



Interactive overlays: A new method for generating global journal maps from Web-of-Science data

Loet Leydesdorff^{a,*}, Ismael Rafols^b

^a Amsterdam School of Communications Research (ASCoR), University of Amsterdam, Kloveniersburgwal 48, 1012 CX Amsterdam, The Netherlands

^b SPRU (Science and Technology Policy Research), University of Sussex, Freeman Centre, Falmer Brighton, East Sussex BN1 9QE, United Kingdom

ARTICLE INFO

Article history:

Received 18 July 2011
Received in revised form 6 October 2011
Accepted 7 November 2011

Keywords:

Map
Journal
Overlay
VOSViewer
Gephi

ABSTRACT

Recent advances in methods and techniques enable us to develop interactive overlays to a global map of science based on aggregated citation relations among the 9162 journals contained in the *Science Citation Index* and *Social Science Citation Index* 2009. We first discuss the pros and cons of the various options: cited versus citing, multidimensional scaling versus spring-embedded algorithms, VOSViewer versus Gephi, and the various clustering algorithms and similarity criteria. Our approach focuses on the positions of journals in the multidimensional space spanned by the aggregated journal–journal citations. Using VOSViewer for the resulting mapping, a number of choices can be left to the user; we provide default options reflecting our preferences. Some examples are also provided; for example, the potential of using this technique to assess the interdisciplinarity of organizations and/or document sets.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

In this study, we extend our previous construction of an overlay which enables users to assess a document set—downloaded from the Web of Science (WoS)—in terms of its journal composition. In the previous overlay (Leydesdorff & Rafols, 2009; Rafols, Porter, & Leydesdorff, 2010) we used the 222 ISI Subject Categories (SC) for this purpose. Although the maps using these categories can perhaps be considered reliable at a sufficiently aggregated level, in many cases—particularly, in the case of smaller sets—one would like to use a more detailed journal map instead. However, the *Journal Citation Reports* 2009 (JCR) of the (*Social*) *Science Citation Index* (SCI) contains 9162 journals. Hitherto, the mapping of this comprehensive set and its clustering has been beset with problems (Leydesdorff, 2006). More recently, technical and methodological advances enable us to provide an interactive overlay at the journal level.

Following up on these previous studies, the objective in this study is to generate an equivalently interactive global map at the journal level. As with our previous toolkit, the basemap should provide the option of using any set downloaded from the Web of Science (WoS) and visualize the sets esthetically in terms of a global map of science. It would be best to have this option available interactively at the Internet. For mapping against a baseline one may wish to use relatively small or specifically defined sets, and in such cases error introduced by a specific classification may matter.

There are cases in which one needs more finely grained, journal-based maps instead of maps based on disciplines. If the publication subset to be positioned is relatively small (e.g., below hundreds) and/or dispersed, in such cases the lack of accuracy in the delineation of ISI Subject Categories matters (Boyack, Klavans, & Börner, 2005; Pudovkin & Garfield, 2002;

* Corresponding author. Tel.: +31 20 525 6598; fax: +31 84 223 9111.

E-mail addresses: loet@leydesdorff.net, l.a.leydesdorff@uva.nl (L. Leydesdorff), i.rafols@sussex.ac.uk (I. Rafols).

URL: <http://www.leydesdorff.net> (L. Leydesdorff).

Rafols & Leydesdorff, 2009). The journal level would be not only more informative, but also more precise. A second case is when mapping interdisciplinary fields, where the assignment of journals to SCs is sometimes controversial (e.g., nanoscience and nanotechnology). Here, a journal map may provide a different perspective on the fields under study, one not necessarily convergent with that of the broader classifications.

The paper is organized as follows: First, we discuss the project of journal mapping from a historical perspective and then specify the current problems in extending from local maps to global maps and interactivity. Two software packages (Gephi and VOSViewer) have recently become available that can visualize and analyze sets as large as the 9000+ journals of the *Journal Citation Reports* of the (Social) Science Citation Reports. We discuss the pros and cons of the two packages.

Whereas Gephi provides options to make interactivity possible in the future at the Internet (using GEXFExplorer), currently VOSViewer in our opinion is the best choice for two reasons: (1) the problem of cluttering of the labels is solved in VOSViewer, and (2) the layout is based on stress-minimization as in multidimensional scaling (MDS) whereas Gephi uses a spring-embedder (Fruchterman & Reingold, 1991). The difference between these two approaches will be explained in a section on visualization techniques (Section 4).

Thereafter, we turn to the introduction and further development of the overlay maps. In addition to the visualization, VOSViewer also provides a clustering algorithm based on community detection (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008; Newman, 2004, 2006; Waltman, van Eck, & Noyons, 2010). However, the resulting classification can be replaced by the user in favor of one's own preferential classification of the journals in (inter)disciplinary categories (e.g., the ISI Subject Categories of Scopus classifications). We elaborate this with examples and further instruction.

2. Mapping the (Social) Science Citation Index at the journal level

The global mapping of the sciences using aggregated journal–journal citations has been on the research agenda of bibliometrics since the mid-eighties (Doreian, 1986; Doreian & Fararo, 1985; Leydesdorff, 1986, 1987; Tijssen, de Leeuw, & van Raan, 1987; cf. Price, 1965). At the time, several research teams concurrently realized that the data aggregated in the JCR make it possible in principle to generate maps of science at the journal level, as distinct from the maps of documents that had been developed as co-citation (e.g., Small & Garfield, 1985) or co-word mappings (Callon, Courtial, Turner, & Bauin, 1983; Callon, Law, & Rip, 1986) in previous years. In this experimental phase (of the 1980s and early 90s), however, maps were technically limited to local maps, for example, at the disciplinary or specialty level.

Journals as intellectual forms of organization of the scientific archive were first recognized by pioneers of bibliometrics such as Derek de Solla Price (1965), Francis Narin (e.g., Narin, Carpenter, & Berlt, 1972), Tibor Braun (e.g., Zsindely, Schubert, & Braun, 1982), and Eugene Garfield (e.g., 1972). Most scholars of this first generation focused on journal standing, hierarchy, and impact. Doreian (1986) developed his first journal maps from this perspective. In addition to “being cited,” however, the JCR also provides access to the aggregated citing behavior and thus the performative power of journals in organizing and restructuring the sciences intellectually (Latour, 1987).

The clusters and consequent configurations in journal maps can show how journal groups develop as intellectual organizers in niches (Leydesdorff & Van den Besselaar, 1997; Lucio-Arias & Leydesdorff, 2009; Van den Besselaar & Leydesdorff, 1996). These niches can robustly—that is, relatively independently of the precise clustering algorithms—be recognized as specialties. Neighbouring specialties are often weakly connected to one another across intellectual borderlines. Specialties develop concurrently and therefore heterarchically—that is, with different types of overlap and without an ordered nesting of categories.

The exceptions are “interdisciplinary” journals which relate across disciplinary borders at a next-higher level—with multidisciplinary journals such as *Nature* and *Science* at the top (Carpenter & Narin, 1973; Leydesdorff & Bornmann, 2011)—or at the bottom with reference to engineering, the clinic, or other practical applications (Gibbons et al., 1994; Leydesdorff & Jin, 2005). The emergence of new journals can also serve as an indicator of new developments in science by means of “visual analytics” (e.g., Cook & Thomas, 2005; Thomas & Cook, 2006; cf. Leydesdorff & Schank, 2008). Such tools can be used as indicators for policy purposes in the early stages of an emergent field and/or technology (Leydesdorff, Cozzens, & van den Besselaar, 1994).

Developments at the journal level provide the analytical advantage of being largely beyond the control of individual authors or funding agencies (Leydesdorff, 1987; Studer & Chubin, 1980). (Journals also operate in a market that is populated with publishing houses, libraries, etc.) Authors can contribute to and sometimes change the reproduction of these structures by publishing new manuscripts. Although aggregated journal citation patterns among major journals are often stable over the years, new developments can be indicated using animations of journal maps for different years (Leydesdorff et al., 1994; Leydesdorff & Schank, 2008). These dynamics of the communications aggregated at the level of journals are non-linear. In other words, they develop most of the time incrementally and recursively, building on previous citation patterns, but may exhibit evolutionary change and path dependencies. The development of specialties in terms of sets of journals can reinforce borderlines by signaling an emerging order in the respective discourses (e.g., Leydesdorff & Probst, 2009; Rice, 1990).

Perhaps, in this age of the internet and Google Scholar where authors can read across journals more easily, this organizing power of journals may have waned (Bollen, Sompel, Smith, & Luce, 2005), but counteracting this is an increased emphasis in research management on legitimacy and citation impact. Institutional incentives have continued to point towards publishing in journals with the highest possible prestige (Halffman & Leydesdorff, 2010). Prestige is often associated first with inclusion in the (Social) Science Citation Index, or to a lesser extent in Scopus (because the latter database covers a considerably

larger—and therefore less restricted—number of journals). The internet interfaces of these databases, and the freely available program *Publish or Perish*,¹ respectively, allow researchers to track publication and citation records in these various domains.

In summary, the use of aggregated journal–journal citation relations for the mapping of the sciences has flourished since the *Journal Citation Reports* (JCR) were made available for the *Science Citation Index* in 1975 and the *Social Science Citation Index* in 1978. Until the JCRs became electronically available in 1994, however, the analysis was confined to the construction of local maps. Whereas new developments within the sciences occur locally, mapping these developments in the contexts of larger sets became possible with electronic data and improved visualization techniques since the mid-90s.

3. Visualization techniques

In the 1980s and the early 1990s attempts to map the sciences were also limited by the graphic interfaces of the line-based DOS and Unix environments, and by the intricacies of plotting alternatively by using mainframe computers. The versions of UCINET available at the time, for example, included multidimensional scaling (MDS), but the plotting power was confined to the 26 upper-case and 26 lower-case character sets for the labeling of 52 data points. The advent of the new technologies from Apple and Microsoft's Windows (in 1995) brought more ambitious programming efforts, such as the development of the network analysis and visualization program Pajek in 1996.² The higher resolution and increased computer power triggered a new spurt in developments during the first decade of the 2000s (Börner, 2010; Börner, Chen, & Boyack, 2003).

Boyack et al. (2005), for example, used large-scale computational power (VxOrd) to organize journals for the first time into a global map which these authors called programmatically “the backbone of science.” Rosvall and Bergstrom (2008) developed a global journal map based on statistical decomposition analysis (Leydesdorff, 1991, 2005; Theil, 1972). Rosvall and Bergstrom (2010) extended this approach to include the dynamic analysis. Still, the results of the algorithmic clustering remained unconvincing (Rafols & Leydesdorff, 2009), and the choice of similarity criteria has remained a cause for concern (Ahlgren, Jarneving, & Rousseau, 2003). Although at a large scale the global maps may seem robust (Börner, 2010; Klavans & Boyack, 2009), the lower-level structures (for example, at the journal level) can depend on small changes in the relevant parameters (Klavans & Boyack, 2006; Leydesdorff, 2006).

The analytical implications of parameter choices are not always under control: visualization necessarily involves the reduction of a multidimensional space into the two dimensions of a sheet or screen. Community-finding algorithms based on modularity (Blondel et al., 2008; Newman, 2004, 2006) may enable us to decompose large matrices in terms of their nearly decomposable components, but the results depend also on a randomly chosen seed number, and therefore cannot reliably be reproduced (Renaud Lambiotte, *personal communication*, December 16, 2009). Are the fluctuations among the solutions indicative of relevant changes in the data, or are they statistical effects? The solutions may remain dependent on the parameter choices made by the users (Van Eck, Waltman, Dekker, & van den Berg, 2010).

Two solutions to the problem of the projection to a two-dimensional map have prevailed: multidimensional scaling (MDS) based on stress-minimization (Kruskal, 1964; Kruskal & Wish, 1978) or minimization of an energy function in the graph conceptualized as a system of springs (e.g., Fruchterman & Reingold, 1991; Kamada & Kawai, 1989). MDS can be considered as equivalent to factor analysis or principal component analysis (Schiffman, Reynolds, & Young, 1981, at p. 13), but the primary objective is visualization (and not clustering). Some network visualization programs (such as UCINET and Visone) offer MDS directly as a possible subroutine for the visualization.

Using MDS, the network is first conceptualized as a multi-dimensional space that is then reduced stepwise to lower dimensionality. At each step, the stress increases using, for example, Kruskal's stress function formulated as follows:

$$S = \sqrt{\frac{\sum_{i \neq j} (||x_i - x_j|| - d_{ij})^2}{\sum_{i \neq j} d_{ij}^2}} \quad (1)$$

In this formula $||x_i - x_j||$ is equal to the distance on the map, while the distance measure d_{ij} can be, for example, the Euclidean distance in the data under study (Borgatti, 1997). The choice of a distance or similarity measure remains open. In the 1980s and 90s, for example, one of us used MDS to illustrate factor-analytic results in tables, and in this case the Pearson correlation obviously provided the best match (e.g., Leydesdorff, 1986).

Spring-embedded or force-based algorithms can be considered as a generalization of MDS, but were inspired by further developments in graph theory during the 1980s. Kamada and Kawai (1989) were the first to reformulate the problem of achieving target distances in a network in terms of energy optimization. They formulated the ensuing stress in the graphical representation as follows:

$$S = \sum_{i \neq j} s_{ij} \quad \text{with } s_{ij} = \frac{1}{d_{ij}^2} (||x_i - x_j|| - d_{ij})^2 \quad (2)$$

¹ Publish or Perish is available at <http://www.harzing.com/pop.htm>.

² Pajek is available for non-commercial usage at <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>.

Eq. (2) differs from Eq. (1) by taking the square root in Eq. (1), and because of the weighing of *each* term with $1/d_{ij}^2$ in the numerator of Eq. (2). This weight is crucial for the quality of the layout, but defies normalization with $\sum d_{ij}^2$ in the denominator; hence the incomparability between the two stress values.

The ensuing difference at the conceptual level is that spring-embedding is a graph-theoretical concept developed for the topology of a network. The weighing is achieved for each individual link. MDS operates on the multivariate space as a system, and hence refers to a different topology. In the multivariate space, two points can be close to each other without entertaining a relationship (Burt, 1982; Granovetter, 1973). For example, they can be close or distanced in terms of the correlation between their *patterns* of relationships (Burt, 1995).

In the network topology, Euclidean distances and geodesics (shortest distances) are conceptually more meaningful than correlation-based measures. In the vector space, correlation analysis (factor analysis, etc.) is appropriate for analyzing the main dimensions of a system. The cosines of the angles among the vectors, for example, build on the notion of a multi-dimensional space. The remaining discussion is over the choice of using Pearson correlations or cosine values (Egghe & Leydesdorff, 2009). In bibliometrics, Ahlgren et al. (2003) have argued convincingly in favor of the latter because of the skewedness of the citation distributions and the abundant zeros in citation matrices.

Technically, one can also input a cosine-normalized matrix into a spring-embedded algorithm. The value of $(1 - \text{cosine})$ is then considered as a distance, albeit in the vector space (Leydesdorff & Rafols, 2011a). In sum, there is a wealth of possible combinations in the parameter space of clustering algorithms and similarity criteria. Analytical arguments and not the aesthetics or policy-relevance of the representations should guide us towards the choices to be made.

4. Methods and data

The data for this project was harvested from the Internet version of the *Journal Citation Reports* 2009 in August 2010. (Currently, an equivalent database using Scopus data is not available; cf. Leydesdorff, de Moya-Anegón, & Guerrero-Bate, 2010). On the basis of this data an aggregated journal–journal citation matrix of 9162 journals was constructed. Of the $9162^2 = 83,942,244$ cells only 2,790,876 (=3.32%) are filled with values larger than zero; the grand total of the matrix is 33,190,645 citations, or on average 11.89 per cell with a value larger than zero. This data was gathered from the “citing” side. Although the long tails of low values are sometimes summed on this (citing) side as “all others,” the file contains 1209 cells with a value of one. The cutoff at the lower end varies in the JCR with the sizes of the tails.³

The matrix was transformed into a cosine-normalized matrix both in the being-cited and the citing dimensions. Matrices with thresholds for the cosine values can then be exported in formats that can be read by the various visualization programs. We used SPSS for the cosine-normalization and Pajek/UCInet for the data manipulation. VOSViewer (based on an MDS-like algorithm)⁴ and Gephi (containing a spring-embedder)⁵ will be compared below for the visualization.

For example, after normalization in terms of the “being cited” patterns and with a threshold of cosine > 0.2, the matrix contains 346,771 (off-diagonal) values larger than zero. More than 99% (346,746) of these edges are contained within a largest component of 8817 nodes (96.23%). The other nodes are mainly isolates. At the threshold level of cosine > 0.5, however, only 5954 nodes (64.99%) are connected by 51,030 edges (less than 15%) in the largest component. We use this latter (much smaller) matrix to first sort out our methodological questions—e.g., about the choice of the visualization software—and then return to the larger set for the overlay map. Visualization software uses largest components because the isolates and non-related components cannot be positioned unambiguously with reference to the larger set.

As noted, we focus on two programs recently made available for the visualization: VOSViewer (Van Eck & Waltman, 2010) and Gephi. Both programs are Java-based and thus their capacity is limited only by the hardware configuration. We worked with a standard desktop PC with 4 GB internal memory under Windows 7, 32-bits. This PC has problems handling the large component of 8817 journals at the threshold level of cosine > 0.2, although Gephi manages to handle the large file after a long processing time. VOSViewer gave no error message for this file size using 8 GB of RAM and a 64-bits operating system.⁶ One needs to generate these results only once since they can be saved as the coordinates of the basemap (to be discussed below).

Gephi has an option to export files in the so-called gexf format which can be embedded in the html using GEXFExplorer. Leydesdorff, Hammarfelt, and Salah (2011) used this option to visualize the aggregated journal citation structures of the *Arts & Humanities Citation Index* 2008 at the Internet. Furthermore, a spring-embedder—the algorithm of Fruchterman and Reingold (1991)—is available in Gephi, and the algorithm for modularization of Blondel et al. (2008) is available for clustering the data; the resulting clusters can be colored accordingly. In summary, the graphic options of Gephi are state-of-the-art. As noted, the program is also efficient in handling large-size files.

³ As of 2009, Thomson Reuters no longer provides the CD-Rom versions of the JCR which allowed for more precise database management. The data for this study was harvested from the web-version of this database.

⁴ VOSViewer is freely available at <http://www.vosviewer.com>.

⁵ Gephi is freely available at <http://gephi.org>. Gephi is open source (GNU APGL).

⁶ Ludo Waltman at CWTS (in Leiden) was so kind as to run our files at a large machine with 64 GB. The results were similar (*personal communication*, 26 April 2011).

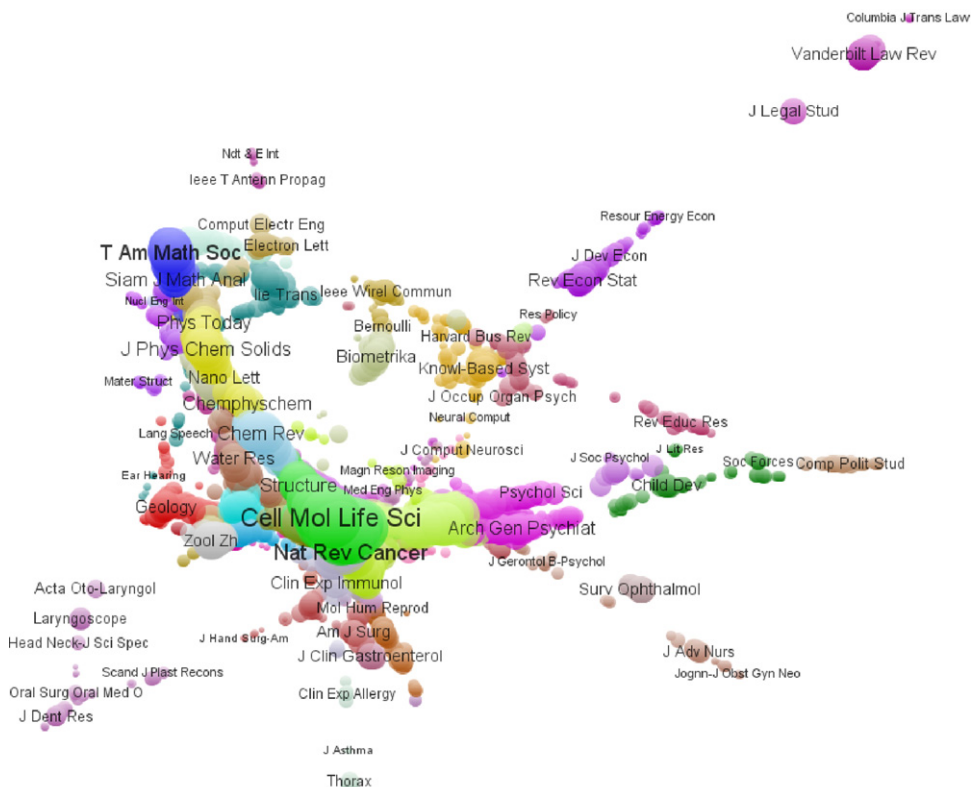


Fig. 1. 5954 journals similar in their being-cited patterns above cosine > 0.5; 70 colors (clusters) distinguished by VOSViewer ($\gamma = 1$).

The abbreviation “VOS” in VOSViewer stands for “visualization of similarities.” The algorithm used for this is akin to that of MDS: VOSViewer minimizes a stress function at the systems level (Van Eck et al., 2010). More recently, Van Eck and Waltman (2010) have integrated a clustering algorithm into the program that operates on the basis of the same principles as the positioning of the nodes. The cluster results are automatically colored into the map. The clustering algorithm operates with a parameter (γ) that can be changed interactively in order to generate more or fewer clusters in the solution. Additionally, a representation of the map as a density or heat map is provided in VOSViewer.

From our perspective, VOSViewer is closer to our objectives than Gephi because it relies on an MDS-like algorithm to position the data in the multidimensional vector space, whereas Gephi provides graph-analytical routines that operate on the networked relations. However, both programs offer a wealth of possibilities for the visualization of a comprehensive map. We use the (smaller) set—normalized with cosine > 0.5—in the cited dimension first for comparison of the options, then provide arguments for our choice of VOSViewer, and change thereafter to the larger set of 8800+ journals included in the largest component when the threshold is set at cosine > 0.2.

5. Construction of the basemap in VOSViewer or Gephi?

5.1. VOSViewer

VOSViewer allowed us to import the Pajek-file for the set of 5954 journals without a problem. The larger set of 8817 journals generated an error message because of memory shortage, as did also the in-between sets for cosine > 0.3 (8378 journals) or cosine > 0.4 (7504 journals). Extending the available memory by changing the parameter for the Java .jar file with 8 GB and 64-bits solved this problem. Thus, it appears that the generation of the basemap will soon be within the reach of an average user of VOSViewer. The current limitation is a minor disadvantage because the coordinates of the global map can be made available as an ASCII text file that can be loaded into the program or webstarted at the Internet. Another minor disadvantage of VOSViewer is perhaps that the computation using MDS-type positioning requires a longer time than using spring-embedded algorithms.

Fig. 1 provides the basemap for the file of 5954 journals that are similar in their being-cited patterns above cosine > 0.5. In addition to the torus-like construct normally found in these maps (Klavans & Boyack, 2009), this visualization also suggests concentrations in specific directions. At the top right, for example, some law journals are set apart. The bio-medical cluster at the bottom left shows a fine structure. Seventy clusters are distinguished in this case and colored differently.

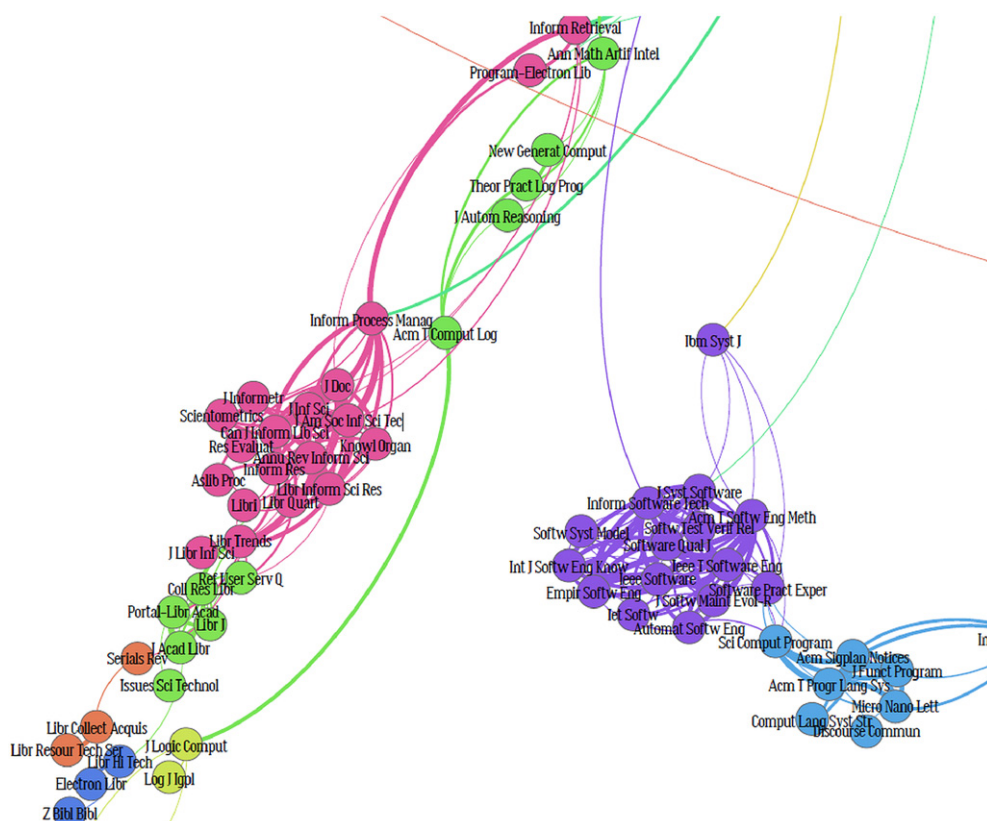


Fig. 2. Details of the map of 400+ communities (modularity 0.83); enlarged 8 times. See the online map at <http://www.leydesdorff.net/journalmaps/forceatlas.htm> for more details.

One can associate the right upper half of the figure with the social sciences and the left-lower half with the natural and life sciences, which are then aligned on a continuum from mathematics at the top to specialties in clinical medicine at the bottom. (In MDS and VOSViewer, results can freely be rotated or translated without changing the stress values.) However, the distinctions among these larger structures may be caused by the relatively high value of the threshold (cosine > 0.5) which partitioned the network into 70 communities. As noted, we shall use a lower threshold value for the overlay below.

5.2. Gephi and GEXFExplorer

Although Gephi had no problems importing the larger file of 8817 journals (cosine > 0.2), the exported file in GEXF becomes so large (>60 MB) that using this file at the Internet may lead to problems. Thus, we worked again with the smaller file of 5954 journals (cosine > 0.5). Because of this high threshold, the number of communities detected by the modularity algorithm (Blondel et al., 2008) increases to slightly above 400, with modularity values (Q) between 0.82 and 0.84. Using the larger file (cosine > 0.2), the connections are denser and modularity drops to values of $Q \approx 0.64$ distinguishing approximately 160 communities.

As noted, these values are statistical. Consequently, one cannot obtain twice the same output—and hence coloring—using the algorithm. In a previous project (Rafols & Leydesdorff, 2009), we found it impossible to make comparisons across years using this algorithm. The community-finding algorithm not only generates fluctuations (because of using random seeds), but the data may also fluctuate from year to year (Renaud Lambiotte, *personal communication*, 16 December 2009).⁷

Gephi comes with two algorithms for the layout that are particularly suited to our purpose: the spring-embedder of Fruchterman and Reingold (1991) and a new algorithm called ForceAtlas that was used by the Gephi team for this purpose (Hachul & Junger, 2005; Harel & Koren, 2002; cf. Hachul & Jünger, 2006). Sébastien Heyman of the Gephi team was so kind as to run our file on his computer using the second version of ForceAtlas at the time still under development (*personal*

⁷ Rosvall and Bergstrom (2010) claimed that nephrology, although separate in 2001, joined the area of medicine in 2003, but more or less departed from it again in 2005. Because this was tested for significance in each year, the authors claim that these changes are significant. Without significance testing, however, the visible changes may be due to fluctuations in the intensity with which authors in nephrology cite and are cited in the literature in general medicine. Such problems may be unavoidable in the dynamic map.

Table 1
19 clusters of journals in citing patterns using VOSViewer ($\gamma = 1$).

	Clusters 1–19	Nr of journals	Factor 1–18 (Rafols et al., 2010)
1	Social Sciences	1962	Biomedical Sciences
2	Agricultural Sciences	1349	Materials Sci.
3	Chemistry	1342	Computer Sciences
4	Clinical Medicine	1177	Clinical Medicine
5	Computer Sciences	1062	Economics, Politics and Geography
6	Biomedical Sciences	864	Psychology
7	Infectious Diseases	288	Ecological Sci.
8	Biophysics	285	Chemistry
9	Economics & Management Sci.	266	Geosciences
10	Neurosciences	173	Cognitive Sciences
11	Occupational Hygiene	34	Health and Social Issues
12	Rheumatology	33	Engineering
13	Multidisciplinary	16	Environmental S&T
14	Biomedical Engineering	4	Agricultural Sci.
15	Micro & Nano Letters	1	Infectious Diseases
16	Journals of Gerontology Series A-Biological Sciences And Medical Sciences	1	Social Studies
17	Pteridines	1	Physics
18	Pediatric Research	1	Business and Management
19	Turkish Journal of Biochemistry-Turk Biyokimya Dergisi	1	

Only 21 communities are distinguished by VOSViewer at this level of connectedness (using $\gamma = 1$; as against 70 at the level of cosine > 0.5).⁸ In other words, the system is more tightly knit than above (in Fig. 1). However, this strong reduction in the number of colors (clusters) can be convenient for our purpose: our objective is to map at the level of journals and not in terms of (hundreds of) specialties. We shall discuss the quality of the clustering below in the case of the citing map.

Using the property that the coordinates of all nodes can be saved in a (comma-separated) database file, the abbreviations of the titles could be used to generate an interface between any download from the WoS and this map. Unfortunately, the abbreviations are not completely standardized between the JCR and the WoS. Thus, for the matching we use the full titles of the journals instead.

7. The citing map as a baseline

The same 9162 journals can be evaluated alternatively in terms of their citing patterns. Whereas the “being-cited” patterns refer to the archive of science, “citing” is the running variable which is exclusively set to the current publication year. Journals are perceived in their environment mainly in terms of how they are cited as an archive (Bensman, 2007). The citing variable, however, does not accumulate over the years.

The positioning in this (citing) dimension is coupled to the activity of the citing authors, and thus can indicate changes in citation patterns without a lag or dampening. For this reason of continuous updating by the community (of authors) itself, we prefer this direction for the map in the case of the previous overlay in terms of ISI Subject Categories. Citing represents the knowledge base of new knowledge claims, while cited represents the accumulated impact of journals on audiences. However, this choice can also be left to users according to their specific research questions.

The largest “citing” component connects 8860 of the 9162 journals and contains 928,878 of the 928,882 edges ($>99.9\%$). In other words, the set is virtually complete. The network is less decomposable than the previous one, with a modularity of $Q \approx 0.475$ distinguishing approximately 40 communities (using Blondel et al.’s (2008) algorithm for the unfolding). In other words, in their citing behavior authors reach across specialty boundaries, but this can be considered as variation, while the accumulated structures in the cited journals are reproduced more regularly as boundaries in the intellectual organization of the sciences.

In Table 1 the 19 clusters distinguished by the algorithm in VOSViewer are compared with the 18 factors that we distinguished on the basis of factor analysis for coloring the overlay among the 222 ISI Subject Categories in Rafols et al. (2010).

⁸ At $\gamma = 1.5$ and $\gamma = 2$, the number of clusters are 31 and 39, respectively.

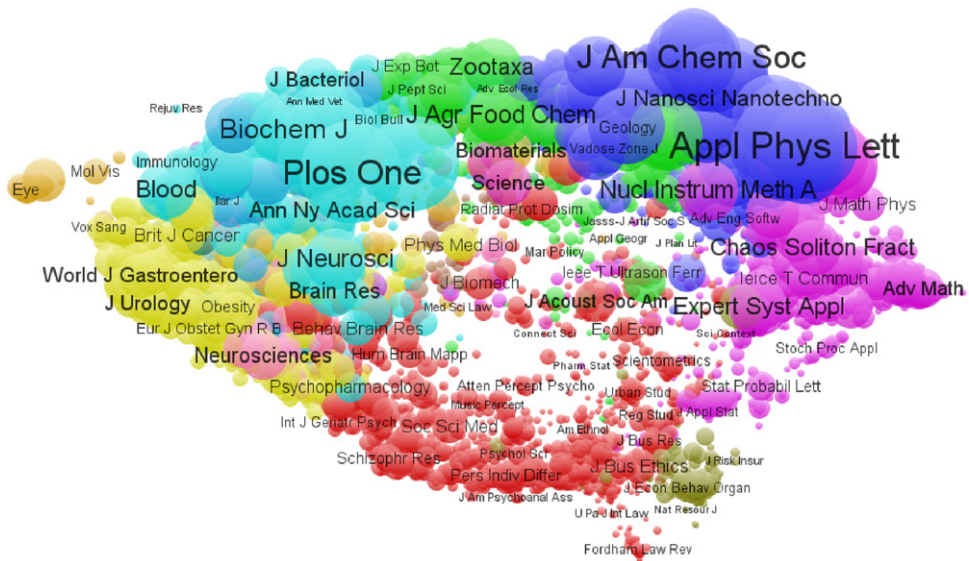


Fig. 4. 8860 journals similar in their citing patterns above cosine > 0.2; 19 clusters ($\gamma = 1$) using the algorithm built into VOSViewer (Van Eck et al., 2010). (At $\gamma = 1.5$ and $\gamma = 2$, the number of clusters are 77 and 121, respectively.) (Sizes of the nodes correspond to numbers of publications in 2009.) (One can webstart this map at <http://www.vosviewer.com/vosviewer.php?map=http://www.leydesdorff.net/journalmaps/citing02.txt>.)

Five clusters distinguished by VOSViewer contain only a single journal, and one has only four journals. These six clusters are barely visible in the map. Actually, only the first ten clusters can be considered meaningful, and these dominate the map. Using these colors, in our opinion, is informative, although this clustering is as yet far from convincing.⁹

In a previous study (Rafols & Leydesdorff, 2009), we discussed various possible classification systems of journals, yet without a conclusive recommendation. One advantage of these new maps is that the user can change the classification system to suit specific needs. After reading the file citing02.txt (available at <http://www.leydesdorff.net/journalmaps/citing02.txt>) into Excel, the user can replace the column labeled “cluster” with another classification. For example, the file at <http://www.leydesdorff.net/journalmaps/citing/blondel.txt> contains the same mapping information, but uses the 41 communities distinguished by the algorithm of Blondel et al. (2008). This provided us with the following figure (Fig. 5).

The community structure is broadly similar, but more finely grained particularly in the relatively poorly populated area contained by the torus. In summary, one can freely change the clustering and accordingly the coloring while maintaining the basemap for the overlay. The mapping and clustering are thus uncoupled. One analytical advantage of such uncoupling is to forestall potential erroneous interpretations arising from either the reduction of dimensions in the mapping or the aggregation in the clustering.

If one wishes to replace the *default* coloring for generating the overlays, the column “cluster” in the table files at <http://www.leydesdorff.net/journalmaps/citing.dbf> and/or <http://www.leydesdorff.net/journalmaps/cited.dbf>, respectively, must be changed. The programs for generating the overlays will then use the newly inserted values.

8. The generation of the overlay files

Two programs are made available online at <http://www.leydesdorff.net/journalmaps/citing.exe> and <http://www.leydesdorff.net/journalmaps/cited.exe>, respectively, for processing a file containing downloaded data from the WoS. These .exe files also need the table file citing.dbf and/or cited.dbf, respectively, in the same folder. (These table files can be downloaded from the same page at <http://www.leydesdorff.net/journalmaps>.) In addition to the coordinate information for the maps, the full titles of the journals as provided by the JCR are listed in these files. As noted, the titles of the journals are used as keys for the matching. (In the case of an unforeseen mismatch one is advised to adapt the title in the corresponding table file.)

When the programs and tables are brought into a single folder with the input file, which is downloaded from the WoS and renamed “data.txt,” an output file can be generated. This file is called either “cited.txt” or “citing.txt” depending on the routine in question. These files can be used as input to VOSViewer and thus visualized. Fig. 6, for example, shows the publication profile of the first author of this article using the “citing” map for the projection.

⁹ In the new version of VOSViewer 1.4 (September 2011), the mapping and the clustering are no longer combined. Note that the clustering is independent of the MDS since based on another (community-finding) algorithm.

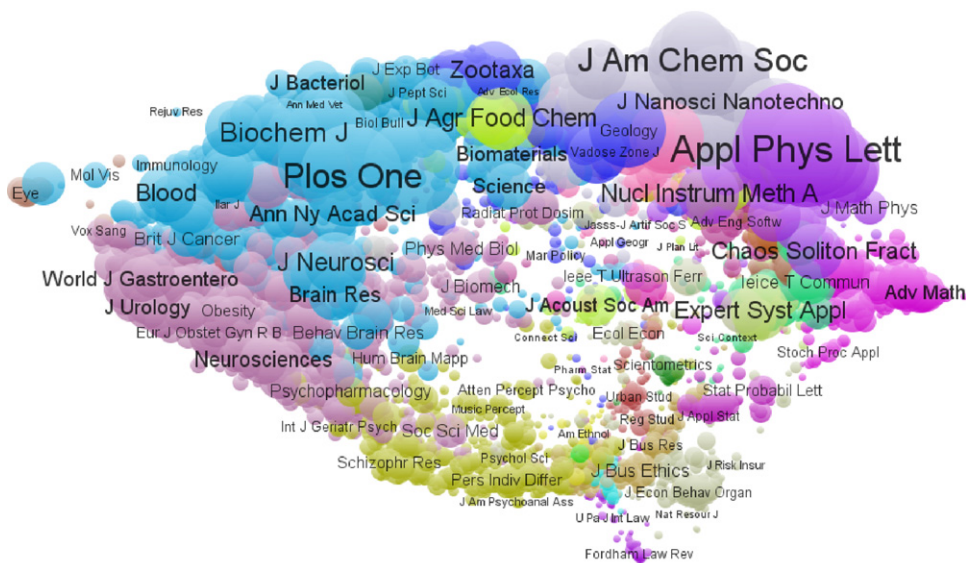


Fig. 5. 8860 journals similar in their citing patterns above $\cosine > 0.2$; 41 clusters (colors) distinguished by the rapidly unfolding algorithm of Blondel et al. (2008). (Colors adapted within VOSViewer.) One can webstart this map at <http://www.vosviewer.com/vosviewer.php?map=http://www.leydesdorff.net/journalmaps/citing/blondel.txt>.

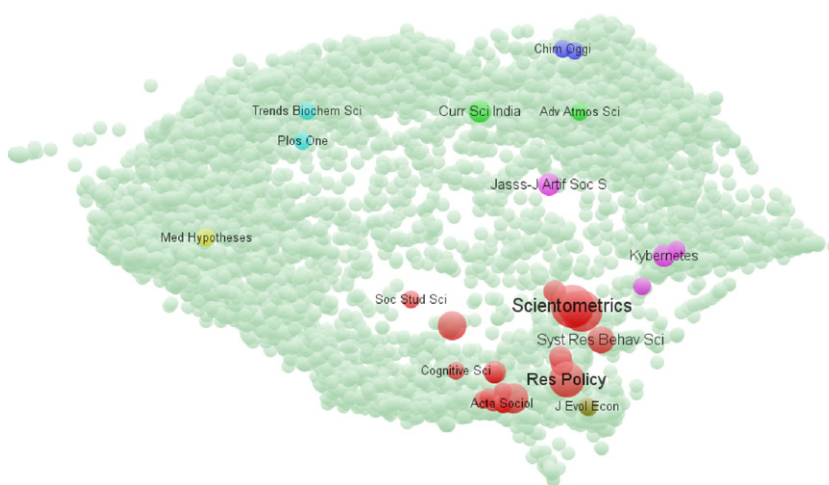


Fig. 6. Overlay map of 168 articles published by the first author (Leydesdorff) in journals included in the ISI journal set.

The colors in the overlay map correspond to the colors in the global map provided above as Fig. 4. Thereafter, all options available in VOSViewer are available to the user. Thus, in addition to this label view, one can choose for the density view or the scatter view, and the options enable one to make all labels equally visible, to change the colors of clusters, etc.

The default setting of our overlay programs presumes that the size of the nodes will be equal to the $\log_{10}(n + 1)$. The value of n is augmented by one in order to prevent the disappearance of a node in the case of a single publication (since $\log(1) = 0$). Depending on the relative sizes, one may wish to use a function other than the logarithm. The values of n can be found in the table file `overlay.dbf` which is generated at each run; this file can be read into Excel. By replacing the column labeled “weight” with the values in the column `NPubl` in the file `overlay.dbf`, one obtains a map with a linear relation between size and publication volume.

One problem is that VOSviewer, as against Pajek, contains an algorithm for relative size normalization by dividing all weights by the average weight. This algorithm has to be turned off when comparing different overlays quantitatively (e.g., in an animation). For example, when analyzing the diffusion and growth of emergent topics, one cannot compare the visual maps thus normalized, but should use the underlying quantitative data (Kiss, Broom, Craze, & Rafols, 2010; Leydesdorff &

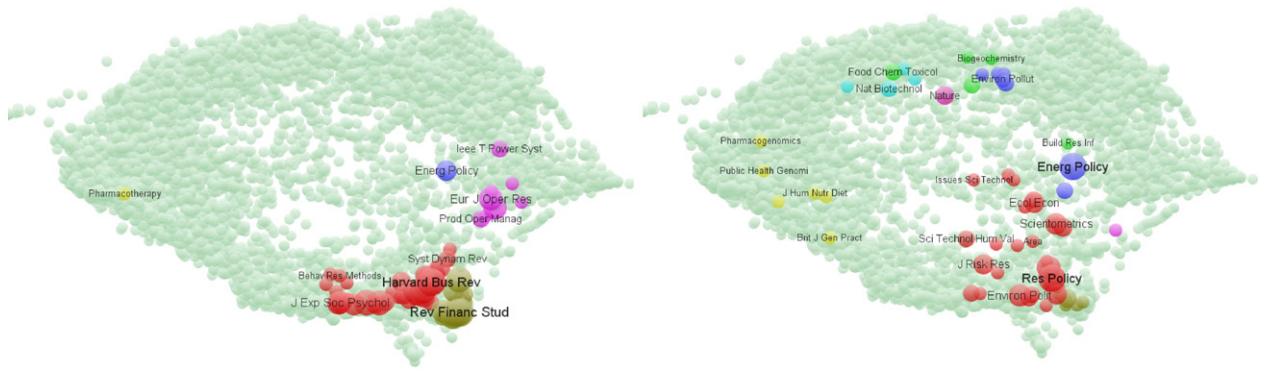


Fig. 7. (a and b) Overlay maps comparing journal publication portfolios from 2006 to 2010 between the London Business School (on the left) and the Science and Technology Policy Research Unit SPRU at the University of Sussex (on the right).

Rafols, 2011b). In response to our signaling this problem, the newest version of VOSViewer (1.4; September 2011) contains the option to replace the column label “weight” with “normalized weight” in the map file; this suppresses the additional normalization by the program itself. The user is able to change this column heading in the first lines of the files “citing.txt” and “cited.txt” after running the programs citing.exe or cited.exe, respectively, but before running VOSViewer.

As noted, the file overlay.dbf also contains the Blondel classification. If one wishes to assume this classification (or any other one) as the default for generating overlays, it is necessary to change the cluster indication in the tables cited.dbf and/or citing.dbf in this respect (in Excel or SPSS) and save these files thereafter as .dbf tables with the same name.

9. An application: mapping interdisciplinary organisations and fields

One of our motivations for developing these overlays has been the wish to evaluate interdisciplinary developments that are potentially relevant to science policy (Leydesdorff et al., 1994; Leydesdorff & Rafols, 2011a, 2011b). Let us therefore test these newly developed overlay methods in two areas of contestation about the function of interdisciplinarity: (1) the assessment of interdisciplinary units such as university departments (Rafols, Leydesdorff, O’Hare, Nightingale, & Stirling, in press), and (2) the delineation of newly emerging specialties such as nanoscience and nanotechnology (Leydesdorff & Schank, 2008).

9.1. Innovation studies versus business & management schools

In a recent study, Rafols et al. (in press) compared the research portfolios of business and management (BM) schools with those of innovation studies departments (many of which are sometimes located in business and management schools) in the case of the UK. We found that mainstream BM schools such as the London Business School focus their publications and citations narrowly in the core journals of business and economics, while departments that are oriented towards innovation studies, publish across disciplinary lines (e.g., SPRU, the Science and Technology Policy Research Unit at the University of Sussex). This analysis was also performed using overlay maps based on the ISI SCs, and provided quantitative evidence that the evaluation tends to rank mono-disciplinary output more highly than interdisciplinary sets, as qualitative studies had previously suggested (Laudel & Origgi, 2006).

Fig. 7 shows that SPRU publication (on the right side of Fig. 7) cover a diverse range of disciplines, whereas the London Business School (on the left side) has a heavily concentrated profile that focuses to a much larger extent on business and management studies journals. These representations are static. One can animate the representation by downloading data for a sequence of years, generate the map for each year, and save the screenshot into PowerPoint (Leydesdorff & Rafols, 2011b).

The comparison between these two organizations could also be made in terms of numbers. However, a table or bar chart cannot fully show the disciplinary diversity of a set, in the sense that it does not convey information about whether the journals in questions are proximate or distant in cognitive space (Leydesdorff & Rafols, 2011a, 2011b; Rafols et al., 2010). The visualization in Fig. 7 enables us to recognize at a glance differences in patterns with reference to the full journal set. In the case of journal maps, the quality of the representation is more precise and convincing than when using the SCs, since the journal names in the individual documents are used directly for the mapping. The so-called “indexer effect” in a classification can thus be avoided.

9.2. Nanoscience and nanotechnology

After the emergence of nanoscience and nanotechnology at the turn of the century (e.g., Leydesdorff & Schank, 2008), the ISI of Thomson Reuters introduced a new subject category in 2005 that contained 27 journals. In 2009, 59 journals were

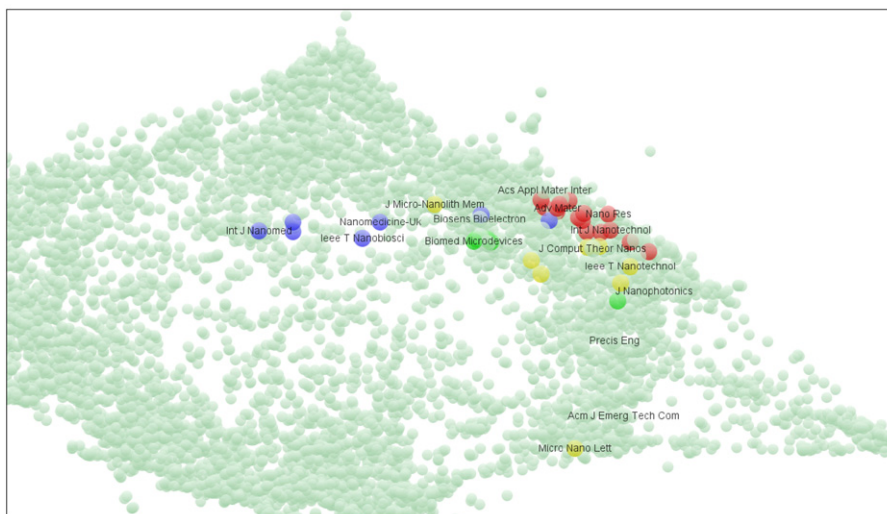


Fig. 8. Main factors in the journal set of Nanoscience and Nanotechnology, overlaid on the cited map.

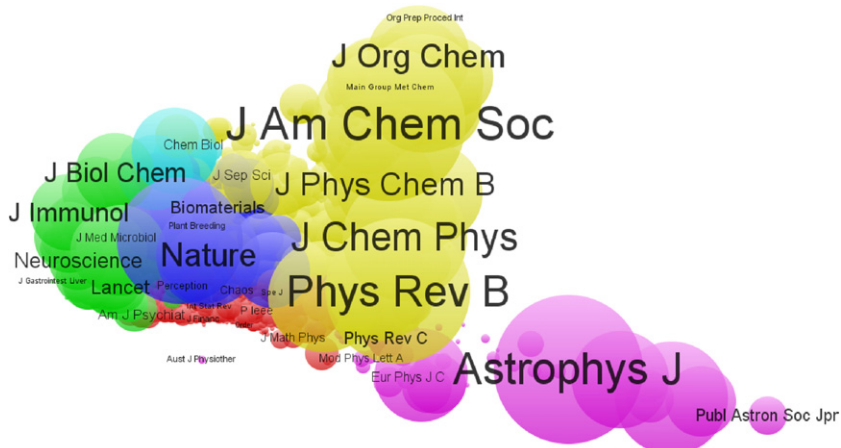


Fig. 9. Map based on the association strength among the being-cited patterns of 9162 journals in the JCR 2009 (without cosine-normalization).

subsumed under this category. How heterogeneous is this category in terms of disciplinary affiliations? The new overlay tool enables us to position these journals in relation to larger journal structures.

For this purpose we first generated the aggregated journal–journal citation matrix among 58 of these 59 journals,¹⁰ and factor-analyzed it. Four factors in the cited patterns explain 35.9% of the variance. We used these factor delineations to divide the journals into four groups: “chemistry and condensed matter physics” (20 journals with primary factor loadings), “microfluidic devices” (7 journals), “nanomedicine” (7 journals), and “electronics” (11 journals). Only journals with factor loadings larger than 0.4 on a specific factor were used to color the map in Fig. 8, for the sake of clarity.

Fig. 8 provides a visual answer to the question of whether and how closely these disciplinary-specific journals are related to the intellectual organization of the journal structures of nanoscience and nanotechnology: journals related to “nanomedicine” are colored purple (left); those related to “microfluidic devices” green (at the center); those related to “chemistry and condensed matter physics” red (upper right); and those associated with “electronics” yellow (lower right). The large distances between the four groups raises questions about whether placing all these journals under a single “interdisciplinary” category is appropriate from the perspective of science dynamics.

¹⁰ ACS Applied Materials & Interfaces was added in 2009 and is not yet covered by the JCR 2009.

10. Conclusions

The results of this exploration exceeded our expectations: VOSViewer can serve as a platform for journal overlays. One can either use the default values to size and color or else feed the tables with one's own preferences. For example, various classification schemes can be used for the journals and the nodes can be sized differently. One drawback concerned the scaling of weight relative to the average weight, but this was made optional in a new version (1.4) in the meantime (Van Eck, *personal communication*, August 29, 2011).

VOSViewer allows for uploading the results at the Internet; for example, one can webstart the basemap in the citing direction using the following hyperlink: <http://www.vosviewer.com/vosviewer.php?map=http://www.leydesdorff.net/journalmaps/citing02.txt>. The direct transport in the form of text files allows for easy reproduction and for changes in the figures.

Our exploration of the options with Gephi—the other freeware program—were less successful. A number of hurdles could be specified which may be solved in future versions of this program. However, the use of a spring-embedder was also conceptually less attractive for our project because of the different topologies involved. VOSViewer follows the MDS tradition and considers the network as a multivariate system in which individual nodes (journals) are positioned (Van Eck et al., 2010). In the case of a spring-embedder the focus is on individual relations, and the positions of the node follow as a result of the spanning of the network among the normalized links. The topology is that of the network graph. The latter perspective can perhaps be considered as an advantage in the case of focusing on local maps, because in this case the relevant contexts sometimes cannot be specified conclusively.

In general, the relations in the graph indicate where information is communicated, but the positions in the multidimensional space enable us to specify the position of the communication from a systems perspective. Furthermore, the cosine is a spatial measure well suited to map the multidimensional space. In summary, we see a great potential and no major problems with using VOSViewer for the purpose of overlays at the level of journals when the purpose is visualization.

11. Discussion

Unlike Gephi, VOSViewer does not allow for statistical analysis, but of course the same data can be fed into statistical software such as Pajek/UCInet or SPSS. We did not find the results of the clustering using VOSViewer convincing, but as noted one can replace this categorization with another one. Part of the problem may be that VOSViewer normalizes the cosine-normalized data once more using the “association strength” (Van Eck et al., 2010) or probabilistic affinity index (Zitt et al., 2000). This measure ($c_{ij}/c_i c_j$) is different from the cosine measure (used above) in that it is probabilistic.

We tested the full matrix of 9162 journal patterns with frequency values in the cited dimension. (The corresponding co-occurrence matrix can be generated by multiplication of this matrix with its transposed using Pajek or UCInet.) The results of using the association strength in this case are provided in Fig. 9.

Different from the comparable results in Fig. 3 (above), the association strength does not sufficiently normalize for size (Leydesdorff, 2008), and thus the larger journals overshadow the smaller ones in this visualization. However, if the cosine values are used without the further normalization provided by VOSViewer—an option also available—a figure is generated (not shown) in which correlations between smaller journals dominate the spanning of the dimensions of the network as axes; in this case, the resolution is strongly reduced among the remaining set in the torus-like shape. The fortunate combination of the two normalization used incidentally may yet have contributed to the unsatisfactory solution of the clustering. As noted, this classification can be replaced by a more informed one without having to change the overlays.

Acknowledgements

We are grateful to Ludo Waltman for running VOSViewer with the basemap on a large computer facility. Sébastien Heyman was so kind as to run our basemap with ForceAtlas2 in Gephi. Andrea Scharnhorst, Nees-Jan van Eck, Ludo Waltman, and two anonymous referees provided comments on previous drafts. Thomson Reuters agreed that we could use their data for generating the base maps.

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.
- Bensman, S. J. (2007). Garfield and the impact factor. *Annual Review of Information Science and Technology*, 41(1), 93–155.
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics*, 8(10), 10008.
- Bollen, J., Sompel, H. V. d., Smith, J. A., & Luce, R. (2005). Toward alternative metrics of journal impact: A comparison of download and citation data. *Information Processing and Management*, 41(6), 1419–1440.
- Borgatti, S. P. (1997). Social Network Analysis Instructional Website, at <http://www.analytictech.com/borgatti/mds.htm>.
- Börner, K. (2010). *Part two: The history of science maps*. Cambridge, MA: MIT Press.
- Börner, K., Chen, C., & Boyack, K. W. (2003). Visualizing knowledge domains. *Annual Review of Information Science and Technology*, 37(1), 179–255.
- Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics*, 64(3), 351–374.
- Burt, R. S. (1982). *Chapter 9: Toward a structural theory of action*. New York, etc.: Academic Press.
- Burt, R. S. (1995). *Chapter 1: The social structure of competition*. Cambridge, MA: Harvard University Press.

- Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. *Social Science Information*, 22(2), 191–235.
- Callon, M., Law, J., & Rip, A. (Eds.). (1986). *Chapter 7: Qualitative scientometrics*. London: Macmillan.
- Carpenter, M. P., & Narin, F. (1973). Clustering of scientific journals. *Journal of the American Society for Information Science*, 24, 425–436.
- Cook, K. A., & Thomas, J. J. (2005). *Illuminating the path: The research and development agenda for visual analytics*. Richland, WA (US): Pacific Northwest National Laboratory (PNNL). Document Number
- Doreian, P. (1986). A revised measure of standing of journals in stratified networks. *Scientometrics*, 11, 63–72.
- Doreian, P., & Fararo, T. J. (1985). Structural equivalence in a journal network. *Journal of the American Society for Information Science*, 36, 28–37.
- Egghe, L., & Leydesdorff, L. (2009). The relation between Pearson's correlation coefficient r and Salton's cosine measure. *Journal of the American Society for Information Science and Technology*, 60(5), 1027–1036.
- Fruchterman, T., & Reingold, E. (1991). Graph drawing by force-directed replacement. *Software—Practice and Experience*, 21, 1129–1166.
- Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, 178(4060), 471–479.
- Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: Sage.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.
- Hachul, S., & Junger, M. (2005). Drawing large graphs with a potential field based multilevel algorithm (extended abstract). In *GD'04: Proceedings of the 12th International Symposium on Graph Drawing*, Vol. 3383/2005 (pp. 285–295). Berlin, Heidelberg: Springer.
- Hachul, S., & Junger, M. (2006). An experimental comparison of fast algorithms for drawing general large graphs. In *GD'05: Proceedings of the 13th International Symposium on Graph Drawing*, Vol. 3843/2006 (pp. 235–250). Berlin, Heidelberg: Springer.
- Hallfman, W., & Leydesdorff, L. (2010). Is inequality among universities increasing? Gini coefficients and the elusive rise of elite universities. *Minerva*, 48(1), 55–72.
- Harel, D., & Koren, Y. (2002). Graph drawing by high-dimensional embedding. In *GD'02: Proceedings of the 10th International Symposium on Graph Drawing* (pp. 299–345). Berlin, Heidelberg: Springer.
- Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Information Processing Letters*, 31(1), 7–15.
- Kiss, I. Z., Broom, M., Craze, P. G., & Rafols, I. (2010). Can epidemic models describe the diffusion of topics across disciplines? *Journal of Informetrics*, 4(1), 74–82.
- Klavans, R., & Boyack, K. W. (2006). Identifying a better measure of relatedness for mapping science. *Journal of the American Society for Information Science and Technology*, 57(2), 251–263.
- Klavans, R., & Boyack, K. (2009). Towards a consensus map of science. *Journal of the American Society for Information Science and Technology*, 60(3), 455–476.
- Kruskal, J. B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29(2), 115–129.
- Kruskal, J. B., & Wish, M. (1978). *Chapter 1: Basic concepts of multidimensional scaling*. Beverly Hills, CA: Sage Publications.
- Latour, B. (1987). *Chapter 5: Tribunals of reason*. Milton Keynes: Open University Press.
- Laudel, G., & Origg, G. (2006). Introduction to a special issue on the assessment of interdisciplinary research. *Research Evaluation*, 15(1), 2–4.
- Leydesdorff, L. (1986). The development of frames of references. *Scientometrics*, 9(3–4), 103–125.
- Leydesdorff, L. (1987). Various methods for the mapping of science. *Scientometrics*, 11, 291–320.
- Leydesdorff, L. (1991). The static and dynamic analysis of network data using information theory. *Social Networks*, 13(4), 301–345.
- Leydesdorff, L. (2005). Similarity measures, author cocitation analysis, and information theory. *Journal of the American Society for Information Science and Technology*, 56(7), 769–772.
- Leydesdorff, L. (2006). Can scientific journals be classified in terms of aggregated journal–journal citation relations using the journal citation reports? *Journal of the American Society for Information Science and Technology*, 57(5), 601–613.
- Leydesdorff, L. (2008). On the normalization and visualization of author co-citation data: Salton's cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology*, 59(1), 77–85.
- Leydesdorff, L., & Bornmann, L. (2011). Integrated Impact Indicators (I3) compared with Impact Factors (IFs): An alternative design with policy implications. *Journal of the American Society for Information Science and Technology*, 62(7), 1370–1381.
- Leydesdorff, L., Cozzens, S. E., & Besselaar, P. v. d. (1994). Tracking areas of strategic importance using scientometric journal mappings. *Research Policy*, 23, 217–229.
- Leydesdorff, L., Hammarfelt, B., & Salah, A. A. A. (2011). The structure of the Arts & Humanities Citation Index: A mapping on the basis of aggregated citations among 1,157 journals. *Journal of the American Society for Information Science and Technology*, 62(12), 2414–2426.
- 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., de Moya-Aneón, F., & Guerrero-Bate, V. P. (2010). Journal Maps on the Basis of Scopus Data: A comparison with the Journal Citation Reports of the ISI. *Journals of the American Society for Information Science and Technology*, 61(2), 352–369.
- Leydesdorff, L., & Probst, C. (2009). The delineation of an interdisciplinary specialty in terms of a journal set: The case of communication studies. *Journal of the American Society for Information Science and Technology*, 60(8), 1709–1718.
- 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., & Rafols, I. (2011a). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*, 5(1), 87–100.
- Leydesdorff, L., & Rafols, I. (2011b). Local emergence and global diffusion of research technologies: An exploration of patterns of network formation. *Journal of the American Society for Information Science and Technology*, 62(5), 846–860.
- Leydesdorff, L., & Schank, T. (2008). Dynamic animations of journal maps: Indicators of structural change and interdisciplinary developments. *Journal of the American Society for Information Science and Technology*, 59(11), 1810–1818.
- Leydesdorff, L., & Van den Besselaar, P. (1997). Scientometrics and communication theory: Towards theoretically informed indicators. *Scientometrics*, 38, 155–174.
- Lucio-Arias, D., & Leydesdorff, L. (2009). An indicator of research front activity: Measuring intellectual organization as uncertainty reduction in document sets. *Journal of the American Society for Information Science and Technology*, 60(12), 2488–2498.
- Narin, F., Carpenter, M., & Berlt, N. C. (1972). Interrelationships of scientific journals. *Journal of the American Society for Information Science*, 23, 323–331.
- Newman, M. E. J. (2004). Fast algorithm for detecting community structure in networks. *Physical Review E*, 69(6), 066133.
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23), 8577–8582.
- Price, D. J. d. S. (1965). Networks of scientific papers. *Science*, 149(3683), 510–515.
- Pudovkin, A. I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, 53(13), 1113–1119.
- Rafols, I., & Leydesdorff, L. (2009). Content-based and algorithmic classifications of journals: Perspectives on the dynamics of scientific communication and index effects. *Journal of the American Society for Information Science and Technology*, 60(9), 1823–1835.
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (in press). How journal rankings can suppress interdisciplinarity: The case of innovation studies and business and management. *Research Policy*.
- Rafols, I., Porter, A., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9), 1871–1887.
- Rice, R. E. (1990). Hierarchies and clusters among communication and library and information science journals, 1977–1987. In C. Borgman (Ed.), *Scholarly communication and bibliometrics* (pp. 138–153). Newbury Park, CA: Sage.

- 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.
- Rosvall, M., & Bergstrom, C. T. (2010). Mapping change in large networks. *PLoS ONE*, 5(1), e8694.
- Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Chapter 1.5: How MDS differs from factor analysis*. New York/London: Academic Press.
- Small, H., & Garfield, E. (1985). The geography of science: Disciplinary and national mappings. *Journal of Information Science*, 11(4), 147.
- Studer, K. E., & Chubin, D. E. (1980). *Appendix C: A rationale and algorithm for structural analysis*. Beverly Hills, etc.: Sage.
- Theil, H. (1972). *Chapter 1.5: The entropy decomposition theorem*. Amsterdam/London: North-Holland.
- Thomas, J. J., & Cook, K. A. (2006). A visual analytics agenda. *IEEE Computer Graphics and Applications*, 10–13.
- Tijssen, R., de Leeuw, J., & van Raan, A. F. J. (1987). Quasi-correspondence analysis on square scientometric transaction matrices. *Scientometrics*, 11, 347–361.
- Van den Besselaar, P., & Leydesdorff, L. (1996). Mapping change in scientific specialties: A scientometric reconstruction of the development of artificial intelligence. *Journal of the American Society for Information Science*, 47(6), 415–431.
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSViewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538.
- Van Eck, N. J., Waltman, L., Dekker, R., & van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, 61(12), 2405–2416.
- Waltman, L., van Eck, N. J., & Noyons, E. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629–635.
- Zitt, M., Bassecoulard, E., & Okubo, Y. (2000). Shadows of the past in international cooperation: Collaboration profiles of the top five producers of science. *Scientometrics*, 47(3), 627–657.
- Zsindely, S., Schubert, A., & Braun, T. (1982). Editorial gatekeeping patterns in international science journals—A new science indicator. *Scientometrics*, 4(1), 57–68.