted the threshold in the ISI case in proportion to the different numbers of total citations because of the name change in 2001 (in accordance with Table 6). With this threshold 18 journals are selected in the ISI environment as relevant. In the *Scopus* database,

TABLE 6. Citations and self-citations of *JASIST* in the *Scopus* and ISI databases, respectively.

	<i>Scopus</i> 1996–2007	ISI 1998–2007	ISI
Total cited	2091 (402)	2595 (165)	3966 ^a (189)
<i>JASIS</i>	686 (232)	0	0
<i>Am Doc</i>	0	0	65 (9)
Total citing	2,843 (575)	3,663 (315)	5,798 (385)
Self-citations	430	588	868

Note. Italic numbers between brackets indicate the number of unique journal-journal citation relations involved.

^aThe ISI provides a number of 3,026 for total cited in 2007 (in both the *SCI* and the *Social SCI*). This would include the category “All others” and must be erroneous since Total Cites should be larger than the sum of the citations in the matrix (excluding the category “All others”). The separate sums for the *SCI* and the *Social SCI* are 1,927 and 2,039, respectively.

16 journals are selected. These include a second set of computer science journals (Figure 4), while the ISI database provides us with a representation of library and information science as a relatively closed set of journals despite the lower threshold level (Figure 5).

We checked again against the matrix based on the same set of journals in the ISI database, but using citation data only for the last 10 years (1998–2007). Although this reduces the grand sum of the matrix from 6,075 to 4,233 (69.7%), the two matrices are again virtually identical ($r = 0.995$; $p < 0.001$). Figure 6 shows the map for the *SCI* 1998–2007. The two figures are marginally different. For example, *Scientometrics* is more distanced from the set of information-science journals if older citations are not taken into account. As noted by Bensman and Leydesdorff (2009), the citation structure

in the ISI database is not significantly affected by discarding the historical record. The citation structures are reproduced in these data yearly and can, hence, be considered as robust.

In summary, the networks in the cores of specialties are connected in the ISI database more strongly than in the journal-journal citation matrix based on the *Scopus* database. This is not because of the wider casting of the network in the latter case, but because of an erosion of structure as a consequence of lower average numbers of citations (cf. López-Illasca et al., 2008). Paradoxically, the maps from the *Scopus* database may appear more informative than the ones from the ISI database. The presence of more relevant environments provides more information in these representations. In using the ISI data, this same effect should be achievable by using a lower threshold. In other words, the ISI data are richer in structure than the *Scopus* data in this case. This is *not* caused by the historical record before 1996, which is available in the *JCRs*, but not in the *Scopus* (1996–2007) data. This historical record is large, but not important for the mapping.

Extension to the Humanities

One obvious shortcoming of the ISI databases is the limitation of the *JCRs* to the sciences and the social sciences. The *Arts & Humanities Citation Index (A&HCI)* does not include a *JCR*. One of the promises of the *Scopus* database is the inclusion of journals from the humanities. As noted, the recent announcement of the *Scopus* team to increase the coverage of journals in the arts and humanities to 3,500, by adding approximately 1,450 top global journals using the ERIH of the European Science Foundation, is not covered by the current study which uses *Scopus* data until 2007.¹²

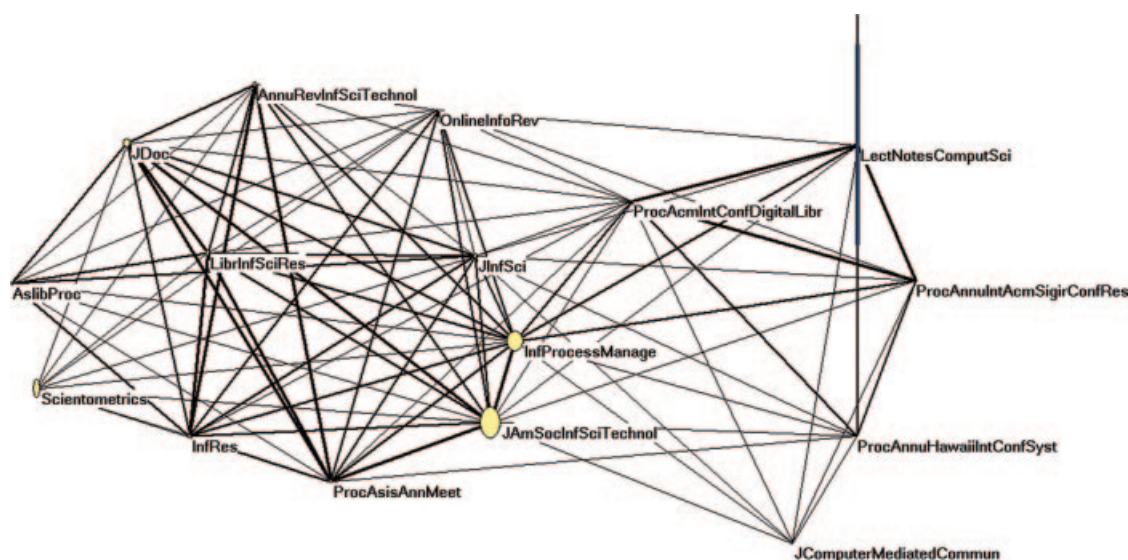


FIG. 4. The citation impact environment of *Journal of the American Society for Information Science and Technology* in the *Scopus* database; $N = 16$; threshold 1%; cosine > 0.2 .

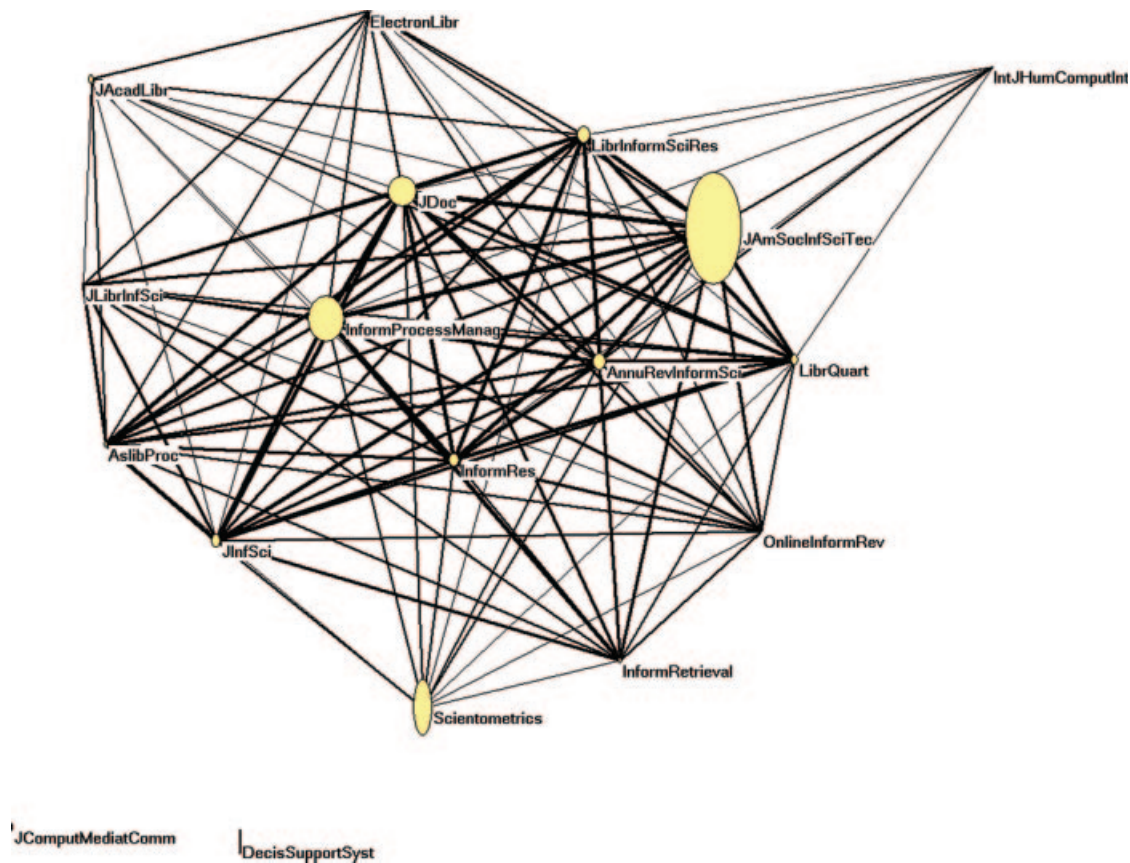


FIG. 5. The citation impact environment of *Journal of the American Society for Information Science and Technology* in the Institute of Scientific Information database; $N = 18$; threshold 0.69%; cosine > 0.2 .

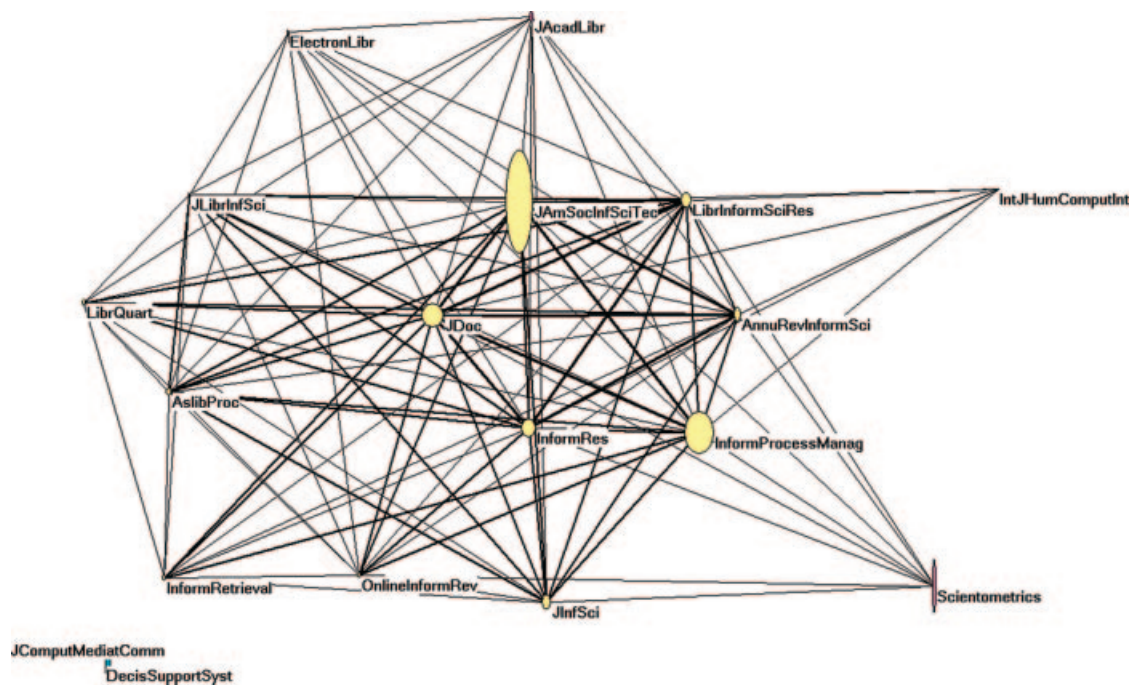


FIG. 6. The citation impact environment of *Journal of the American Society for Information Science and Technology* in the *Science Citation Index* 1998–2007 using the same set as in Figure 5 (including all citations).

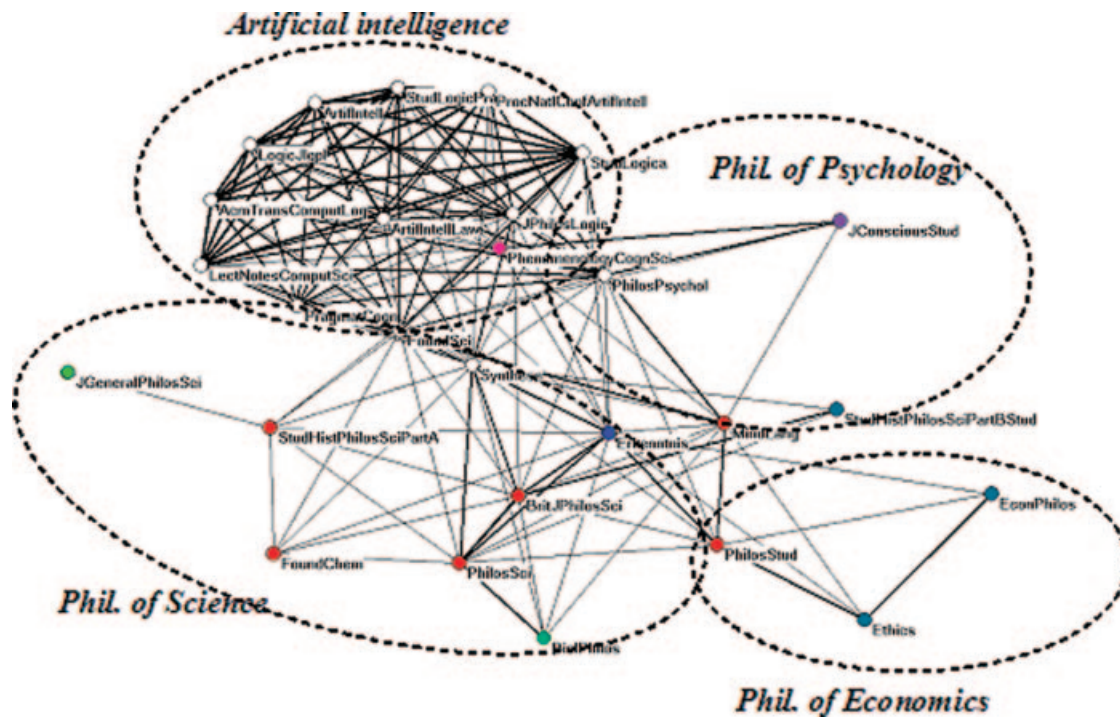


FIG. 7. Citation impact environment of *Erkenntnis*; $N = 32$; threshold: 1%; cosine > 0.2.

Is the promise of more coverage than the ISI set in the social sciences and the arts & humanities fulfilled in the *Scopus* 1996–2007 set?¹³ Using the journal *Erkenntnis* as the ego,¹⁴ Figure 7 shows the citation impact environment of this journal using the same parameters as before. Thirty-two journals are included. The screeplot suggests a model with four or five factors. Table 7 shows the four-factor solution.

Four factors explain 54.8% of the variance. The citation pattern of *Erkenntnis* itself loads on a second factor with five other journals in the philosophy of science. These five other journals are all included in the *SSCI*. However, three more journals included in the *Scopus* database can also be categorized as philosophy of science (*Studies in History and Philosophy of Science Part A*, *Foundations of Chemistry*, and the *Journal for General Philosophy of Science*). In Table 7, the things these three journals have in common is that they load with a negative sign on factor 4.

Unfortunately, the *Scopus* (1996–2007) database did not otherwise contain the rich set of humanities journals that are included in the *Arts and Humanities Citation Index* of the ISI. Three journals, for example, had the word “art” in their title: the *International Journal of Art and Design Education*, the *American Journal of Art Therapy*, and *Art Therapy*. The first of these three journals is included in the *SSCI*; the latter two are not. However, they cannot be considered as art journals.

A large set of journals (more than 100) containing the word “art” in their titles is included in the *A&HCI*. *Leonardo*, for example, can be considered as a leading journal for readers

interested in the applications of contemporary science and technology to the arts (Salah & Salah, 2008). Despite its relative ranking as a C-journal on the lists of the European Research Index for the Humanities, the journal is included in the *A&HCI* but not yet in the *Scopus* database at the time of this investigation.¹⁵ The inclusion in the *A&HCI* provides us with references in this journal from the citing side. The citations to the journal in the ISI domain can also be retrieved.

It is less well-known that the all ISI-indexes (including the *A&HCI*) allow us to construct a journal citation map from the cited references in the articles aggregated into journals. The cited reference field is highly standardized across the three ISI-Indices. For example, Figure 8 can be constructed using the 3,259 references in the 107 articles which cited *Leonardo* in 2008. The aggregate of their reference lists can be used for a “cited” journal map. Similarly, one can use all references in the 157 articles published in *Leonardo* in 2008 for a “citing” journal map.¹⁶

Three factors in this case explain 50.3% of the variance. (Because the organization of the factors does not map in terms of contingent areas in the cosine-base map, we used colors to indicate the three factors in this case.) The self-citation pattern of articles in *Leonardo* citing *Leonardo* is attributed to the third factor (designated by us as “perception research”), but with a factor loading as low as 0.165. Its loading on the “neuro” factor is negative.

Using this methodology, one can make these maps for all journals in the three citation indices of the ISI. Because the structure of the reference patterns in the *Scopus* database is more complex than the uniform format in the ISI data,

TABLE 7. Four-factor model of the citation impact environment in the ego-network of *Erkenntnis* in the *Scopus* database, 2007.

	Component			
	1	2	3	4
<i>Acm Trans Comput Log</i>	0.986			
<i>Logic J Igpl</i>	0.982			
<i>Lect Notes Comput Sci</i>	0.980			
<i>Artif Intell</i>	0.979			
<i>Stud Logic Pract Reasoning</i>	0.979		0.100	
<i>Stud Logica</i>	0.953			
<i>Proc Natl Conf Artif Intell</i>	0.946			
<i>Artif Intell Law</i>	0.828			
<i>J Philos Logic</i>	0.789	0.138	−0.117	
<i>Pragmat Cogn</i>	0.554		0.283	
<i>Found Sci</i>	0.529		0.239	−0.356
<i>Philos Sci</i>	−0.102	0.871	−0.124	−0.173
<i>Brit J Philos Sci</i>		0.860	−0.202	−0.174
<i>Synthese</i>	0.241	0.769		
<i>Erkenntnis</i>		0.689		0.188
<i>Biol Philos</i>		0.454		−0.189
<i>Stud Hist Philos Sci Part B Stud</i>		0.359	−0.258	−0.266
<i>J Radiol</i>		−0.186	−0.128	
<i>European Review</i>		−0.183	−0.122	
<i>Zool Scr</i>		−0.171	−0.130	
<i>J Conscious Stud</i>		−0.103	0.755	−0.132
<i>Phenomenology Cogn Sci</i>	0.115	−0.131	0.748	−0.241
<i>Philos Psychol</i>	0.463	0.144	0.737	
<i>Mind Lang</i>		0.261	0.612	0.349
<i>J Hist Biol</i>		−0.116	−0.123	−0.107
<i>Ethics</i>	−0.132			0.750
<i>Econ Philos</i>			−0.133	0.593
<i>Philos Stud</i>		0.473	0.229	0.585
<i>Stud Hist Philos Sci Part A</i>	−0.103	0.142	0.155	−0.386
<i>Found Chem</i>		0.199		−0.300
<i>J General Philos Sci</i>				−0.248

Note. Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

^aRotation converged in 12 iterations.

it hitherto is not possible to use *Scopus* data for this mapping by aggregating documents. The *Scopus* data require some formatting for an unequivocal translation into the ISI format as used by *HistCite*TM and other programs.¹⁷

In summary, the *Scopus* database enriches our possibilities to map the arts and humanities because of its larger journal set. However, the possibilities to map also journals in the *Arts & Humanities Citation Index* of the ISI are larger than hitherto exploited. The coverage of the arts and humanities in the *A&HCI* was at the time of this research considerably larger than in the *Scopus* database.

Extension to Interdisciplinary Journals

The extension of the journal set in the *Scopus* database allows us to address the position of peripheral journals, which could previously not be addressed. Recently, one of us used citations to nonsource journals in the ISI database to map the position of *Science & Public Policy* (Leydesdorff, 2008, p. 283). The *Scopus* database, however, includes this journal (Figure 9).

To pursue the comparison, let us add in this case the citation data about *Science & Public Policy* as a single nonsource journal to the *Social SCI* (1998–2007). Figure 9 shows an organization on the basis of the *Scopus* (1996–2007) database that is very different from the one possible on the basis of the data in the *Social SCI* (1998–2007) in Figure 10.

In both cases, *Science & Public Policy* connects a number of journals in science and technology studies. The *Scopus* database shows a relatively heterogeneous set connected in a core group, while Figure 10 provides a more differentiated map of science and technology studies. Four factors¹⁸ explain 30.1% of the variance in this case,¹⁹ while in the *Scopus* case, the configuration is to a large extent (18.0%) defined by the first factor containing the core group which is indicated in both sets. This core group shows factorial complexity—sometimes used as an indicator of interdisciplinarity (Leydesdorff & Cozzens, 1993; Van den Besselaar & Heimeriks, 2001)—in the ISI set.

The function of the interdisciplinary journal connecting different parts of this field of studies can be made visible in both representations. In previous studies, the interdisciplinary field of science and technology studies was found to be so weakly connected that it could no longer be mapped as integrated using the *JCRs* of the (*Social*) *SCI* (Leydesdorff, 2007; Leydesdorff & Van den Besselaar, 1997; Van den Besselaar, 2000). The addition of interdisciplinary journals, thus, shows the cross-connections as relatively weak ties with potentially important functions. One is more likely to find these journals included in the *Scopus* set than in the *SCI*, but the latter database allowed in this case for a richer representation by including the interdisciplinary journal from the list of nonsource items.

National Subfields

The claim is often made that the ISI databases are not suited for the evaluation of the social sciences and the humanities because these scholarly discourses are embedded in national cultures with a specific character. Hagendijk and Smeenk (1989) introduced the terminology of “national subfields” for field sciences like ecology, but the point has mainly been made more emphatically with reference to the specific character of continental European sociology and philosophy, where the premises are different from those in American and English literature (Nederhof et al., 1989). The *Scopus* database promises to enable us to study these national sets in more detail using leading journals in the respective cultures. Let us focus on sociology in Germany and France as relevant examples (Glänzel & Schöpfli, 1999; Yitzhaki, 1998). Because of the potentially low numbers of citations in this field, we included the complete citation impact environments—that is, without using a threshold—in these cases.

Figure 11 provides the representation of the ego-network of the German *Zeitschrift für Soziologie* in the *Scopus* (1996–2007) domain. The journal is visible in its national context of

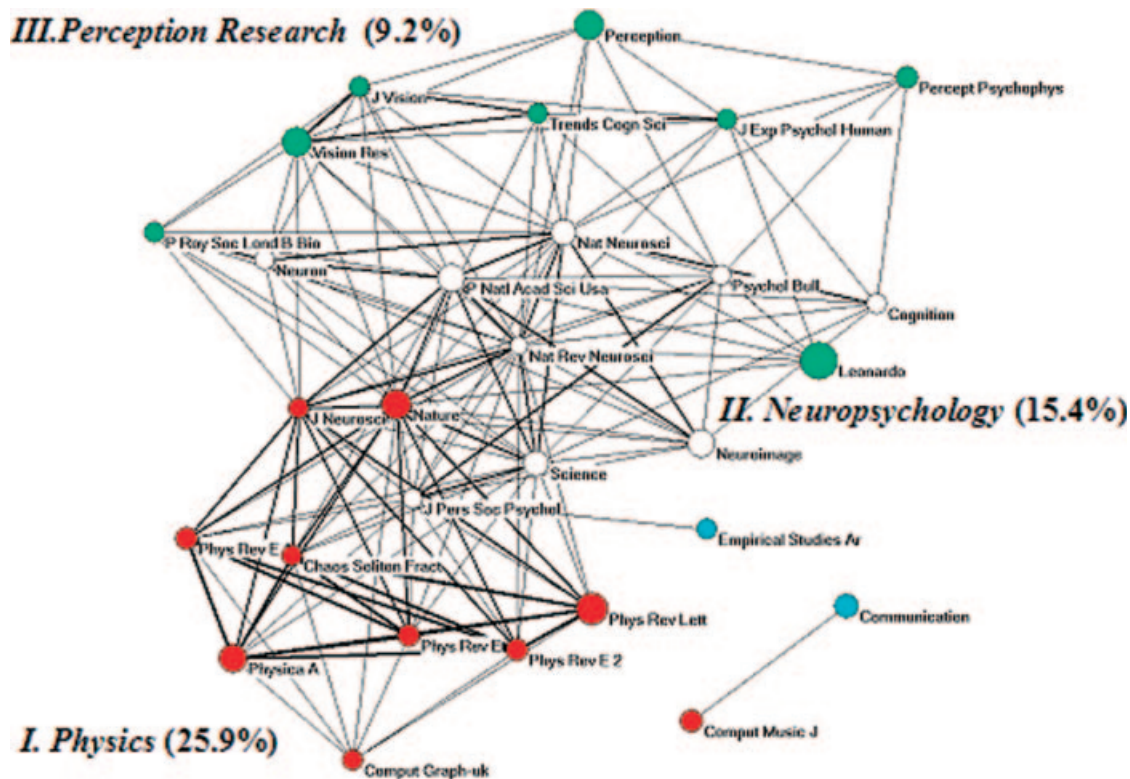


FIG. 8. Cosine relations among 3,259 references in 107 articles citing *Leonardo* during 2008; threshold 0.5%; cosine > 0.2.

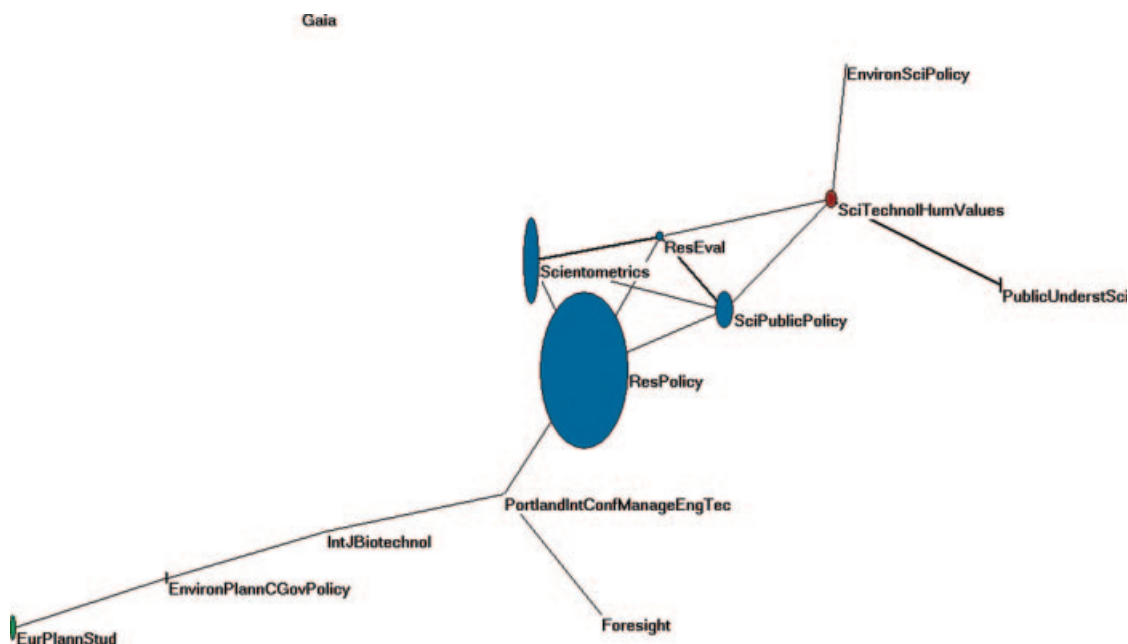


FIG. 9. The citation impact environment of *Science & Public Policy* in the Scopus database in 2007; $N = 13$; threshold: 1%; cosine > 0.2.

a core group of journals in the social sciences in the German language. The international dimension is connected via the political science journals involved in the German network, on the one side, and mainly via the European dimension, on the other. There are no citation relations to the leading American journals in this (“being cited”) direction.

The *Zeitschrift für Soziologie* is also included in the *SSCI*. Figure 12 provides the map of its ego-network in this database when the historical tail is removed (*SCI* 1998–2007). Adding the tail makes only a single additional relation visible between the *British Journal of Sociology* and the *British Journal of Sociology of Education*, and the network then is a bit

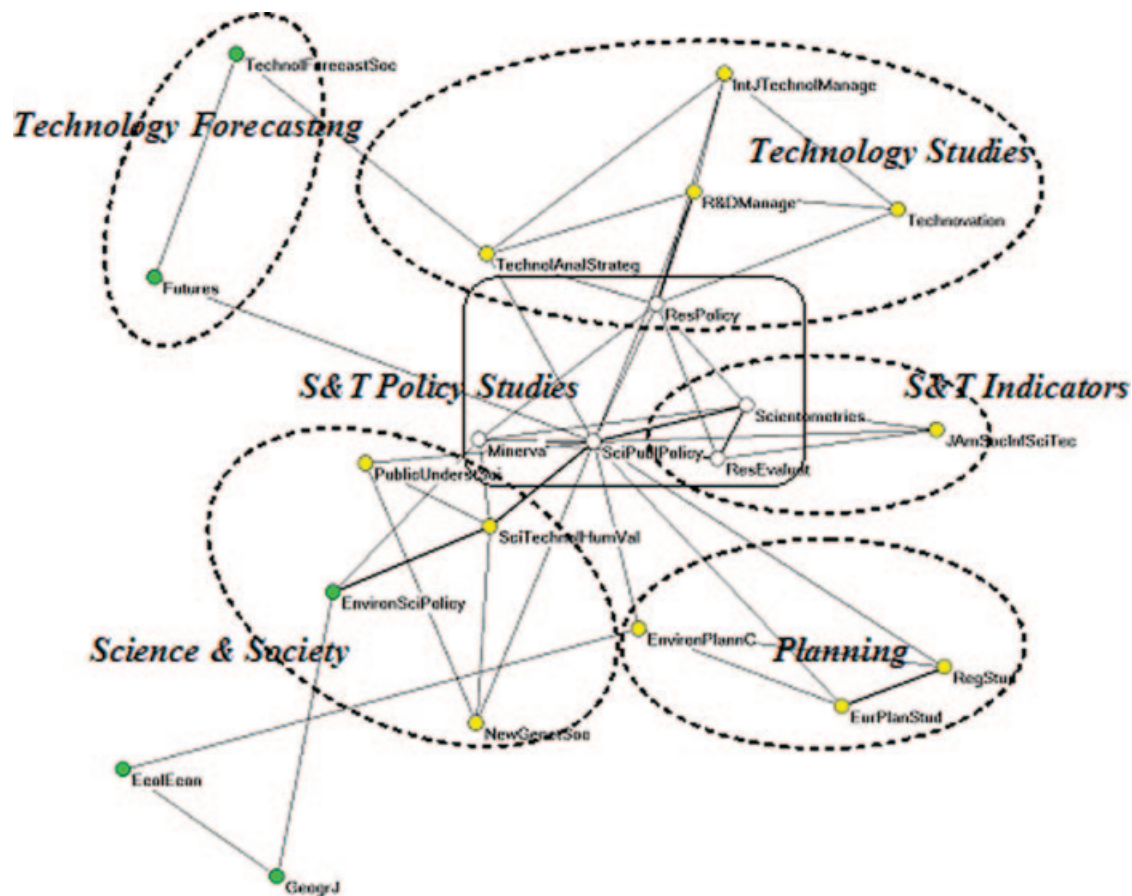


FIG. 10. The citation impact environment of *Science & Public Policy* in the *Social Science Citation Index* (1998–2007); $N = 21$; ¹⁸ threshold: 1%; cosine > 0.2.

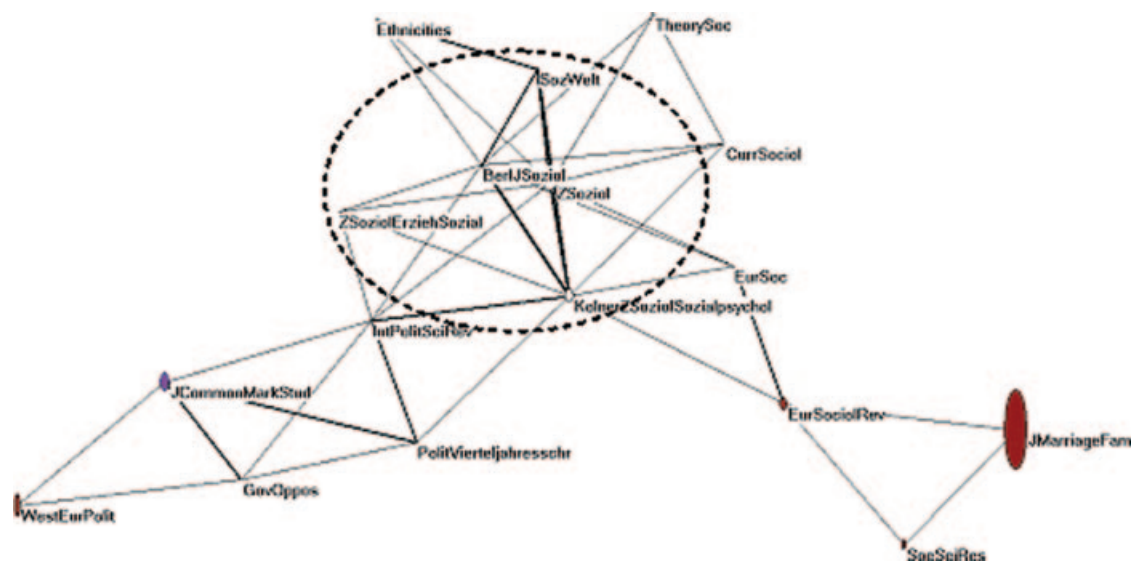


FIG. 11. The ego-network of the *Zeitschrift für Soziologie* in the Scopus database; $N = 17$; no threshold; cosine > 0.2.

denser. Although rather different again, the map based on the *Scopus* data seem to outperform the map based on the *SCI* even if the latter is not restricted to the 1998–2007 period.

In France, the national tradition in sociology is more integrated with the Anglosaxon literature than in Germany. Several (e.g., Canadian) journals accept manuscripts in both languages. A leading journal in French sociology—initiated

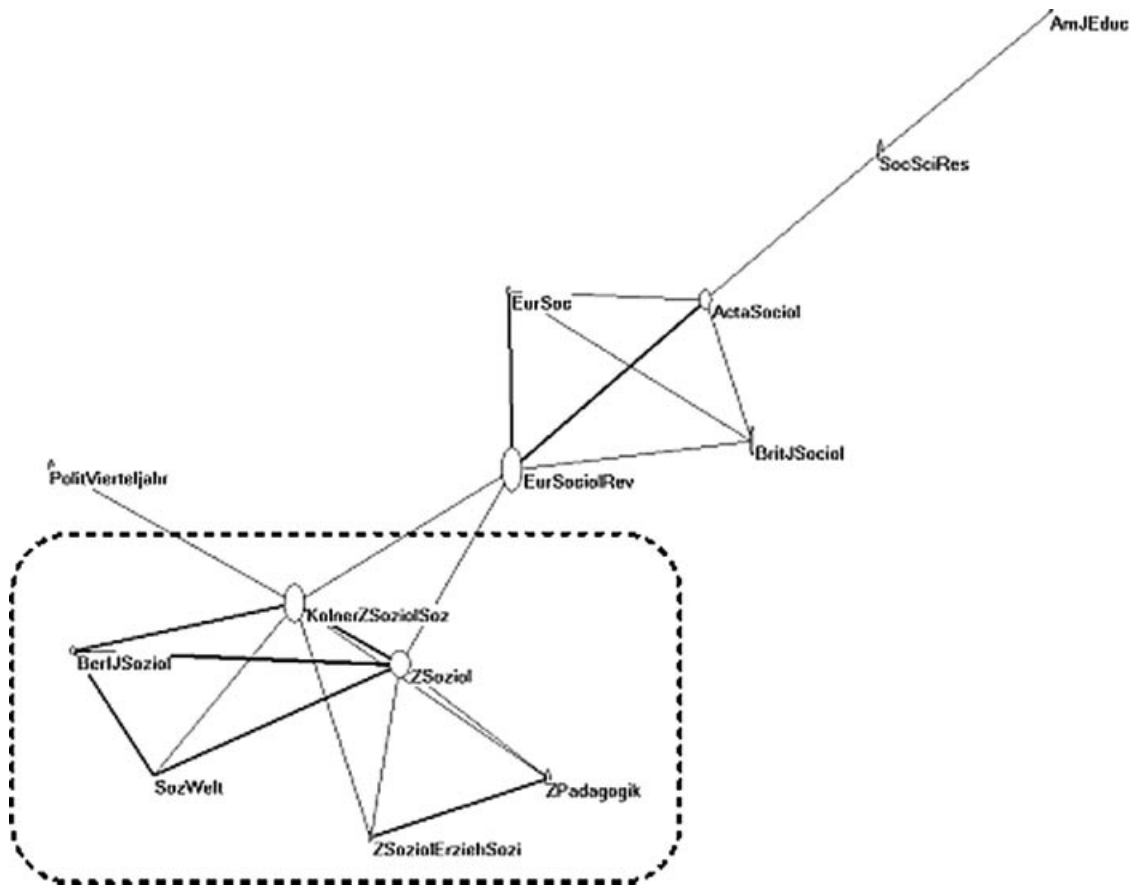


FIG. 12. The citation impact environment of the *Zeitschrift für Soziologie* in the *Social Science Citation Index* 1998–2007; $N = 13$; no threshold; cosine > 0.2 .

by Pierre Bourdieu in 1975—is *Actes de Recherche en Sciences Sociales*. This journal belongs to the group of journals not cited in the *Scopus* database, but it is included in the *SSCI* 2007. According to the *Scopus* database, *Actes de Recherche en Sciences Sociales* contains seven references during 2007 as against 766 references according to the *SSCI* for that same year. *Actes de Recherche en Sciences Sociales*, however, had an impact factor of 0.074 in 2007.

A journal in French sociology, which is included in both databases, is the *Revue Française de Sociologie* with an impact factor 2007 of 0.222. We used this journal for the comparison. Figure 13 provides the map for the *Scopus* 1996–2007 data and Figure 14 for the complete set in the *SSCI*. The restricted set in this case provides a map even poorer than Figure 14 (based on the full *SCI* data), while the latter is still outperformed by the one based on the *Scopus* (1996–2007) database in Figure 13.

Both figures show that the French sociological tradition is less distanced from the Anglosaxon discourse than the noble German one. In both the French and the German cases, the maps made on the basis of the *Scopus* data outperform the ones on the basis of the *SSCI*. Because of this focus on the social sciences, one would perhaps expect the historical tails of the distributions (in the *SSCI*) to be more influential than they are. In other words, the loss of the tail in

the *Scopus* database is more than compensated by the wider reach of this database.

Nonscholarly Journals and Newspapers

The *New York Times*, *Washington Post*, and the *Wall Street Journal* can be analyzed in terms of their citation patterns in the ISI database using the non-ISI source references. However, the journal names and abbreviations are not standardized in this list. In the *Scopus* database, these names are standardized, but as in the case of the ISI, the newspapers are not processed in terms of their citing references but registered only when they are cited in the *Scopus* domain.

There are only few citations to newspapers in the matrix based on the *Scopus* database (e.g., 43 to the *NY Times* and 14 to the *Wall Street Journal*), because newspapers are not among the source journals. In contrast, the *Wall Street Journal* was cited 2,196 times by 354 source journals of the ISI databases in 2007. (This newspaper is mostly cited in law review journals.) However, neither database allows for systematic research at the interfaces between scholarly communication and newspapers. Citation relations are typical for the scientific domain; co-word analysis among titles or full texts may be a better tool if one is interested in analyzing

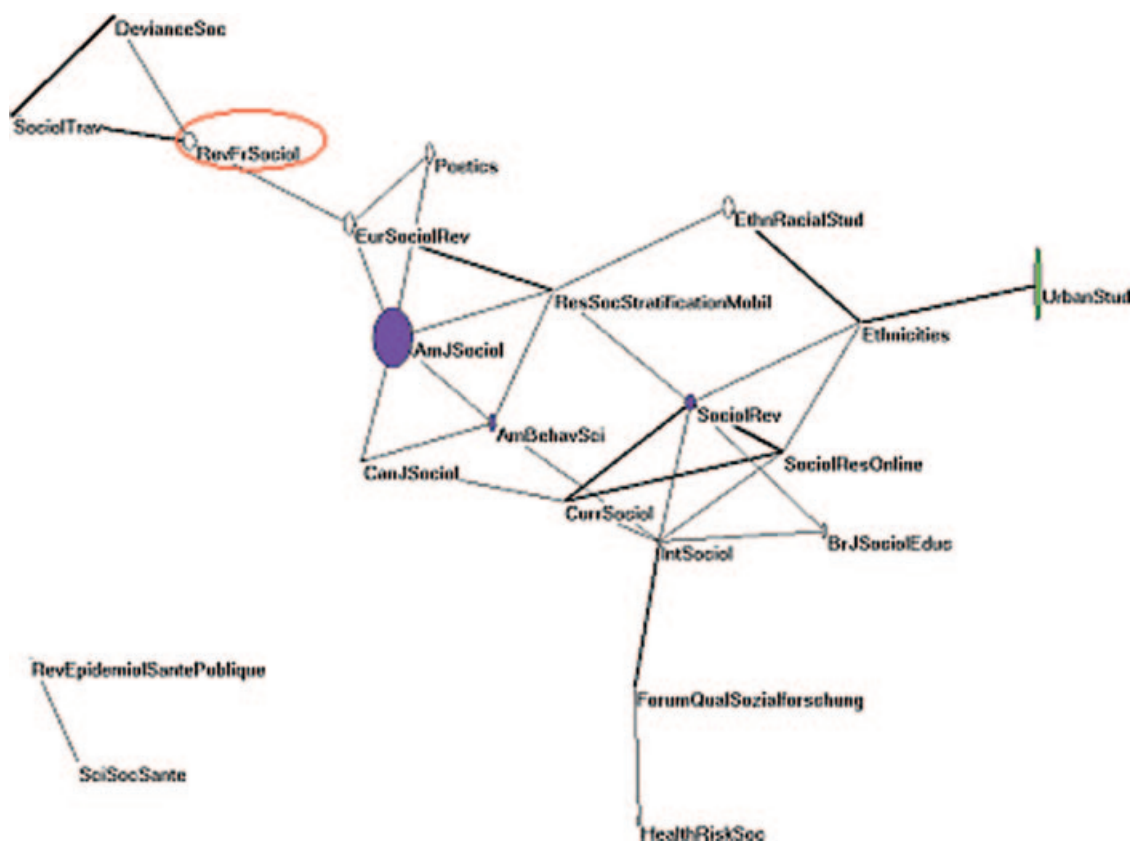


FIG. 13. Complete citation impact environment of the *Revue Française de Sociologie* in the Scopus database 1996–2007; no threshold; cosine > 0.2.

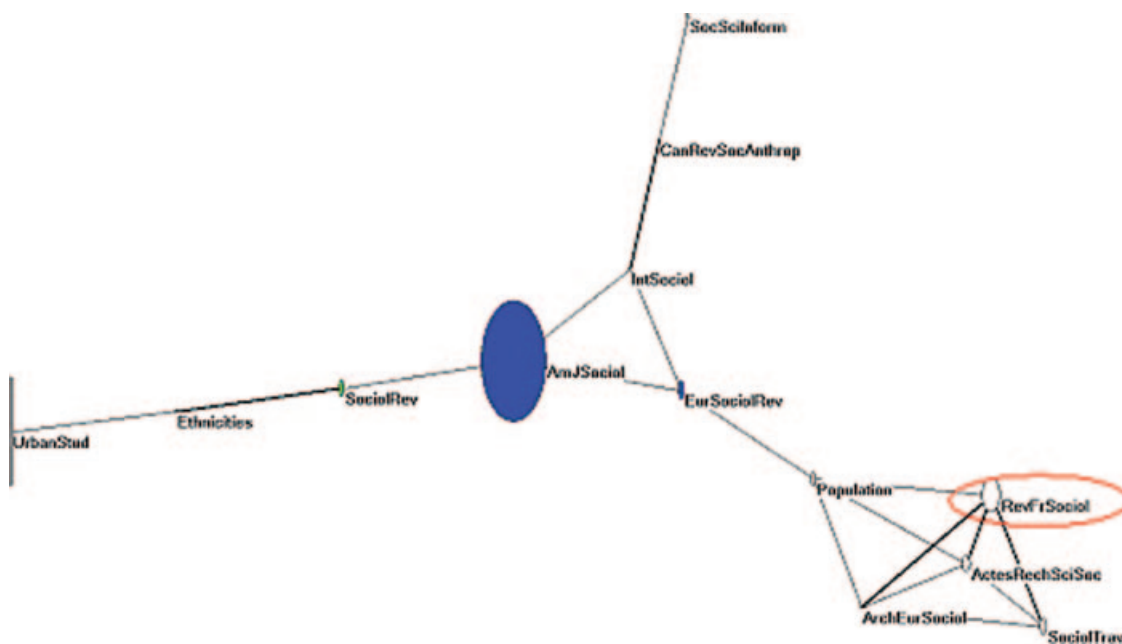


FIG. 14. Complete citation impact environment of the *Revue Française de Sociologie* in the *Social Science Citation Index* (all years); no threshold; cosine > 0.2.

developments at the interfaces of scientific communication (Callon, Law, & Rip, 1986; Leydesdorff & Hellsten, 2005).

Conclusions and Discussion

By providing aggregated journal-journal citation data, both the *Scopus* and the ISI databases enable us to map journals and to delineate specialties, while this cannot easily be done using *Google Scholar*. As is well known, the coverage of journals by *Scopus* is larger than by the ISI. Because the core groups of scholarly journals are largely the same, *Scopus* can be expected to provide mainly an enlargement when compared with the ISI set. One can expect that the added journals are more peripheral, and that the added networks will thus be less dense than those among core journals (López-Illescas et al., 2008).

In addition to this effect, the journal-journal citation matrix based on the *Scopus* data contains sometimes fewer citations to and references from individual journals than the one based on ISI data. Although the aggregations under the category “all others” have a considerable effect in the latter database, even after correction for this category, the numbers in the ISI database tend to remain higher than in the *Scopus* (1996–2007) database. The matrix based on *Scopus* is not only sparser but also less pronounced in terms of the values in the cells. Although this may significantly affect the analysis, a user may feel well-served with a map showing a less pronounced but wider net.

One major difference between the two databases is the inclusion of citations from the historical record in the ISI database, while the *Scopus* database begins only in 1996. We noted that this historical record is of the order of more than 30%, although this can be expected to vary among fields of science. In the case of the ISI database, the historical record is also reorganized in the event of a name change of a journal, but this affects the structure present in the database only marginally (Bensman & Leydesdorff, 2009). Otherwise, the citation structure is continuously reproduced and updated by the *JCRs*. The differences with the matrix based on the *Scopus* database find their origin in other rules for the inclusion of citations during the period covered by the database.

The extension of *Scopus* to include journals in the humanities could provide us with a new opportunity as long as a *JCR* for the *A&HCI* is missing. Figure 7 provided a journal map for a philosophical domain (using *Erkenntnis* as a source journal), which one cannot make with the ISI databases. Similarly, a large number of journals processed by *Scopus* are not processed by the ISI. We took *Science & Public Policy* as an example. Because these journals relate in relatively sparse parts of the network, inclusion of their relatively “weak ties” may provide new perspectives on the data (Granovetter, 1982).

In the case of *Science & Public Policy*, for example, a specific coherence among journals in the area of science & technology studies can be revealed, but this cannot be easily found using the ISI set. *Scientometrics*, for example, is firmly

embedded in another cluster of journals—that is, library and information science—in the latter database. However, a relevant map could also be provided about the journal structures in this case using the references to nonsource items in the ISI data. It seems worth the additional effort to use these ISI data in such cases. With some additional effort, journals in the arts and humanities can also be mapped in terms of aggregated journal-journal citations directly from the *Arts & Humanities Citation Index*.

Within the social sciences, it is common to complain about the inaptness of the American-based citation indices for the evaluation of nationally oriented cultural analyses, such as those published in sociology journals in the respective languages. Does the wider coverage of *Scopus* solve this problem? In the case of leading journals in German sociology, we found the two databases to point in different directions, but both indicated a prevailing isolation of this tradition in terms of its being-cited patterns. The French tradition is less isolated and better perceived in mainstream sociological research, notably, in the United Kingdom. This becomes better visible using the *Scopus* data than the *SCI* data.

Let us repeat that both databases do a fine job in providing source materials for the mapping of science. Some of the differences that are being shown in this article are a result of the differences in maturity of the two databases. The ISI data are more mature in that the ISI has had decades longer to develop cleaning, standardization, and normalization procedures that show up in the raw data and *JCR*. The *Scopus* data are newer and, thus, less clean (despite unique item identifiers) and less standardized in the cited references. However, the competition between the two may lead to further improvements such as the recent extension of the ISI database with 700 regional journals. Improvements in the *Scopus* database could focus on covering the source materials more systematically. As noted, a further extension of the coverage of journals in the arts and humanities was recently realized.

Acknowledgment

We are grateful to Almila Salah for her advice about using the *Arts & Humanities Citation Index* for the purpose of journal mapping. We are also grateful to an anonymous referee who provided us with additional information about the *Scopus* database.

Endnotes

¹In 2008, the number of included journals was increased to 23,553.

²Additionally, the *Scopus* database contains a large number of journals which were imported from MedLine (as a public database) or previously organized databases of Elsevier such as Embase and Compendex. The record before 1996 contains another 19 million references, but these were not systematically processed on the basis of the selection of source journals (Ove Kähler at *Scopus*, personal communication, 27 August 2009).

³http://isiwebofknowledge.com/products_tools/multidisciplinary/webofscience/contentexp/expansionessay/

⁴“Scopus works with European foundation to expand Arts & Humanities coverage”, at http://info.scopus.com/news/press/pr_100609.asp, June 10, 2009.

⁵At <http://info.scopus.com/detail/what/>

⁶Pajek is freely available for academic use at <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>.

⁷Degree centralization could not be computed since the networks contain loops; closeness centralization could not be computed since the networks are not strongly connected.

⁸This number was kindly provided to us by one of the referees.

⁹The number of references to the 7,793 journals which contain citations to the last ten years was 24,554,342 of which 8,005,289 are given to the cited volumes older than 1998. The resulting 16,549,053 (= 24,554,342 – 8,005,289) is provided in Table 2 as the sum of journal-journal relations in this set.

¹⁰The references to publications older than 1998 sum to 7,850,607 (out of 23,674,098) in the case of the *Science Citation Index*, and to 579,731 (out of 1,305,293) in the case of the *Social Science Citation Index*.

¹¹As explained in the help files of the JCR (at http://admin-apps.isiknowledge.com/JCR/JCR?RQ=TITLE_CHANGES), the ISI data is not affected because the name change did not change this title's position alphabetically.

¹²“Scopus works with European foundation to expand Arts & Humanities coverage”, available at http://info.scopus.com/news/press/pr_100609.asp

¹³For example: “Actually Scopus includes all of the Social Sciences titles in Thomson Scientific Social Sciences Citation Index®, as well as an additional few hundred titles.” at http://info.scopus.com/detail/what/julie_arnheim.asp (retrieved on August 9, 2009).

¹⁴*Erkenntnis* was rated as an A-journal on both the provisional lists for philosophy and the history and philosophy of science of the European Reference Index for the Humanities at <http://www.esf.org/research-areas/humanities/research-infrastructure-including-erih/erih-initial-lists.html>.

¹⁵In the meantime (August 2009), *Leonardo* is covered by *Scopus* for the years 2008 and 2009, and partly for 2004.

¹⁶Figure 7 was made using BibJourn.Exe which is freely available for non-commercial usage at <http://www.leydesdorff.net/software/bibjourn/index.htm>.

¹⁷A number of these programs are available as freeware at <http://www.leydesdorff.net/indicators/index.htm>. One of these programs is Scop2ISI.exe which enables the user to reformat downloads from *Scopus* into the tagged format of the ISI data.

¹⁸*Science & Public Policy* itself has to be added to the list using a threshold because it is not cited by itself in this representation, as it is not processed by the ISI as a source journal.

¹⁹The corresponding value is 31.6% in the case of the full set including the tails.

References

- Ahlgren, P., Jarneving, B., & Rousseau, R. (2003). Requirement for a cocitation similarity measure, with special reference to Pearson's correlation coefficient. *Journal of the American Society for Information Science and Technology*, 54(6), 550–560.
- Bar-Ilan, J. (2008). Which h-index?—A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2), 257–271.
- Bensman, S.J., & Leydesdorff, L. (2009). Definition and identification of journals as bibliographic and subject entities: Librarianship vs. ISI journal citation reports (JCR) methods and their effect on citation measures. *Journal of the American Society for Information Science and Technology*, 60(6), 1097–1117.
- Callon, M., Law, J., & Rip, A. (Eds.). (1986). *Mapping the dynamics of science and technology*. London: Macmillan.
- De Looze, M.-A., Cronini, R., Legentil, M., Jeannin, P., & Magri, M.-H. (1996). Determining the core of journals of a research centre: The example of researchers from the department of rural economy and sociology of the Institut National de la Recherche Agronomique, France. *Scientometrics*, 36(2), 167–183.
- De Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Muñoz-Fernández, F.J., & Herrero-Solana, V. (2007). Visualizing the marrow of science. *Journal of the American Society for Information Science and Technology*, 58(14), 2167–2179.
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178(4060), 471–479.
- Garfield, E. (1979). *Citation indexing: Its theory and application in science, technology, and humanities*. New York: John Wiley.
- Garfield, E. (1990). How ISI selects journals for coverage: Quantitative and qualitative considerations. *Current Contents*, 28(May), 5–13.
- Garfield, E. (1998). Journal impacts as surrogates. Paper presented at the 41st Annual Meeting of the Council of Biology Editors, Salt Lake City, UT.
- Glänzel, W., & Schoepflin, U. (1999). A bibliometric study of reference literature in the sciences and social sciences. *Information Processing and Management*, 35(1), 31–44.
- Granovetter, M.S. (1982). The strength of weak ties. A network theory revisited. In P.V. Marsden & N. Lin (Eds.), *Social structure and network analysis* (pp. 105–130.). London/New York: Sage.
- Hagendijk, R.P., & Smeenk, J.W. (1989). The analysis of national subfields: A case study of Dutch fresh-water ecology. *Scientometrics*, 15(5), 485–508.
- Jacso, P. (2005). As we may search—comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. *Current Science*, 89(9), 1537–1547.
- Jacso, P. (2008). Testing the calculation of a realistic h-index in Google Scholar, Scopus, and Web of Science for FW Lancaster. *Library Trends*, 56(4), 784–815.
- Jin, B., Jiangong, Z., Chen, D., & Zhu, X. (2002). Development of Chinese scientometric indicators (CS). *Scientometrics*, 54(1), 145–154.
- Jin, B., & Leydesdorff, L. (2004). A comparison of the aggregated journal citation structures in the Chinese science citation database and the science citation index. Paper presented at the Eighth International Conference on Science and Technology Indicators, Leiden, 23–25 September.
- Klavans, R., & Boyack, K. (2007). Is there a convergent structure of science? A comparison of maps using the ISI and Scopus databases. *Proceedings of the 11th International Conference of Scientometrics and Informetrics*, D. Torres-Salinas & H. Moed (Eds.), Vol. 1, pp. 437–448, CSIC, Madrid, 21–25 June 2007.
- Klavans, R., & Boyack, K.W. (2009a). Towards a consensus map of science. *Journal of the American Society for Information Science and Technology*, 60(3), 455–476.
- Klavans, R., & Boyack, K.W. (2009b). Identifying distinctive competencies in science. *Journal of Higher Education*, under submission.
- Leydesdorff, L. (1991). The static and dynamic analysis of network data using information theory. *Social Networks*, 13, 301–345.
- Leydesdorff, L. (1994). The generation of aggregated journal-journal citation maps on the basis of the CD-ROM version of the Science Citation Index. *Scientometrics*, 31(1), 59–84.
- Leydesdorff, L. (2007). Mapping interdisciplinarity at the interfaces between the Science Citation Index and the Social Science Citation Index. *Scientometrics*, 71(3), 391–405.
- Leydesdorff, L. (2008). Caveats for the use of citation indicators in research and journal evaluation. *Journal of the American Society for Information Science and Technology*, 59(2), 278–287.
- Leydesdorff, L., & Bensman, S.J. (2006). Classification and powerlaws: The logarithmic transformation. *Journal of the American Society for Information Science and Technology*, 57(11), 1470–1486.
- Leydesdorff, L., & Cozzens, S.E. (1993). The delineation of specialties in terms of journals using the dynamic journal set of the science citation index. *Scientometrics*, 26, 133–154.
- Leydesdorff, L., & Hellsten, I. (2005). Metaphors and diaphors in science communication: mapping the case of ‘stem-cell research.’ *Science Communication*, 27(1), 64–99.
- Leydesdorff, L., & Jin, B. (2005). Mapping the Chinese science citation database in terms of aggregated journal-journal citation relations. *Journal*

- of the American Society for Information Science and Technology, 56(14), 1469–1479.
- Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348–362.
- Leydesdorff, L., & Van den Besselaar, P. (1997). Scientometrics and communication theory: Towards theoretically informed indicators. *Scientometrics*, 38, 155–174.
- Leydesdorff, L., & Wagner, C.S. (2009). Macro-level indicators of the relations between research funding and research output. *Journal of Informetrics*, 3(4), 353–362.
- López-Illescas, C., de Moya-Anegón, F., & Moed, H.F. (2008). Coverage and citation impact of oncological journals in the Web of Science and Scopus. *Journal of Informetrics*, 2(4), 304–316.
- Maricic, S. (1997). The mainstream-peripheral science communication. *Technoscience*, 10(Winter). Also available at <http://cybermetrics.cindoc.csic.es/cybermetrics/pdf/394.pdf>
- Meho, L.I., & Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: A comparison between Scopus and Web of Science. *Journal of the American Society for Information Science and Technology*, 59(11), 1711–1726.
- Meho, L.I., & Yang, K. (2007). Impact of data sources on citation counts and rankings of LIS faculty: Web of science versus Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*, 58(13), 2105–2125.
- Moed, H.F. (2002). Measuring China's research performance using the science citation index. *Scientometrics*, 53(3), 281–296.
- Narin, F., Carpenter, M., & Berlt, N.C. (1972). Interrelationships of scientific journals. *Journal of the American Society for Information Science*, 23, 323–331.
- Nederhof, A.J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. *Scientometrics*, 66(1), 81–100.
- Nederhof, A.J., Zwaan, R.A., Bruin, R.E., & Dekker, P.J. (1989). Assessing the usefulness of bibliometric indicators for the humanities and the social sciences: A comparative study. *Scientometrics*, 15, 423–436.
- Price, D.J. de Solla (1961). *Science since Babylon*. New Haven: Yale University Press.
- Price, D.J. de Solla (1965). Networks of scientific papers. *Science*, 149, 510–515.
- Price, D.J. de Solla (1970). Citation measures of hard science, soft science, technology, and nonscience. In C.E. Nelson & D.K. Pollock (Eds.), *Communication among scientists and engineers* (pp. 3–22). Lexington, MA: Heath.
- Rafols, I., & Leydesdorff, L. (2009). Content-based and algorithmic classifications of journals: Perspectives on the dynamics of scientific communication and indexer effects. *Journal of the American Society for Information Science and Technology*, 60(9), 1823–1835.
- Ren, S., & Rousseau, R. (2002). International visibility of Chinese scientific journals. *Scientometrics*, 53(3), 389–405.
- Rosvall, M., & Bergstrom, C.T. (2008). Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4), 1118–1123.
- Salah, A.A.A., & Salah, A.A. (2008). Technoscience art: A bridge between neuroesthetics and art history? *Review of General Psychology*, 12(2), 147–158.
- Sivertsen, G. (2003). Bibliografiske datakilder til dokumentasjon av vitenskapelige publikasjoner. Oslo: NIFU skriftserie nr. 22/2003. Available at: http://www.nifustep.no/norsk/publikasjoner/bibliografiske_datakilder_til_dokumentasjon_av_vitenskapelige_publikasjoner
- Testa, J. (1997). The Thomson Scientific Journal selection process. Available at http://www.thomsonreuters.com/business_units/scientific/free/essays/journalselection/.
- Van den Besselaar, P. (2000). Communication between science and technology studies journals. *Scientometrics*, 47, 169–193.
- Van den Besselaar, P., & Heimeriks, G. (2001). Disciplinary, multidisciplinary, interdisciplinary: Concepts and indicators. The 8th International Conference on Scientometrics and Informetrics, M. Davis & C.S. Wilson (Eds.), 705–716. Sydney: University of New South-Wales.
- Van Leeuwen, T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S., & Van Raan, A.F.J. (2001). Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance. *Scientometrics*, 51(1), 335–346.
- Yitzhaki, M. (1998). The 'language preference' in sociology: Measures of 'language self-citation,' 'relative own-language preference indicator,' and 'mutual use of languages.' *Scientometrics*, 41(1), 243–254.