Science Overlay Maps: A New Tool for Research Policy and Library Management

Ismael Rafols
SPRU – Science and Technology Policy Research, University of Sussex, Brighton, BN1 9QE, England. E-mail: i.rafols@sussex.ac.uk.

Alan L. Porter
Technology Policy and Assessment Center, Georgia Institute of Technology, Atlanta, GA 30332. E-mail: alan.porter@isye.gatech.edu.

Loet Leydesdorff
Amsterdam School of Communication Research (ASCoR), University of Amsterdam, Kloveniersburg 48, 1012 CX Amsterdam, The Netherlands. E-mail: loet@leydesdorff.net

We present a novel approach to visually locate bodies of research within the sciences, both at each moment of time and dynamically. This article describes how this approach fits with other efforts to locally and globally map scientific outputs. We then show how these science overlay maps help benchmarking, explore collaborations, and track temporal changes, using examples of universities, corporations, funding agencies, and research topics. We address their conditions of application and discuss advantages, downsides, and limitations. Overlay maps especially help investigate the increasing number of scientific developments and organizations that do not fit within traditional disciplinary categories. We make these tools available online to enable researchers to explore the ongoing sociocognitive transformations of science and technology systems.

Introduction

Most science and technology institutions have undergone or are undergoing major reforms in their organization and in their activities in order to respond to changing intellectual environments and increasing societal demands for relevance. As a result, the traditional structures and practices of science, built around disciplines, are being bypassed by organizations in various ways to pursue new types of differentiation that react to diverse pressures (such as service to industry needs, translation to policy goals, openness to public scrutiny). However, no clear alternative sociocognitive structure has yet replaced the "old" disciplinary classification. In this fluid context, in which social structure often no longer matches with the dominant cognitive classification in terms of disciplines, it has become increasingly necessary for institutions to understand and make strategic choices about their positions and directions in moving cognitive spaces. "The ship has to be reconstructed while a storm is raging at sea" (Neurath, 1932/1933). The overlay map of science we present here is a technique that intends to be helpful in responding to these needs, elaborating on recently developed global maps of science (Leydesdorff & Rafols, 2009).

Although one would expect global maps of science to be highly dependent on the classification of publications, metrics, clustering algorithms, and visualization techniques used, recent studies comparing maps created using very different methods revealed that at a coarse level, these maps are surprisingly robust (Klavans & Boyack, 2009; Rafols & Leydesdorff, 2009). This stability allows "overlaying" publications or references, produced by a specific organization or research field, against the background of a stable representation of global science and producing comparisons that are visually attractive, very readable, and potentially useful for science policy making or research and library management. In this study, we present one such overlay technique and introduce its possible usages by practitioners by providing some demonstrations. For example, one can assess a portfolio at the global level or animate a diffusion pattern of a new field of research. We illustrate the former application with examples

Received December 16, 2009; revised March 16, 2010; accepted April 19, 2010

 • Published online 26 May 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.21368

from universities, industries, and funding agencies, and the latter for an emergent research topic (carbon nanotubes). In Appendix A, we provide the technical information for making these overlays using software available in the public domain.

Our first objective is to introduce the method and provide a tool for making or utilizing the global maps to prospective users in the wider science policy and research management communities who are not able to follow the developments in scientometrics in detail. Because the article addresses a wide audience, we shall not discuss technical bibliometric issues but provide references to further literature. Second, we reflect on issues about the validity and reliability of these maps. Third, this study explores the qualitative conditions of application of the maps, proposing examples of meaningful usage and flagging out potential misreadings and misunderstandings.

As classifications, maps can become embedded into working practices and turn into habits, or be taken for granted away from public debate, and yet still shape policy or management decisions that may benefit some groups at the expense of others (Bowker & Star, 2000, pp. 319–320). In our opinion, scientometric tools remain error-prone representations and fair use can be defined only reflexively. Maps, however, allow for more interpretative flexibility than rankings. By specifying the basis, limits, opportunities, and pitfalls of the global and overlay maps of science, we try to avoid the widespread problems that have beset the policy and management (mis-)use of bibliometric indicators such as the *impact factor* (Martin, 1997; Gläser & Laudel, 2007). By specifying some of the possible sources of error, we aim to set the conditions so that this novel tool remains open to critical scrutiny and can be used in an appropriate and responsible manner (Rip, 1997, p. 9).

We do not claim that these overlay maps are the most appropriate way of mapping scientific fields. The multidimensional character of science is best captured by combining various perspectives, including disparate mapping techniques. Hence, overlay maps are not a silver bullet for solving policy disputes or allocation decisions, but they can serve as a tool for gaining a limited but informed perspective on how a given issue or actor in science is related to coarse-grained disciplinary domains.

The Dissonance Between the Epistemic and Social Structures of Science

The traditional representation of science was derived from the so-called "tree of knowledge," according to which metaphor knowledge is split into branches, then into major disciplines, and further differentiated into subdisciplines and specialties. The modern universities mainly organized their social structure along this model (Lenoir, 1997), with a strong belief that specialization was key for successful scientific endeavour (Weber, 1919). However, many (if not most) scientific activities no longer align with disciplinary boundaries (Whitley, 2000; Klein, 2000; Weingart & Stehr, 2000).[1] As Lenoir (p. 53) formulated:

> Scientists at the research front do not perceive their goal as expanding a discipline. Indeed most novel research, particularly in contemporary science, is not confined within the scope of a single discipline, but draws upon work of several disciplines. If asked, most scientists would say that they work on problems. Almost no one thinks of her- or himself as working on a discipline.

The changing social contract of science has brought during the last 20 years a stronger focus on socioeconomic relevance and accountability (Gibbons, Limoges, Nowotny, Schwartzman, Scott, & Trow, 1994; Etzkowitz & Leydesdorff, 2000), which has exacerbated the dissonances between epistemic and organizational structures. Descriptions of recent transformations emphasize interdisciplinary, multidisciplinary, or transdisciplinary research as a key characteristic of the new forms of knowledge production (reviewed by Hessels & Van Lente, 2008).

These ongoing changes pose challenges to the conduct of and institutional management of science and higher education. New "disciplines" that emerged in the last decades, such as computer or cognitive sciences, do not fit neatly into the tree of knowledge. Demands for socially relevant research have also led to the creation of mission-oriented institutes and centers targeting societal problems, such as mental health or climate change, that spread (and sometimes cross-fertilize) across disciplines. At the institutional level or in evaluation, however, one cannot avoid the key question of the relative position of these emergent organizations and fields in relation to "traditional" disciplines. Can changes in research areas be measured against a baseline (Leydesdorff, Cozzens, & Van den Besselaar, 1994; Studer & Chubin, 1982)? Are the new developments transient (Gibbons et al., 1994) or, perhaps, just relabeling "old wine" (Van den Daele, Krohn, & Weingart, 1979; Weingart, 2000)? Such questions point to our endeavour: How can science overlay maps be a tool to explore the increasingly fluid and complex dynamics of the sciences? Do they allow us to throw light upon the cognitive and organizational dynamics, thereby facilitating research-related choices (e.g., funding, organization)?

Approaches to Mapping the Sciences

Science maps are symbolic representations of scientific fields or organizations in which the elements of the map are associated with topics or themes. Elements are positioned in the map so that other elements with related or similar characteristics are located in their vicinity, while those elements that are dissimilar are positioned at distant locations (Noyons, 2001, p. 84). The elements in the map can be

[1] The 'tree of knowledge' (e.g., Maturana and Varela, 1984) has strong similarities with the 'tree of life' developed by biology (via the subdicipline of systematics) to explain the diversity of species out of a common origin. Interestingly, the 'tree of life' has also been increasingly challenged by evidence of massive horizontal gene transfer among prokaryotes (Bapteste et al. 2009).

DOI: 10.1002/asi

authors, publications, institutes, scientific topics, or instruments, etc. The purpose of the representation is to enable the user to explore relations among the elements.

Science maps were developed in the 1970s (Small 1973; Small & Griffith, 1974; Small & Sweeny, 1985; Small, Sweeney, & Greenlee, 1985). They underwent a period of development and dispute regarding their validity in the 1980s (Leydesdorff, 1987; Hicks, 1987; Tijssen, de Leeuw, & van Raan, 1987) and a slow process of uptake in policy during the 1990s, which fell below the expectations created (Noyons, 2001, p. 83). The further development of network analysis during the 1990s made new and more user-friendly visualization interfaces available. Enhanced availability of data has spread the use and development of science maps during the last decade beyond the scientometrics community, in particular, with important contributions by computer scientists specialized in the visualization of information (Börner, Chen, & Boyack, 2003), as illustrated by the educative and museological exhibition, *Places and Spaces* (http://www.scimaps.org/).

Most science maps use data from bibliographic databases, such as PubMed, Thomson Reuters' Web of Science (WoS), or Elsevier's Scopus, but they can also be created using other data sources (e.g., course prerequisite structures; Balaban & Klein, 2006). Maps are built on the basis of a matrix of similarity measures computed from correlation functions among information items present in different elements (e.g., cooccurrence of the same author in various articles). The multidimensional matrices are projected onto two or three dimensions. Details of these methods are provided by Leydesdorff (1987), Small (1999), and reviewed by Noyons (2001, 2004) and Börner et al. (2003).

In principle, there are several advantages of using maps rather than relying just on numeric indicators. Maps position units in a (two-dimensional [2D]) network instead of ranking them on a (one-dimensional) list. As in any data visualization technique, maps furthermore facilitate the reading of bibliometric information by nonexperts—with the downside that they also leave room for manipulating the interpretation of data structures. Second, maps allow for the representation of diverse and large sets of data in a succinct way. Third, precisely because they make it possible to combine different types of data, maps also enable users to explore *different* views on a given issue. This interpretive flexibility induces reflexive awareness about the phenomenon the user is analysing and about the analytical value (and pitfalls) of these tools. Implicitly, science maps convey a key message: bibliometrics cannot provide definite, "closed" answers to science policy questions, such as "picking the winners." Instead, maps remain more explicitly heuristic tools to explore and potentially open up plural perspectives to inform decisions and evaluations (Roessner, 2000; Stirling, 2008).

Although the rhetoric of numbers behind indicators can easily be misunderstood as objectified and normalized descriptions of a reality (the "top-10," etc.), the heuristic, toy-like quality of interactive science maps is self-exemplifying. These considerations are important because "there is a lot of misunderstanding [by users] about the validity and utility of the maps" (Noyons, 2004, p. 238). This is compounded with a current lack of ethnographic or sociological validation of the actual use of bibliometric tools (Woolgar, 1991; Rip, 1997; Gläser & Laudel, 2007).

The vast majority of science maps have aimed at portraying *local* developments in science, using various units of analysis and similarity measures. To cite just a few techniques:

- Cocitations of articles (e.g., research on collagen; Small, 1977)
- Coword analysis (Callon, Law, & Rip, 1986), e.g., translation of cancer research (Cambrosio, Keating, Mercier, Lewison, & Mogoutov, 2007)
- Coclassification of articles (e.g., neural network research; Noyons & Van Raan, 1998)
- Cocitations of journals (e.g., artificial intelligence; Van den Besselaar & Leydesdorff, 1996)
- Cocitation of authors (e.g., information and library science; White & McCain, 1998)
- Various combinations of the previous, such as joint use of cocitation and cowords, for example, to capture temporal dynamics (Braam et al. 1991a, 1991b; Leydesdorff, 1989; Zitt & Bassecoulard, 1994).

Some of these maps already used what we here call the "overlay." This technique comprises, first, making a map based on the relations of a type of elements, and, second, "overlaying" on each element information, such as the number of articles, growth, etc., of some of the actors studied. For example, Noyons Moed, and Luwel (1999, p. 120) provided information on the relative activity of seven different institutes overlaid on a map of the subdomains of microelectronics; Boyack, Wylie, and Davidson (2002) located the position of different type of documents over a microsystem technology landscape.

The local maps are very useful for understanding the *internal* dynamics of a research field or emergent discipline, but typically they cover only a small area of science. Local maps have the advantage of being potentially accurate in their description of the relations *within* a field studied, but the disadvantage is that the units of analyses and the positional coordinates remain specific to each study. As a result, these maps cannot teach us how a new field or institute relates to other scientific areas in its (interdisciplinary?) environments. Furthermore, comparison among different developments is difficult because of the different methodological choices (thresholds and aggregation levels) used in each map.

Shared units of representation and positional coordinates are needed for proper comparisons between maps. To arrive at stable positional coordinates, a full mapping of science is needed. In summary, two requirements can be formulated as conditions for a global map of science: mapping of a full bibliographic database and robust classification of the sciences. Both requirements were computationally difficult until the last decade and mired in controversy. The next section explains how some of these controversies are in the process of being resolved and a consensus on the core structure of science is emerging.

Global Maps of Science: The Emerging Consensus

The vision that a comprehensive bibliographic database contained the structure of science was already present in the seminal contributions of Price (1965). From the 1970s, Henry Small and colleagues at the Institute of Scientific Information (ISI) started efforts to achieve a global map of science. In 1987, the ISI launched the first World Atlas of Science (Garfield, 1987) based on cocitation clustering algorithms, hence at the paper level (Small & Garfield, 1985; Small, 1999). However, the methods used (single-linked clustering) were seen as unstable and problematic (Leydesdorff, 1987). Bassecoulard and Zitt (1999) proposed a first global map based on journal clustering. A combination of new algorithms and increased computational power led to a flurry of new global science maps being developed since the mid 2000s (e.g., Moya-Anegón et al., 2004, 2007; Boyack, Klavans, & Börner, 2005; Rosvall & Bergstrom, 2008; Leydesdorff & Rafols, 2009). Klavans and Boyack (2009) provide a detailed review of these global science maps.

Given the many choices that can be made in terms of units of analysis, measures of similarity and distance, reduction of dimensions, and visualization techniques (Börner et al., 2003), most researchers in the field (including ourselves) expected any global science representation to remain heavily dependent on these methodological choices (Leydesdorff, 2006). Against these expectations, recent results of a series of global maps suggest that the basic structure of science is surprisingly robust.

First, Klavans and Boyack (2009) reported a remarkable degree of agreement in the core structure of 20 maps of science, generated by independent groups, despite different choices of unit of analysis, similarity measure, classification (or clustering algorithms), or visualization technique.[2] Then, Rafols and Leydesdorff (2009) showed that similar global maps can be obtained using significantly "dissenting" journal classifications. These validations emphasize bibliometric rather than expert assessment (Rip, 1997, p. 15), but this seems suitable in considering global science mappings, given that no experts are capable of making reliable judgement on the interrelations of all parts of science (Boyack et al., 2005, p. 359; Moya-Anegón et al, 2007, p. 2172). The consensus is more about the coarse structure of science than on final maps. The latter may show apparent discrepancies due to different choices of representation. This is the case, for example, when one compares Moya-Anegón et al.'s (2007) use of fully centric maps as opposed to Klavans and Boyack's (2008) fully circular ones.

Let us explore key features of the emerging consensus on the global structure, illustrated in Figure 1. The first feature is that science is not a continuous body, but a fragmentary structure that comprises both solid clusters and empty spaces—in geographical metaphors, a rugged landscape of high mountains and deep valleys or faults, rather than plains with rolling hills. This quasi-modular structure (or "near decomposability" in terms of the underlying [sub]systems) can be found at various levels.[3] These discontinuities are consistent with qualitative descriptions of the disunity of science (Dupré, 1993; Galison & Stump, 1996; Abbot, 2001).

A first view of Figure 1 at the global level reveals a major biomedical research pole (left side), with molecular biology and biochemistry at its center, and a major physical sciences pole (right side), including engineering, physics, and material sciences. A third pole comprises the social sciences and the humanities (bottom left).[4]

The second key feature is that the poles described above are arranged in a somewhat circular shape (Klavans & Boyack, 2009)—rather than a uniform ring, more like an uneven doughnut (a torus-like structure) that thickens and thins at different places of its perimeter. This doughnut shape can best be seen in three-dimensional (3D) representations; it is not an artefact produced by the reduction of dimensions or choice of algorithm used for the visualization. The torus-like structure of science is consistent with a pluralistic understanding of the scientific enterprise (Knorr-Cettina, 1999; Whitley, 2000): In a circular geometry, no discipline can dominate by occupying the center of science, and at the same time, each discipline can be considered as the center of its own world.

The torus-like structure explains additionally how the great disciplinary divides are bridged. Moving counter-clockwise from 3 o'clock to 10 o'clock in Figure 1 (see Figure 2 for more details), the biomedical and the physical sciences poles are connected by one bridge that reaches from material sciences to chemistry, and a parallel elongated bridge that stretches from engineering and materials to the earth sciences (geosciences and environmental technologies), and then through biological systems (ecology and agriculture) to end in the biomedical cluster. Moving from 10 o'clock to 6 o'clock, one can observe how the social sciences are strongly connected to the biomedical cluster via a bridge made by cognitive science and psychology, and a parallel bridge made by disciplines related to health services (such as occupational health and health policy). Finally, moving from 6 o'clock to 3 o'clock, we observe that the social sciences link back to the physical sciences via the weak interactions in mathematical applications and between business and computer sciences.[5]

[2]More recently developed maps also show a high degree of agreement, in spite of using very different methods such as hybrid text/citation clustering (Janssens et al. 2009) or click-stream by users of journal websites (Bollen et al. 2009).

[3]Whether and how this multi-level cluster structure is related to the power-law distributions in citations (Katz, 1999) is an issue open to debate (Leydesdorff & Bensman, 2006).

[4]The social sciences appear as a rather diffuse and small area in these science maps due to lower citation rates. However, a recent study of social sciences on their own shows a cluster as large as the natural sciences (Bollen et al, 2009).

[5]Notice that although the global map has circular symmetry, some branches develop in parallel over the torus(e.g. geosciences and chemistry). As a result, the creation of a fully uni-dimensional *wheel or circle of science* is very elegant and perhaps very useful, but it involves some distortions beyond the consensus structure (see Klavans and Boyack, forthcoming or http://www.scival.com/).

DOI: 10.1002/asi

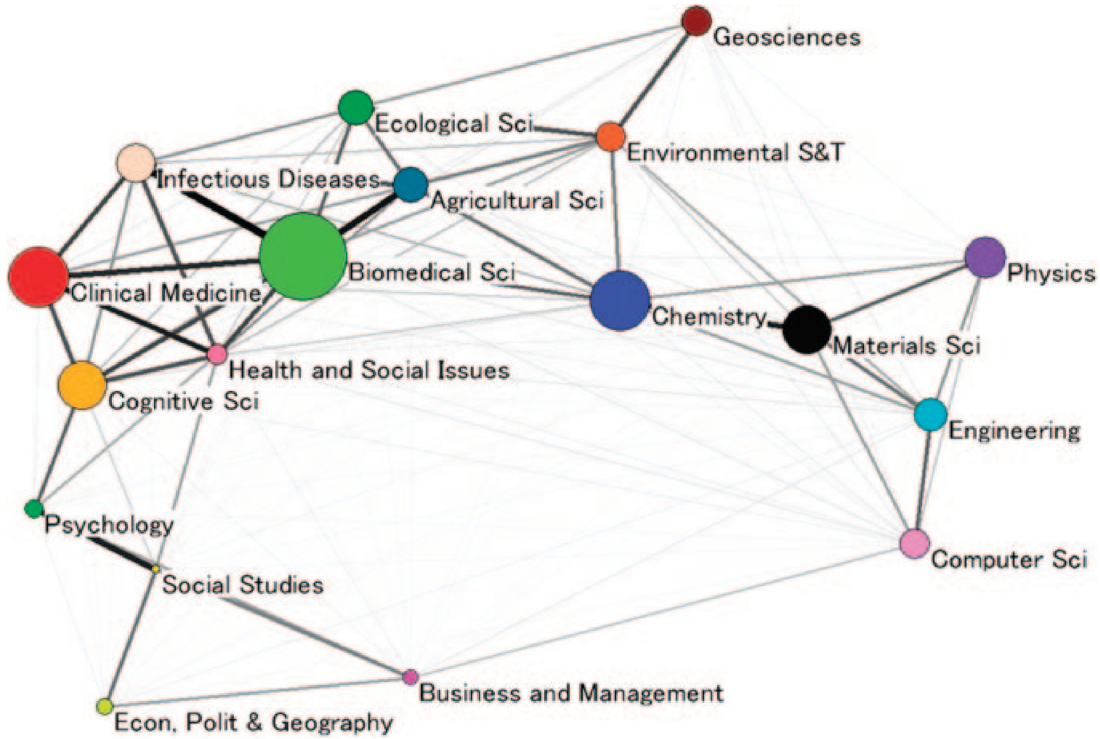

FIG. 1. The core structure of science. Cosine similarity of 18 macro-disciplines created from factor analysis of Institute of Scientific Information (ISI) subject categories in 2007. The size of nodes is proportional to number of citations produced.

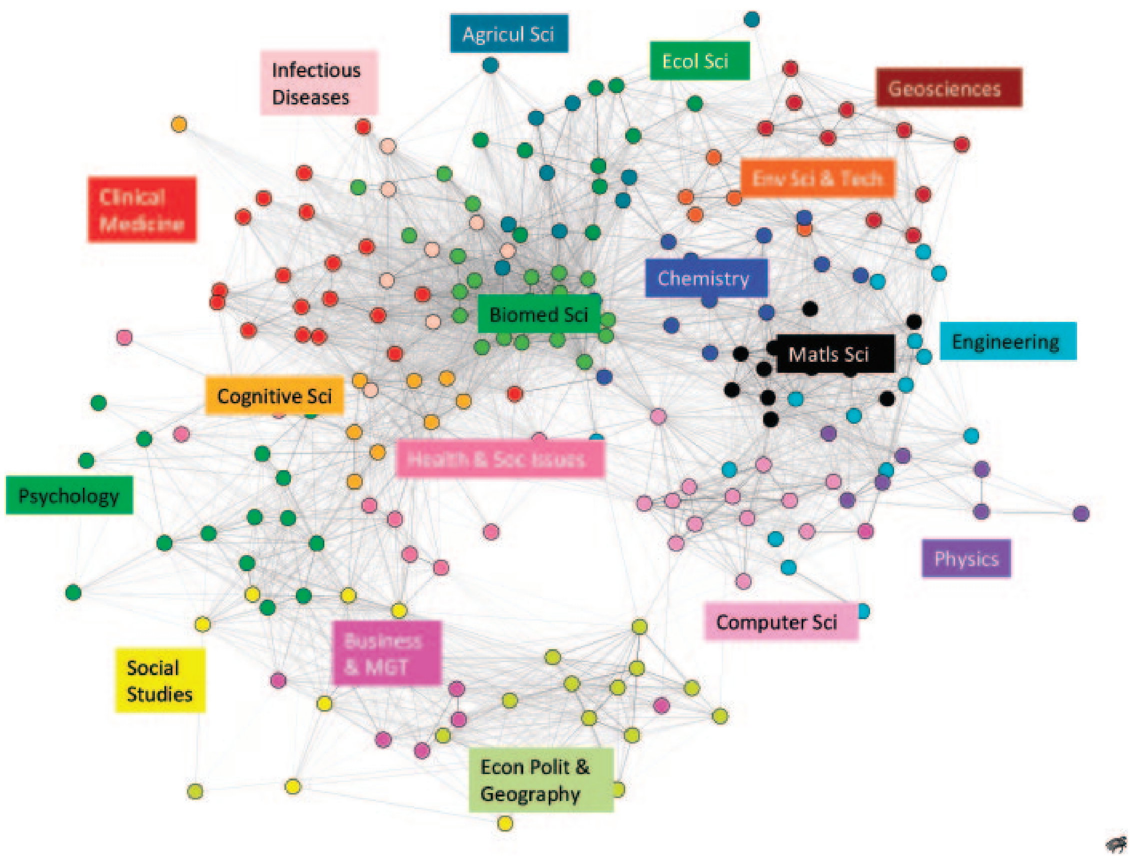

FIG. 2. Global science map based on citing similarities among ISI subject categories (2007).

DOI: 10.1002/asi

The idea behind the emergent consensus is that the most important relations among disciplines are robust—i.e., they can be elicited in the different maps even when their representations differ in many details of the global science map because of other methodological choices. However, one should not underestimate the differences among maps, particularly because they can illuminate biases. In some cases, the disagreements are mainly visual like those between geographic portrayals (e.g., in Mercator vs. Peters projections): Although there are different choices regarding the position and area size of Greenland, they all agree that Greenland lies between North America and west Eurasia. However, in some other cases, disagreements can be significant and meaningful. For example, the position of mathematics (all math subject categories) in the map remains open to debate. Since different strands of mathematics are linked to different major fields (medicine, engineering, social sciences), these may show as diverse entities in distant positions, rather than as a unitary corpus, depending on metrics, classifications, and clustering algorithms used. In controversial cases such as this, the differences between the different maps can be taken as sources of complementary understandings and fed into the discussions by experts, which are also likely to be beset with plural views.

It is important to recognize that the underlying relationships are multidimensional, so various 2D (and 3D) representations can result. For example, we depicted (in Figure 1) chemistry in the center and geosciences at the periphery, but a 3D representation would show that an opposite representation is also legitimate. Furthermore, due to reduction of dimensions, relative distances among categories need to be interpreted with caution because two categories may appear to be close without being similar. This is the case, for example, for the categories "paper and wood materials science" in relation to "paleontology" (at the top of our basemap) or "dairy and animal science" in relation to "dentistry" (top left). Categories that are only weakly linked to a few other categories are particularly prone to generate this type of positional "mirage." On the other hand, dimensional reduction also means that one can expect "tunnels," whereby hidden dimensions closely connect apparently distant spaces in the map. For example, "clinical medicine" and a small subset of engineering are connected via a slim "tunnel" made by "biomedical engineering and nuclear medicine."

In summary, the consensus on the structure of science enables us to generate and warrant a stable global template to use as a *basemap*. Several representations of this backbone are possible, legitimate, and helpful in bringing to the fore different lights and shadows. By standardizing our mapping with a convenient choice (as shown in Figure 2), we can produce comparisons that are potentially useful for researchers, science managers, or policy makers. For example, one can assess a portfolio at the global level or animate a diffusion pattern of a new field of research.

Science Overlay Maps: A Novel Tool for Research Analysis

The local science maps are problematic for comparisons because they are not stable in the units or positions of representation, as outlined in the Approaches to Mapping the Sciences section. To overcome this, one can use the units and positions derived from a global map of science, but overlay on them the data corresponding to the organizations or themes under study, as first shown by Boyack (2009). In this section, we introduce in detail a method of overlaying maps of science. This method is freely available in our Web site http://www.leydesdorff.net/overlaytoolkit.[6] A step-by-step guide on how to construct overlay maps is provided in Appendix A. A new Web site at http://idr.gatech.edu/maps will provide an interactive version of the overlay method, including detailed node labels for all the maps presented in this article.

To construct the basemap, we use the subject categories (SCs) of the WoS, to which the ISI (Thomson Reuters) assigns journals based on journal-to-journal citation patterns and editorial judgment. The SCs operationalize "bodies of specialized knowledge" (or subdisciplines) to enable one to track the position of articles. The classification of articles and journals into disciplinary categories is controversial and the accuracy of the ISI classification is open to debate (Pudovkin & Garfield, 2002, at p. 1113n). Other classifications and taxonomies are problematic as well (Rafols & Leydesdorff; 2009; NAS, 2009, p. 22). Bensman and Leydesdorff (2009) argued for using the classification of the Library of Congress, but this extension would lead us beyond the scope of this study. Despite its shortcomings, we pragmatically choose the ISI SC classification simply because it has been the most widely used and it is the most easily accessible by the potential users of the overlay map tool. Because the global maps have been shown to be relatively robust, even when there is 50% disagreement about classifications, the ISI SCs may provide reliable representations for sufficiently large data sets (see Appendix B for the statistically required data size). As discussed in Rafols and Leydesdorff (2009), we believe that classifications based on algorithmic methods (e.g., clustering) might provide a more accurate and transparent choice, but, unfortunately, these alternative full-science classifications are presently beyond the reach of many bibliometricians and policy makers.

We follow the same method outlined in Leydesdorff and Rafols (2009), inspired by Moya-Anegón et al. (2004). First, data were harvested from the CD-ROM version of the Journal Citation Reports (JCR) of the Science Citation Index (SCI) and the Social Science Citations Index (SSCI) of 2007, comprising 221 SCs. This data are used to generate a matrix of citing SCs to cited SCs with a total of 60,947,519 instances of citations among SCs. Salton's cosine was used

[6]A user-friendly toolkit using freeware *Pajek* is available at http://www.leydesdorff.net/overlaytoolkit/sc2007.zip.

DOI: 10.1002/asi

for normalization in the citing direction. *Pajek* is used for the visualizations (http://pajek.imfm.si) and SPSS (version 15) for the factor analysis. Figure 2 shows the global map of science obtained using the 221 ISI SCs in 2007. Each of the nodes in the map shows one SC, representing a subdiscipline. The lines indicate the degree of similarity (with a threshold cutoff at a *cosine* similarity > 0.15) between two SCs, with darker and thicker lines indicating stronger similarity. The relative position of the SCs is determined by the pulls of the lines as a system of strings, depending on the extent of similarity, based on the algorithm of Kamada and Kawai (1989). Although in this case we used the ISI SCs, the same method could be reproduced with other classification schemes (Rafols & Leydesdorff, 2009).

The labels and colors in Figure 2 display 18 macro-disciplines (groupings of SCs) that were obtained using factor analysis of this same matrix. The attribution of SCs to factors is listed in the file *221_SCs_2007_Citations&Similarities.xls* provided in the supplementary materials.[7] The choice of 18 factors was set pragmatically, because it was found that the 19th factor did not load strongly to its own elements. Figure 1, which we used above to illustrate the discussion on the degree of consensus, shows the core structure of science according to these 18 macro-categories.

The full map of science, shown in Figure 2, provides the basemap, over which we will explore specific organizations or scientific themes using the overlay technique. The method is straightforward. First, the analyst retrieves a set of documents at the WoS. This set of documents is the body of research to be studied, for example, the publications of an organization, the references (knowledge base) used in an emergent field, or the citations (audience) to the publications of a successful laboratory. By assigning each document to a category, the function *Analyze* provided in the WoS interface can be used to generate a list of the number of documents present in each SC. Uploading this list the visualization freeware *Pajek* produces a map of science, in which the size (area) of a node (SC) is proportional to the number of documents in that category. Full details of the procedure to generate this vector are provided in Appendix A.

Figure 3 illustrates the use of science overlay maps by comparing the profiles of three universities with distinct strengths: the University of Amsterdam, the Georgia Institute of Technology, and the London School of Economics (LSE). For each of them, the publications from 2000 to 2009 were harvested and classified into SCs in the WoS.[8] The maps show that the University of Amsterdam is an organization with a diverse portfolio and extensive research activity in clinical medicine. Georgia Tech is strong in computer sciences, materials sciences, and engineering, as well as in applications of engineering, such as biomedical or environmental technologies. Not surprisingly, LSE's main activity lies in the areas of (a) politics, economics, and geography, and (b) social studies—with some activity in the engineering and computer sciences with social science applications (e.g., statistics, information systems, or operations research) and in the health services (e.g., heath care and public health). To fully appreciate the descriptions, labels for each of the nodes are needed. Although they are not presented in these figures because of lack of resolution in printed material, they can be explored at the interactive maps at http://idr.gatech.edu/maps. Alternatively, labels can be switched on and off in the computer visualization interface Pajek, as explained in Appendix A.

Some of the advantages of overlay maps over local maps are illustrated by Figure 3. First, they provide a visual framework that enables us to make immediate and intuitively rich comparisons. Second, they use cognitive units for the representation (disciplines and specialties) that fit with conventional wisdom, whereas one can expect the analytical aggregates of local maps to be unstable and difficult to interpret. Third, whereas the generation of meaningful local maps requires bibliometric expertise, overlay maps can be produced SCI users, who are not experts in scientometrics. Finally, they can be used for various purposes depending on the units of analysis displayed by the size of the nodes, whether number of publications, citing articles, cited references, growth or other indicators, as shown by a series of recent studies (cf. Rafols & Meyer, 2010; Porter & Rafols, 2009; Porter & Youtie, 2009).

Conditions of Application of the Overlay Maps

As is the case with all bibliometric indicators, the appropriate use of overlay maps should not be taken for granted, particularly because they are tools that can be easily used by nonexperts (Gläser & Laudel, 2007). In this section, we explore the conditions under which overlay maps can be valid for science policy analysis and management, building on Rip (1997). This validation is about not just accuracy of representation but also, crucially, utility for practitioners, which depends on transparency and parsimony. Because there is generally a trade-off between accuracy, on the one hand, and transparency and parsimony, on the other, we argue that for a wide range of users, the most useful maps are not necessarily the most accurate, but are those that satisfy their needs with the most clarity and the least burden.[9]

A first issue concerns the use of journal-based ISI subject categories as the basic unit for classification. This is inaccurate because journals can be expected to combine different epistemic foci, and scientists can be expected to read sections and specific articles from different journals

[7] This matrix is available at http://www.leydesdorff.net/overlaytoolkit/SC2007.xls.

[8] Publications retrieved are as follows: 31,507 for University of Amsterdam; 26,701 for Georgia Tech; and 6,555 for London School of Economics.

[9] This is equivalent to the choice of methods in engineering: the fact the equations of relativity are more accurate than Newton's laws do not render them more useful in practice. In most cases, the gains on accuracy with relativity are negligible in comparison to the burden created by the mathematical tools required. Hence for most practical purposes, Newton's parsimony trumps relativity's accuracy.

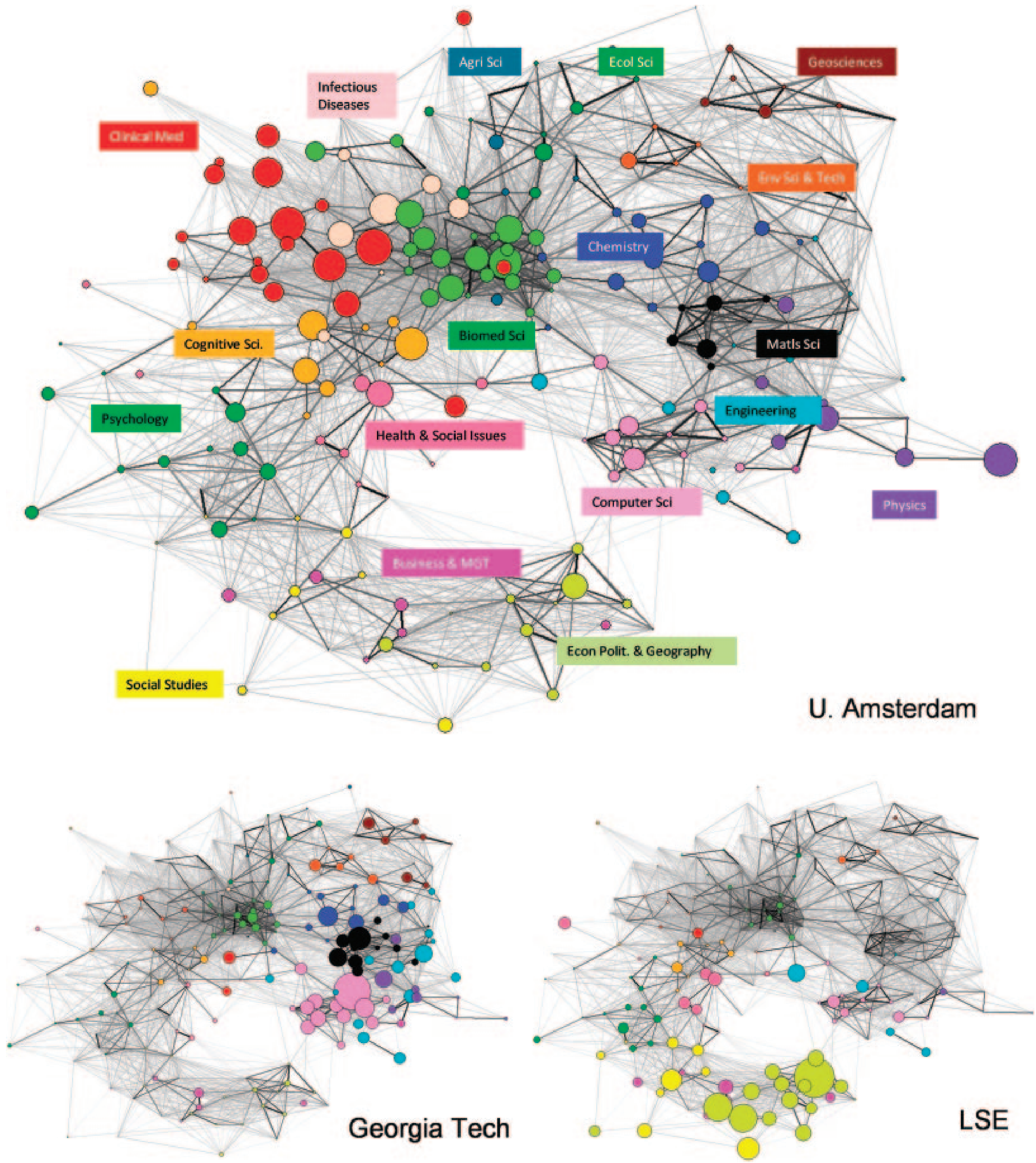

FIG. 3. Publications profiles of the University of Amsterdam, Georgia Tech, and London School of Economics (LSE) overlaid on the map of science.

(Bensman, 2007). Furthermore, journal content may not match specific disciplinary categories. In particular, consider journals such as *Nature* and *Science*, which cover multiple fields. The ISI includes these in their category, "Multidisciplinary Sciences" (which is factored into our Biomedical Sciences macro-discipline, although physics, chemistry, etc., articles appear in it). To date, we just treat this and the seven other interdisciplinary or multidisciplinary SCs (e.g., "Chemistry, Multidisciplinary") the same as any other SCs (Leydesdorff & Rafols, 2010).

On the one hand, we have to stress that the choice of ISI SCs as units of classification is a pragmatic one, determined also by their wide availability to users. On the other hand, the structural similarity of maps obtained with different

DOI: 10.1002/asi

classifications suggests that discrepancies and errors are not biased and, therefore, tend to average out when aggregated (Rafols & Leydesdoff, 2009). Hence, the answer to the problem of generalizing from specific or local classifications to a global map lies in the power of statistics: Given a sufficiently large number of assignations, there is high probability that the largest number of publications will have been assigned correctly.

For example, assuming a category with an expected correct assignation of 50%, the binomial test predicts that about 70 papers are sufficient to guarantee the correct assignation of at least 40% of the papers to this category, with a significance level of 0.05. Thus, one can rely on the laws of statistics if the error is sufficiently random and the sample size sufficiently large. Appendix B provides further details of the binomial test and estimates of the minimum size of samples under different constraints.[10] These results suggest that one should be cautious about asserting how accurately he or she is "locating" a given body of research based on small numbers of papers. Instead, for the study of single researchers or laboratories, it may be best to rely on proxies. For example, if a researcher has 30 publications, the analyst is advised to consider the set of references within these articles as a proxy for the disciplinary profile (Rafols & Meyer, 2010).

A second set of conditions for the overlay maps to be useful for research policy and management purposes is transparency and traceability, i.e., being able to specify, reproduce, and justify the procedures behind the maps in the public domain. Although the majority of the users of the map may not be interested in the scientometric details, the possibility to re-trace the methods and challenge assumptions is crucial for the maps to contribute to policy debates, where transparency is a requirement. For example, Rip (1997) noted that in the politically charged dispute regarding the "decline" of British science in the 1980s, a key issue of debate concerned the use of static versus dynamic journal categories (Irvine, Martin, Peacock, & Turner, 1985; Leydesdorff, 1988).

A further requirement for traceability is relative parsimony, that is, the rule to avoid unnecessary complexity in procedures and algorithms so that acceptable representations can be obtained by counter-expertise or even nonexperts—even at the expense of some detailed accuracy—to facilitate public discussion, if need be. In the case of overlay maps, traceability involves making publicly available the following choices: the underlying classifications used or clustering algorithms to obtain them (in our case, the ISI SC's); the similarity measures used among categories (Salton's cosine similarity); and the visualization techniques (Kamada-Kawai with a cosine > 0.15 threshold). These minimal requirements are needed so that the maps can be reproduced and validated independently.

[10]This result of at least 70 papers for each of the top categories to be identified in an overlay map is obtained under a very conservative estimate of the accuracy of existing classifications—less stringent estimates suggest that some 10–20 papers per top category may provide overlay maps with accuracy within the vicinity of a SC.

A third condition of application concerns the appropriateness of the given science overlay map for the evaluation or foresight questions that are to be answered. Roessner's (2000) critique of the indiscriminate use of quantitative indicators in research evaluation applies also to maps: Without a clear formulation of the question of what a program or an organization aims to accomplish, and its context, science maps cannot provide a well-targeted answer. What type of questions can overlay maps help to answer? We think that they can be particularly helpful for comparative purposes in benchmarking collaborative activities and looking at temporal change, as described in the next section.

Use in Science Policy and Research Management

The changes that science and technology systems are undergoing exacerbate the apparent dissonance between social and cognitive structures—with new cross-disciplinary or transversal coordinates (Whitley, 2000, p. xl ; Shinn & Ragouet, 2005). As a result, disciplinary labels of university or R&D units cannot be relied upon to provide an accurate description of their epistemic activities (Rafols & Meyer, 2007, pp. 639–640). This is because researchers often publish outside the field of their departmental affiliation (Bourke & Butler 1998) and, further, cite outside their field of publication (Van Leeuwen & Tijssen, 2000)—and increasingly so (Porter & Rafols, 2009).

Science overlay maps offer a method to locate or compare positions, shifts, and dissonances in the disciplinary activities at different institutional or thematic levels. This type of map (with a different basemap) was first introduced by Kevin Boyack and collaborators to compare the disciplinary differences in the scientific strength of nations,[11] in the publishing profiles of two large research organizations (Boyack, 2009, pp. 36–37), and the publication outcomes of two funding agencies (Boyack, Börner, & Klavans, 2009, p. 49). Some of us have used previous versions of the current overlay method to

- compare the degree of interdisciplinarity at the laboratory level (Rafols & Meyer, 2010);
- study the diffusion of a research topic across disciplines (Kiss, Broom, Craze, & Rafols, 2009);
- model the evolution over time of cross-disciplinary citations in six established research fields (Porter & Rafols, 2009); and
- explore the multidisciplinary knowledge bases of emerging technologies, namely, nanotechnology, as a field (Porter & Youtie, 2009) and specific subspecialties (Rafols, Park, & Meyer, 2010; Huang, Guo, & Porter, in press).

The following examples focus on applications for the purposes of benchmarking, establishing collaboration, and capturing temporal change, as illustrated with universities (Figure 3), large corporations (Figure 4), funding agencies

[11]http://wbpaley.com/brad/mapOfScience/index.html

DOI: 10.1002/asi

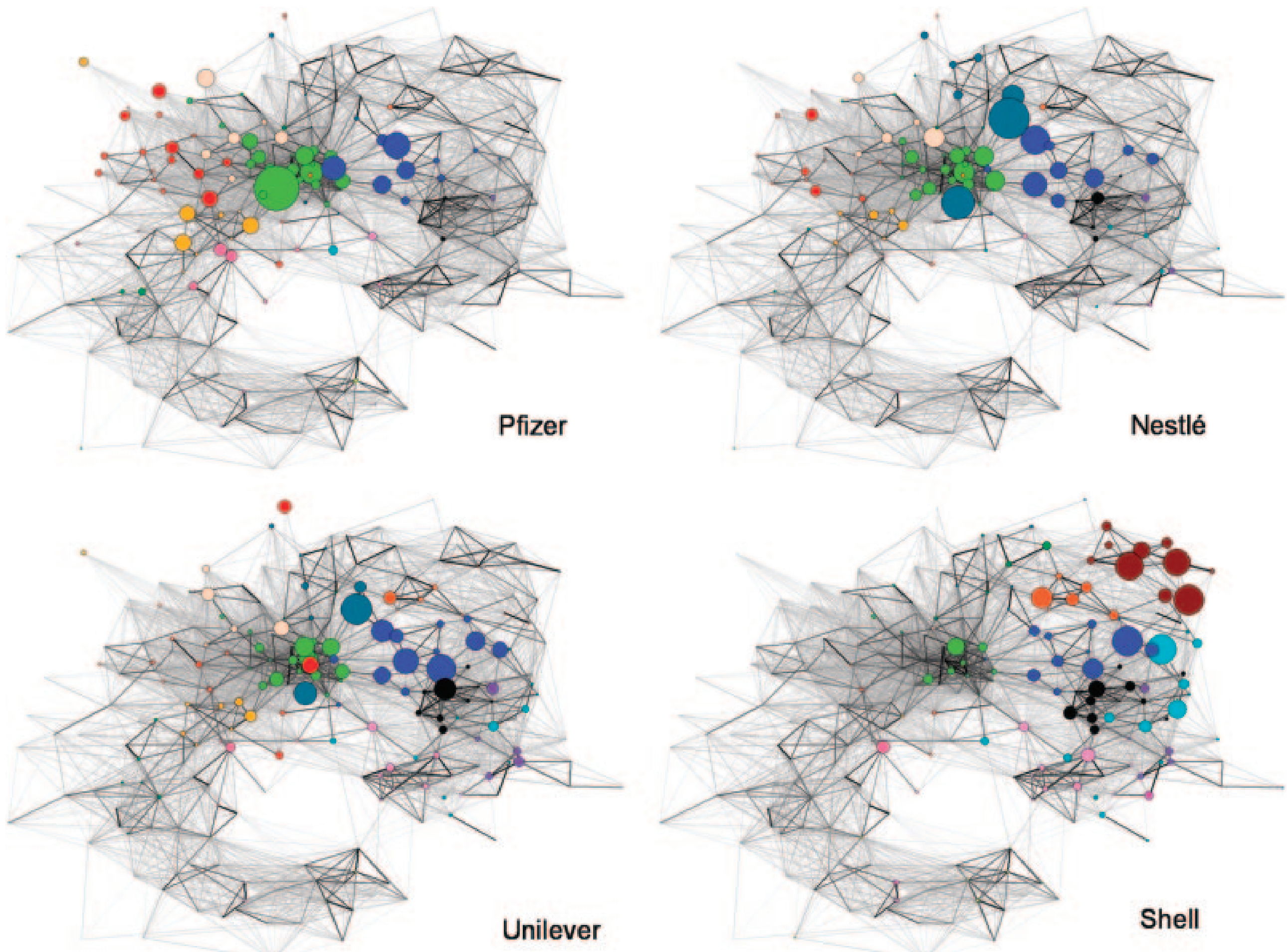

FIG. 4. Profiles of the publications (2000–2009) of large corporations of different economic sectors: pharmaceutical (Pfizer), food (Nestlé), consumer products (Unilever), and oil (Shell).

(Figure 5), and an emergent topic of research (carbon nanotubes, Figure 6).[12]

Benchmarking

A first potential use of organizational comparisons is benchmarking: How is an organization performing in comparison to possible competitors or collaborators? For example, a comparison between Pfizer (Figure 4) and Astrazeneca (not shown, see Web page), reveals at first glance a very similar profile, centered around biomedical research (pharmacology, biochemistry, toxicology, oncology) with activity in both clinical medicine and chemistry. However, a more careful look allows spotting some differences: Whereas Pfizer has a strong profile in nephrology, Astrazeneca is more active in gastroenterology and cardiovascular systems. This description may be too coarse for some purposes (e.g., specific research and development [R&D] investment), but sufficient for policy-oriented analysts to discuss the knowledge base of the firms.[13]

Several choices can be made regarding the data to be displayed in the maps. First, should the map display an input (the categories of the *papers cited by* the organizations), an output (the categories of a set of *publications* of the organization), or an outcome (the categories of the *papers citing* the organization's research)? Second, should the overlay data be normalized by the size of the category or the size of the organization? The figures here are normalized by the size of the organization, but not by the size of the category; normalizing by category will bring to the forefront those categories in which one organization is relatively very active compared with others, even if it represents a small percentage of its

[12] The data shown were retrieved from Web of Science in October, 2009. Figure 4 is based on 8107 publications by Pfizer, 1772 by Nestlé, 2632 by Unilever, and 1617 by Shell between 2000 and 2009. Figure 5 is based on 42,440 publications funded by NIH, 40,283 by NSF, 2,104 by BBSRC, and 5,746 by EPSRC, using the new field of "funding agency." Figure 6 is based on 7,782 publications on carbon-nanotubes in 2008 (cf. Lucio-Arias and Leydesdorff, 2007).

[13] Personal communication with an analyst in a pharmaceutical corporation, November 2009.

DOI: 10.1002/asi

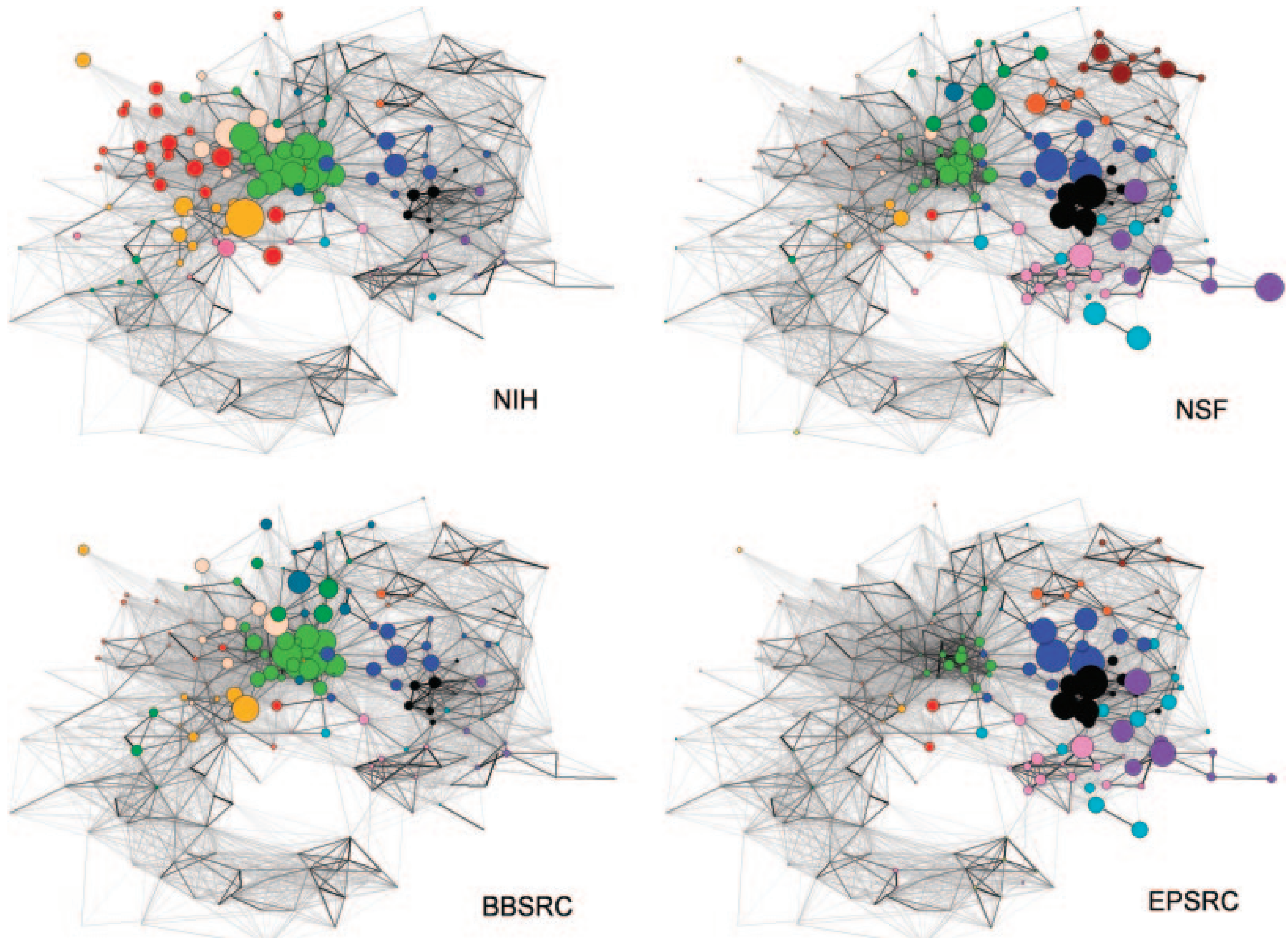

FIG. 5. Publication profile of funding agencies in the United States (top) and the United Kingdom (bottom).

production. Third, in addition to the number or proportion of publications per SC (or macro-discipline), other indicators such as impact factor or growth rate indicators can be mapped (Noyons et al., 1999; Van Raan & Van Leeuwen, 2002; or Klavans & Boyack, 2010).

Exploring Collaborations

A second application of the overlay maps is to explore complementarities and possible collaborations (Boyack, 2009). For example, Nestlé's core activities lie in food-related science and technology. Interestingly, the map reveals that one of its areas of highest research publication activity, the field of nutrition and dietetics (the dark green spot in the light green cluster in Figure 4 for Nestlé), falls much closer to the biomedical sciences than other food-related research. This suggests that the field of nutrition may act as bridge and common ground for research collaboration between the food and pharmaceutical industry—sectors that are approaching one another, as shown by Nestlé's strategic R&D investment in "functional" (i.e., health-enhancing) foods (*The Economist*, 2009).

In Figure 5, we compare funding agencies in terms of potential overlap. The funding agencies in the United States and the United Kingdom have, in principle, quite differentiated remits. In the United States (top of Figure 5), the National Institute of Health (NIH) focuses on biomedical research while the National Science Foundation (NSF) covers all basic research. In the United Kingdom (bottom of Figure 5), the Biotechnology and Biological Sciences Research Council (BBSRC) and the Engineering and Physical Sciences Research Council (EPSRC) are expected to cover the areas described in their respective names. However, Figure 5 reveals substantial areas of overlap. These are areas where duplication of efforts could be occurring, suggesting a case for coordination among agencies. It may also help indentify interdisciplinary topics warranting express collaboration between committees from two agencies, as it is the case of the BBSRC and EPSRC on the area of "Engineering and Biological Systems."

The exploration of collaboration practices is a topic in which overlay maps provide added value, because they implicitly convey information regarding the cognitive distance among the potential collaborators. A variety of studies (Llerena & Meyer-Krahmer, 2004; Cummings & Kiesler, 2005; Nooteboom et al., 2007; Rafols, 2007) have suggested that successful collaborations tend to occur in a middle range of cognitive distance, whereupon the collaborators can

DOI: 10.1002/asi

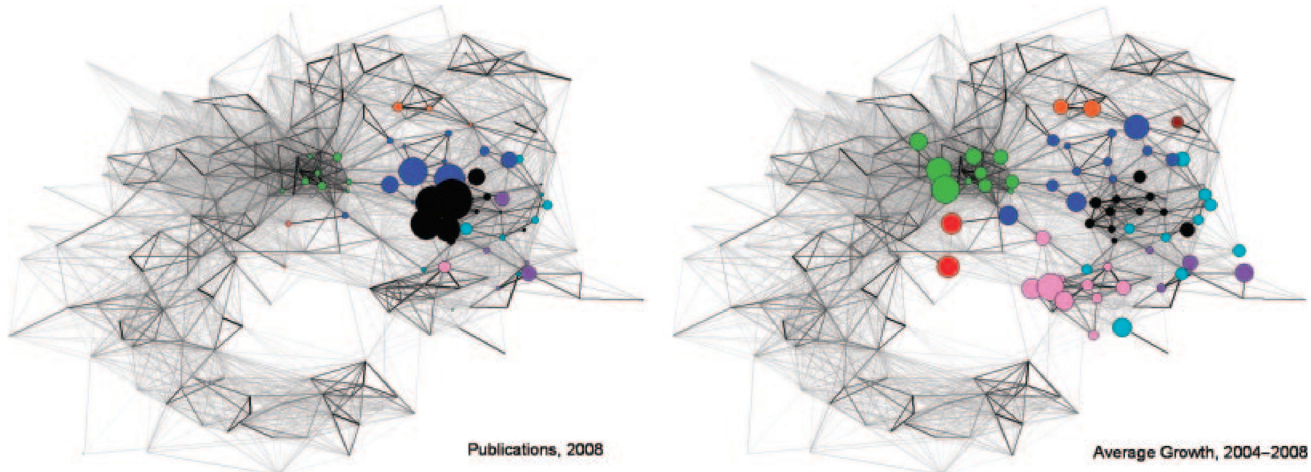

FIG. 6. Publications (2008) and average annual growth of publications (2004–2008) on carbon nanotubes.

succeed at exchanging or sharing complementary knowledge or capabilities, while still being able to understand and coordinate with one another. At short cognitive distances, the benefits of collaboration may be too low to be worth the effort (or competition may be too strong), while at large distances, understanding between partners may become difficult. It remains an empirical question whether one may think of an "optimal cognitive distance," which would allow formulating a research project with "optimal diversity" (Van den Bergh, 2008).

In any case, overlay maps offer a first (yet crude) method to explore complementarities between prospective partners. United States managers of grant programs for highly innovative research pointed out to us that the science overlay maps might be useful for finding partners, as well as for evaluating prospective grantees. The U.S. National Academies Keck Futures Initiative (NAKFI) has found it helpful to overlay research publications pertaining to a prospective workshop topic (synthetic biology) to help identify which research communities to include.

Capturing Temporal Change

A third use of overlap maps is to compare developments over time. This allows exploring the diffusion of research topics across disciplines (Kiss et al., 2009). In cases where the research topic is an instrument, a technique, or a research material, the spread may cover large areas of the science map (as noted by Price, 1984, p. 16). Figure 6 shows the location of publications on carbon nanotubes (left) and its areas of growth (right). The growth rate was computed by calculating the annual growth between 2004 and 2008 and taking the average over the period. Since their discovery in 1991, carbon nanotubes research has shown exponential growth, first in the areas of materials sciences and physical chemistry (Lucio-Arias & Leydesdorff, 2007). However, nowadays the highest growth can be observed in computer sciences due to electronic properties of carbon nanotubes (pink) in medical applications (red: e.g., imaging and biomedical engineering) and in both biomedical research (green: e.g., pharmacology and oncology) and in environmental research (orange). Within the dominant areas of chemistry and materials sciences (blue and black), growth is highest in applied fields, such as materials for textiles and biomaterials. The overlay methodology, thus, offers a perspective of the shift of carbon nanotubes research towards applications and issues of health and environmental safety. Alternatively to a static display of growth rate, the overlay maps can make a "movie" of the evolution of a field (e.g., via a succession of PowerPoint time-slice slides).

Comparison over time can also be interesting to track developments in organizations. For example, Georgia Tech, traditionally an engineering-centered university without a medical school, recently created the School of BioMedical Engineering. Going back to Figure 3, we can see a medium-size red spot in Georgia Tech publications corresponding to biomedical engineering. A dynamic analysis would depict how this has grown in the last decade.

Because the rationales of research policy, evaluation, and management are more complex than bibliometric indicators or maps can be, science overlay maps will provide complementary inputs to support (and sometimes to justify) decisions. Other possible uses include checking the match of reviewers for the assessment of interdisciplinary research in emergent fields or finding valid benchmarks when comparing organizations (Laudel & Gläser, 2009).

Advantages and Limitations of Overlays

We noted above some major advantages and downsides of overlay maps: on the plus side, their readability, intuitive, and heuristic nature, and on the minus side, the inaccuracy in the attribution to categories and the possible error by visual inspection of cognitive distance given the reduction of dimensions. In this section, we explore further potential benefits of maps in terms of cognitive contextualization and capturing diversity, and its main limitation, namely, its lack of local relational structures.

DOI: 10.1002/asi

Contextualising Categories

Science overlay maps provide a concise way to contextualise previously existing information of an organization or topic, in a cognitive space. The same information overlaid on the maps may well have been provided in many previous studies in tabular or bar chart format. For example, policy reports (e.g., Van Raan & Van Leeuwen, 2002) may extensively show the outcomes of a research programme via tables and bar charts: fields of publication, user fields, relative impacts, changes of these indicators over time, etc. What would the overlay maps offer more than this?

In our opinion, these maps provide the contextualisation of the data. This extension not only facilitates the comprehension of sets of data, but also their correct interpretation. Unlike bar charts and tables based on categories, the overlay maps remain valid (statistically acceptable) despite possible errors in the classifications. The reason is that, whereas different classifications may produce notably different bar charts, in corresponding maps 'misclassified' articles fall on nearby nodes and the user may still be provided with an adequate pattern. The context can thus reduce perceptual error.

For example, let us consider the new ISI SC of *Nanoscience and Nanotechnology*. A study of a university department in materials science during the 2000s might suggest a strong shift towards nanotechnology based on considering bar charts that show its strong growth in this new SC. However, on a global map of science, this new SC, *Nanoscience and Nanotechnology*, locates extremely closely to other core disciplines in Materials Sciences. Therefore, one would appreciate this change as a relatively small shift in focus, rather than a major cognitive shift. If a department under study had fully ventured into more interdisciplinary nanotechnology, its publications would also increasingly be visible in more disparate disciplines, for example, in the biomedical or environmental areas (Porter & Youtie, 2009).

Capturing Diversity

Science overlay maps provide the user with a perspective of the disciplinary diversity of any given output, yet without the need to rely on combined or composite indices. Research organizations often seek a diverse cognitive portfolio, but find it difficult to assess whether the intended diversity is achieved. However, diversity encapsulates three entangled aspects (variety, balance, and disparity) which cannot be univocally subsumed under a single index (Stirling, 2007), but are differently reflected in these maps:

- First, the maps capture the *variety* of disciplines by portraying the number of disciplines (nodes) in which a research organization is engaged;
- Second, they capture the disciplinary *balance* by plotting the different sizes of the SC nodes;
- Third—different from, say, bar charts—maps can convey the *disparity* (i.e. the cognitive distances) among disciplines by placing these units closer or more distant on the map (Rafols & Meyer, 2010).

This spatial elaboration of diversity measures is particularly important when comparing scientific fields in terms of multi- or interdisciplinarity. For example, Porter & Rafols (2009) show that in fields such as biotechnology, many disciplines are cited (high variety, a mean of 12.7 subject categories cited per article in 2005), but they are mainly cited in the highly dense area around biomedical sciences (low disparity). In contrast, atomic physics publications cite fewer disciplines (a mean of 8.7 per article), but from a more diverse cognitive area, ranging from physics to materials science and chemistry (higher disparity).

This discussion highlights that overlay maps are useful to explore interdisciplinary developments. In addition to capturing disciplinary diversity, they can also help to clarify the relative location of disciplines and thereby enable us to gain insights of another of the aspects of interdisciplinary research, namely their position *in between* or central (or marginal) to other research areas (Leydesdorff, 2007). Unlike indicators that seek to digest multiple facets to a single value or ranking of the extent of "interdisciplinarity," maps invite the analyst to more reflexive explorations and provide a set of perspectives that can help to open the debate. This plurality is highly commendable, given the conspicuous lack of consensus on the assessment of interdisciplinarity (Rinia, Van Leeuwen, Bruins, Van Buren, & Van Raan, 2001, Morillo, Bordons, & Gómez, 2003; Bordons, Morillo, & Gómez, 2004; Leydesdorff & Rafols, 2010; Porter, Cohen, Roessner, & Perreault, 2007; Rafols & Meyer, 2010; see review by Wagner et al., in press).

Missing the Relational Structure

The two characteristics that make overlay maps so useful for comparisons, their fixed positional and cognitive categories, are also, inevitably, their major limitations and a possible source of misreading. Because the position in the map is given only by the attribution in the disciplinary classification, the resulting map does not teach us anything about the direct linkages between the nodes. For example, Figure 3 shows that the University of Amsterdam covers many disciplines, but we do not know at all whether its local dynamics is organized within the disciplines portrayed or according to a variety of themes transversal to a collection of SCs. To investigate this, one would need to create local maps, as described in the Approaches to Mapping the Sciences section. For most local purposes, these maps will be based on smaller units of analysis, such as words, publications, or journals, rather than SCs.

In our opinion, a particularly helpful option is to combine overlay maps (based on a top-down approach, with fixed and given categories) with local maps (based on a bottom-up approach, with emergent structures) to capture the dynamics of an evolving field (Rafols & Meyer, 2010; Rafols et al., 2010; Rosvall & Bergstrom, 2009). A recursive combination of overlay and local maps allows us to investigate the evolution of a field both in terms of its *internal* cognitive coherence

DOI: 10.1002/asi

and the diversity of its knowledge sources with reference to disciplinary classifications (*external*).

Conclusions

Science overlay maps offer an intuitive way of visualizing the position of organizations or topics in a fixed map based on conventional disciplinary categories. By, thus, standardizing the mapping, one can produce comparisons that are easy to grasp for science managers or policy makers. For example, one can assess a research portfolio of a university or animate a diffusion pattern of an emergent field.

In this study, we have introduced the bases for the use of overlay maps to prospective nonexpert users and described how to create them. We demonstrated that the emergent consensus on the structure of science enables us to generate and warrant a stable global template to use as a *basemap*. We introduced the conditions to be met for a proper use of the maps, including a sample size of statistical reliability, and the requirements of transparency and traceability, and indicated potential sources of error and misinterpretations. We provided examples of use for benchmarking, searching collaborations and examining temporal change in applications to universities, corporations, funding agencies, and emergent topics.

In our opinion, overlay maps provide significant advantages in the readability and contextualization of disciplinary data and in the interpretation of cognitive diversity. The downside is that they provide only a coarse-grained perspective and miss changes in relational structure and weak interactions. Hence, we do not claim that overlay maps are *the* best method to map specific domains, but just a helpful and straightforward addition in a toolbox that should be complemented with maps that provide other, more detailed perspectives. As is the case with maps in general, overlays are more helpful than indicators to accommodate reflexive scrutiny and plural perspectives. Given the potential benefits of using overlay maps for research policy, we provide the reader with an interactive Web page to explore overlays (http://idr.gatech.edu/maps) and a freeware-based toolkit (available at http://www.leydesdorff.net/overlaytoolkit).

Acknowledgments

We are very grateful to Ziyo Wang for her enthusiastic design of the interactive Web pages. We are indebted to K. Boyack, R. Klavans, F. de Moya-Anegón, E. Noyons, B. Vargas-Quesada, A. Stirling, and M. Zitt for insights and discussions. Alan P. Porter and Ismael Rafols acknowledge support from the US National Science Foundation (Award no. 0830207, "Measuring and Tracking Research Knowledge Integration," http://idr.gatech.edu/). The findings and observations contained in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation. Ismael Rafols is partially supported by the EU FP7 project FRIDA, grant 225546.

References

Abbot, A. (2001). Chaos of disciplines. Chicago: The University of Chicago Press.

Balaban, A.T., & Klein, D.J. (2006). Is chemistry "the central science"? How are different sciences related? Cocitations, reductionism, emergence, and posets. Scientometrics, 69(3), 615–637.

Bapteste, E., O'Malley, M.A., Beiko, R.G., Ereshefsky, M., Gogarten, J.P., & Franklin-Hall, L., et al. (2009). Prokaryotic evolution and the tree of life are two different things. Biology Direct, 4(34), 1–20.

Bassecoulard, E., & Zitt, M. (1999). Indicators in a research institute: A multilevel classification of journals. Scientometrics, 44(3), 323–345.

Bensman, S.J. (2007). Garfield and the impact factor. Annual Review of Information Science and Technology, 41(1), 93–155.

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.

Bollen, J., Van de Sompel, H., Hagberg, A., Bettencourt, L., Chute, R., Rodriguez, M.A. et al. (2009). Clickstream data yields high-resolution maps of science. PLoS ONE 4(3), e4803.

Bordons, M., Morillo, F., & Gómez, I. (2004). Analysis of cross-disciplinary research through bibliometric tools. In H.F. Moed, W. Glänzel, & U. Schmoch (Eds.), Handbook of quantitative science and technology research (437–456). Dordrecht, the Netherlands: Kluwer.

Börner, K., Chen, C., & Boyack, K.W. (2003). Visualizing knowledge domains. Annual Review of Information Science & Technology, 37, 179–255.

Bourke, P., & Butler, L. (1998). Institutions and the map of science: Matching university departments and fields of research. Research Policy, 26(6), 711–718.

Bowker, G.C., & Star, S.L. (2000). Sorting things out: Classification and its consequences. Cambridge, MA: MIT Press.

Boyack, K.W. (2009). Using detailed maps of science to identify potential collaborations. Scientometrics, 79(1), 27–44.

Boyack, K.W., Börner, K., & Klavans, R. (2009). Mapping the structure and evolution of chemistry research. Scientometrics, 79(1), 45–60.

Boyack, K.W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. Scientometrics, 64(3), 351–374.

Boyack, K.W., Wylie, B.N., & Davidson, G.S. (2002). Domain visualization using VxInsight® for science and technology management. Journal of the American Society for Information Science and Technology, 53(9), 764–773.

Braam, R.R., Moed, H.F., & van Raan, A.F.J. (1991a). Mapping of science by combined co-citation and word analysis: I. Structural aspects. Journal of the American Society for Information Science, 42(4), 233–251.

Braam, R.R., Moed, H.F., & van Raan, A.F.J. (1991b). Mapping of science by combined co-citation and word analysis: II. Dynamical aspects. Journal of the American Society for Information Science, 42(4), 252–266.

Callon, M., Law, J., & Rip, A. (Eds.) (1986). Mapping the dynamics of science and technology: Sociology of science in the real world. London: MacMillan.

Cambrosio, A., Keating, P., Mercier, S., Lewison, G., & Mogoutov, A. (2007). Mapping the emergence and development of translational cancer research. European Journal of Cancer, 42(18), 3140–3148.

Cummings, J.N., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. Social Studies of Science, 35(5), 733–722.

Dupré, J. (1993). The disorder of things. metaphysical foundations of the disunity of science. Cambridge, MA: Harvard University Press.

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From national systems and "mode 2" to a triple helix of university-industry-government relations. Research Policy, 29(2), 109–123.

Galison, P., & Stump, D.J. (Eds.) (1996). The disunity of science: boundaries, context and power. Stanford, CA: Stanford University Press.

Garfield, E. (1987). Launching the ISI Atlas of Science: For the new year, a new generation of reviews. Current Contents, 1, 1–6.

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.

Gläser, J., & Laudel, G. (2007). The social construction of bibliometric evaluations. In R. Whitley & J. Gläser, The changing governance of the sciences (pp. 101–123). Amsterdam: Springer.

Hessels, L.K., & Van Lente, H. (2008). Re-thinking new knowledge production: A literature review and a research agenda. Research Policy, 37(4), 740–760.

Hicks, D. (1987). Limitations of co-citation analysis as a tool for science policy. Social Studies of Science, 17(2), 295–316.

Huang, L., Guo, Y., & Porter, A.L. (forthcoming). Identifying emerging roles of nanoparticles in biosensors, Journal of Business Chemistry.

Irvine, J., Martin, B., Peacock, T., & Turner, R. (1985). Charting the decline of British science. Nature, 316, 587–590.

Janssens, F., Zhang, L., Moorc, B.D., & Glänzel, W. (2009). Hybrid clustering for validation and improvement of subject-classification schemes. Information Processing and Management, 45(6), 683–702.

Kamada, T., & Kawai, S. (1989). An algorithm for drawing general undirected graphs. Information Processing Letters, 31(1), 7–15.

Katz, J.S. (1999). The self-similar science system. Research Policy, 28(5), 501–517.

Kiss, I.Z., Broom, M., Craze, P., & Rafols, I. (2010). Can epidemic models describe the diffusion of research topics across disciplines? Journal of Informetrics, 4(1), 74–82.

Klavans, R., & Boyack, K.W. (2008). Thought leadership: A new indicator for national and institutional comparison. Scientometrics, 75(2), 239–250.

Klavans, R., & Boyack, K.W. (2009). Toward a consensus map of science. Journal of the American Society for Information Science and Technology, 60(3), 455–476.

Klavans, R., & Boyack, K.W. (forthcoming). Toward an objective, reliable and accurate method for measuring research leadership. Scientometrics.

Klein, J.T. (2000). A conceptual vocabulary of interdisciplinary science. In P. Weingart & N. Stehr (Eds.), Practising interdisciplinarity (pp. 3–24). Toronto: University of Toronto Press.

Knorr-Cetina, K. (1999). Epistemic cultures. How the sciences make knowledge. Cambridge, MA: Harvard University Press.

Laudel, G., & Gläser, J. (2009, July). Interdisciplinarity as criterion and condition of evaluations. Paper presented at International Society for Scientometrics and Informetrics Conference, Rio de Janeiro.

Lenoir, T. (1997). Instituting science. The cultural production of scientific disciplines. Stanford, CA: Stanford University Press.

Leydesdorff, L. (1987). Various methods for the mapping of science. Scientometrics, 11(5–6), 291–320.

Leydesdorff, L. (1988). Problems with the 'measurement' of national scientific performance. Science and Public Policy, 15, 149–152.

Leydesdorff, L. (1989). Words and co-words as indicators of intellectual organization. Research Policy, 18(4), 209–223.

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 & Technology, 57(5), 601–613.

Leydesdorff, L. (2007). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. Journal of the American Society for Information Science & Technology, 58(9), 1303–1319.

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., & Van den Besselaar, P. (1994). Tracking areas of strategic importance using scientometric journal mappings. Research Policy, 23(2), 217–229.

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. Journal of the American Society for Information Science & Technology, 60(2), 348–362.

Leydesdorff, L., & Rafols, I. (2010). Indicators of the interdisciplinarity of journals. Diversity, Centrality, and Citations. Manuscript submitted for publication.

Llerena, P., & Meyer-Krahmer, F. (2004). Interdisciplinary research and the organization of the university: General challenges and a case study. Science and innovation. In A. Geuna, A.J. Salter, & W.E. Steinmueller (Eds.), Rethinking the rationales for funding and governance (pp. 69–88). Cheltenham, United Kingdom: Edward Elgar.

Lucio-Arias, D., & Leydesdorff, L. (2007). Knowledge emergence in scientific communication: From "fullerenes" to "nanotubes." Scientometrics, 70(3), 603–632.

Martin, B.R. (1997). The use of multiple indicators in the assessment of basic research. Scientometrics, 36(3), 343–362.

Maturana, H.R., & Varela, F.J. (1984). The tree of knowledge. Boston: New Science Library.

Morillo, F., Bordons, M., & Gómez, I. (2003). Interdisciplinarity in science: A tentative typology of disciplines and research areas. Journal of the American Society for Information Science and Technology, 54(13), 1237–1249.

Moya-Anegón, F., Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Munoz-Fernández, F.J., & Herrero-Solana, V. (2007). Visualizing the marrow of science. Journal of the American Society for Information Science & Technology, 58(14), 2167–2179.

Moya-Anegón, F., Vargas-Quesada, B., Victor Herrero-Solana, Chinchilla-Rodríguez, Z., Elena Corera-Álvarez, & Munoz-Fernández, F.J. (2004). A new technique for building maps of large scientific domains based on the cocitation of classes & categories. Scientometrics, 61(1), 129–145.

NAS Keck Center. (2009). Data on federal research and development investments: A pathway to modernization. Washington DC: National Academies Press. Retrieved November 20, 2009, from http://www.nap.edu/catalog/12772.html

Neurath, O. (1932/1933). Protokollsätze. Erkenntnis [Protocol statements]. 3, 204–214.

Nooteboom, B., Haverbeke, W.V., Duysters, G., Gilsing, V., & van den Oord, A. (2007). Optimal cognitive distance and absorptive capacity. Research Policy, 36, 1016–1034.

Noyons, E. (2001). Bibliometric mapping of science in a science policy context. Scientometrics, 50(1), 83–98.

Noyons, E.C.M. (2004). Science maps within a science policy context. In H.F. Moed, W. Glänzel, & U. Schmoch (Eds.), Handbook of quantitative science & technology research (pp. 237–255). Dordrecht: Kluwer.

Noyons, E.C.M., Moed, H., & Luwel, M. (1999). Combining mapping and citation analysis for evaluative bibliometric purposes. Journal of the American Society for Information Science, 50(2), 115–131.

Noyons, E.C.M., & Van Raan, A.F.J. (1998). Monitoring scientific developments from a dynamic perspective: Self-organized structuring to map neural network research. Journal of the American Society for Information Science, 49(1), 68–81.

Porter, A.L., Cohen, A.S., Roessner, J.D., & Perreault, M. (2007). Measuring researcher interdisciplinarity. Scientometrics, 72(1), 117–147.

Porter, A.L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. Scientometrics, 81(3), 719–745.

Porter, A.L., & Youtie, J. (2009). Where does nanotechnology belong in the map of science? Nature-Nanotechnology, 4(9), 534–536.

Price, D.J. de Solla (1965). Networks of scientific papers. Science, 149(3682), 510–515.

Price, D.J. de Solla (1984). The science/technology relationship, the craft of experimental science, and policy for the improvement of hight technology innovation. Research Policy, 13(1), 3–20.

Pudovkin, A.I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. Journal of the American Society for Information Science and Technology, 53(11), 1113–1119.

Rafols, I. (2007). Strategies for knowledge acquisition in bionanotechnology: Why are interdisciplinary practices less widespread than expected? Innovation: the European Journal of Social Science Research, 20(4), 395–412.

Rafols, I., & Meyer, M. (2007). How cross-disciplinary is bionanotechnology? Explorations in the specialty of molecular motors. Scientometrics, 70(3), 633–650.

DOI: 10.1002/asi

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 & Technology, 60(9), 1823–1835.
Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. Scientometrics, 82(2), 263–287.
Rafols, I., Park, J.-H., & Meyer, M. (2010). Hybrid nanomaterial research: Is it really interdisciplinary? In K. Rurack & R. Martínez- Máñez (Eds.), The supramolecular chemistry of organic- and inorganic hybrid materials. Hoboken, NJ: John Wiley & Sons.
Rinia, E.J., Van Leeuwen, T.N., Bruins, E.P.W., Van Buren, H.G., & Van Raan, A.F.J. (2002). Measuring knowledge transfer between fields of science. Scientometrics, 54(3), 347–362.
Rip, A. (1997). Qualitative conditions of scientometrics: The new challenges. Scientometrics, 38(1), 7–26.
Roessner, D. (2000). Quantitative and qualitative methods and measures in the evaluation of research. Research Evaluation, 9(2), 125–132.
Rosvall, M., & Bergstrom, C.T. (2009). Mapping change in large networks. PLoS ONE, 5(1), e8694.
Shinn, T., & Ragouet, P. (2005). Controverses sur la science. Pour une sociologie transversaliste de l'activité scientifique [Controversies on science for a transversal sociology of scientific activity]. Paris: Raisons d'agir.
Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. Journal of the American Society for Information Science, 24(4), 265–269.
Small, H. (1977). A co-citation model of a scientific specialty: A longitudinal study of collagen research. Social Studies of Science, 7, 139–166.
Small, H. (1999). Visualizing science by citation mapping. Journal of the American Society for Information Science, 50(9), 799–813.
Small, H., & Griffith, B. (1974). The structure of scientific literature: I. Science studies, 4(1), 17–40.
Small, H., & Sweeney, E. (1985). Clustering the Science Citation Index using co-citations: I. A comparison of methods,. Scientometrics, 7(5–6), 391–409.
Small, H., Sweeney, E., & Greenlee, E. (1985). Clustering the Science Citation Index using co-citations: II. Mapping science. Scientometrics, 8(5–6), 321–340.
Stirling, A. (2007). A general framework for analysing diversity in science, technology and society. Journal of The Royal Society Interface, 4(15), 707–719.
Stirling, A. (2008). "Opening up" and "closing down": Power, participation, and pluralism in the social appraisal of technology. Science, Technology & Human Values, 33(2), 262–294.
Studer, K.E., & Chubin, D.E. (1980). The cancer mission. Social contexts of biomedical research. Beverly Hills, CA: Sage.
The Economist, (2009, October 29). Nestlé. The unrepentant chocolotier.
Tijssen, R., de Leeuw, J., & van Raan, A.F.J. (1987). Quasi-correspondence analysis on square scientometric transaction matrices. Scientometrics, 11(5–6), 347–361.
Van den Bergh, J.C.J.M. (2008). Optimal diversity: Increasing returns versus recombinant innovation. Journal of Economic Behavior & Organization, 68(3–4), 565–580.
van den Besselaar, P., & Leydesdorff, L. (1996). Mapping change in scientific specialties: A scientometric reconstruction of the development of artificial intellingence. Journal of the American Society for Information Science, 46(6), 415–436.
Van den Daele, W., Krohn, W., & Weingart, P. (Eds.). (1979). Geplante Forschung: Vergleichende Studien über den Einfluss politischer Programme auf die Wissenschaftsentwicklung [Planned research: Comparative studies of the influence of political programmes on the development of science]. Frankfurt, Germany: Suhrkamp.
Van Leeuwen, T., & Tijssen, R. (2000). Interdisciplinary dynamics of modern science: Analysis of cross-disciplinary citation flows. Research Evaluation, 9(3), 183–187.
Van Raan, A.F.J., & van Leeuwen, T.N. (2002). Assessment of the scientific basis of interdisciplinary, applied research. Application of bibliometric methods in nutrition and food research. Research Policy, 31(4), 611–632.
Wagner, C.S., Roessner, J.D., Bobb, K., Klein, J.T., Boyack, K.W., & Keyton, J., et al. (in press). Approaches to understanding and measuring Interdisciplinary Scientific Research (IDR): A review of the literature. Journal of Informetrics.
Weingart, P. (2000). Interdisciplinarity: The paradoxical discourse. In P. Weingart & N. Stehr (Eds.), Practising interdisciplinarity (pp. 25–41). Toronto: University of Toronto Press.
Weingart, P., & Stehr, N. (Eds.) (2000). Practising interdisciplinarity. Toronto: University of Toronto Press.
White, H.D., & McCain, K.W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. Journal of the American Society for Information Science, 49(9), 327–355.
Whitley, R. (2000). The intellectual and social organization of the sciences (2nd ed.). Oxford: Oxford University Press.
Woolgar, S. (1991). Beyond the citation debate: Towards a sociology of measurement technologies and their use in science policy. Science and Public Policy, 18(5), 319–326.
Zitt, M., & Bassecoulard, E. (1994). Development of a method for detection and trend analysis of research fronts built by lexical or cocitation analysis Scientometrics, 30(1), 333–351.

Appendix A

A User-Friendly Method for the Generation of Overlay Maps

We follow the method introduced in Rafols and Meyer (2010) to create the overlay map on the basis of a global map of science (Leydesdorff & Rafols, 2009). The steps described below rely on access to the WoS and the files available in our mapping kit. The objectives are to obtain the set of SCs for a given set of articles, provide this to network software (we describe for Pajek), and output as overlay information to add to a suitable basemap.

First, the analyst has to conduct a search in the Thomson Reuters WoS (www.isiknowledge.com). Nonexpert users should note that this initial step is crucial and should be done carefully: authors may come with different initials, addresses are often inaccurate, and only some types of documents may be of interest (e.g., only so-called citable items: *articles*, *proceedings papers*, *reviews*, and *letters*). Once the analyst has chosen a set of documents from searches at WoS, one can click the tab, *Analyze results*. In this new Web page, the selected document set can then be analyzed along various criteria (top left-hand tab). The *Subject Area* choice produces a list with the number of documents in each Subject Category. This list can be downloaded as *Analyze.txt*.

In the next step, the analyst can go to our Web page for maps (http://idr.gatech.edu/maps) and upload this file. If one desires more control on the process, one can use the program Pajek and the associated overlay toolkit at http://www.leydesdorff.net/overlaytoolkit. After opening Pajek, press F1 and upload the basemap file *SC2007-015cut-2D-KK.paj*. This file provides the basemap,[14] as shown by selecting *Draw>Draw-Partition-Vector* (or pressing *Ctrl-P*). Then, the previously downloaded *Analyze.txt* file has to be

[14]The matrix underlying the basemap and the grouping of SCs is available at: http://www.leydesdorff.net/overlaytoolkit/sc2007.xls

transformed by the mini-program *SC2007.exe* (in our tool kit) as into the Pajek vector format "*SC07.vec*." This file can be uploaded into Pajek by choosing *File>Vector>Read* from the main Pajek menu. Selecting from the menu *Draw>Draw-Partition-Vector* (alternatively, pressing *Ctrl-Q*), the overlay map will be generated.

At this stage, the size of nodes will often need adjustment, which can be done by selecting *Options>Size of Vertices* in the new draw window. To have the *standard* colour settings, the file *SC2007-18Factors-ColourSettings.ini* can be loaded by going to *Options>Ini File>Load* in the main Pajek window. *Crtl-L* and *Ctrl-D* allow visualize and delete, respectively, the labels of each SC. Selecting nodes moves SCs to other positions. The image can be exported selecting *Export>2D>Bitmap* in the menu of the *Draw* window. Alternatively, a manual for embellishing one's Pajek output can be found at http://www.leydesdorff.net/indicators/lesson6.htm. A further optional step would be to label the map in terms of factors by importing this image into PowerPoint to label groups of clusters, as shown in the file *SC2007 Global maps.ppt*.

An alternative procedure for more experienced users is to download the records of a document set found in the WoS. This is done by adding the desired documents to the *Marked list* (bottom bar), second, going to *Marked list* (top bar), and then downloading the documents in a *Tagged Field* format after selecting *Subject category* as one of the fields to download. The downloaded file should be renamed as *data.txt* and used as input into the program ISI.exe (available at http://www.leydesdorff.net/software/isi). One of the outputs of the programs ISI.exe is the file *SC07.vec*, which can be used in Pajek as explained above. The advantage of this procedure is that ISI.exe also produces other files with information on fields such as author or journal that may be of interest. Feel free to contact the authors in case of difficulty.

Appendix B

Estimation of Number of Papers Needed for Reliability in Overlay Maps

In a previous study, Rafols and Leydesdorff (2009) found that there is between 40%–60% of disagreements between attributions of journals to disciplinary categories. Taking a conservative approach, let us assume that for a sample of N papers, there is a probability $p = 0.5$ that they will be misclassified by a given classification (whichever one is used). How large should a sample of papers be so that in spite of the error, the largest categories in the distribution correctly represent the core discipline of the population?

Let us then assume that we have N papers of one given category A. Given the $p = 0.5$ probability of correct assignation, we expect only 50% of the papers in category A. The analyst has then to arbitrarily choose a lower threshold m (we suggest 40%) as the minimum percentage acceptable, with a given degree of significance (we suggest $\sigma = 0.05$, corresponding to a z-score of 1.65). Because a given paper can be either correctly or incorrectly assigned to a category, we can use the binomial distribution to make a binomial test. For $N \geq 50$ and $Np(1 - p) \geq 9$, the binomial distribution can be approximated to the normal distribution, with the following z-score:

$$z = \frac{N(p - m)}{\sqrt{Np(1 - p)}}$$

For a given degree of significance σ, the associated z-score allows us to calculate the minimum size of the population N to guarantee that the correct category A will have at least a proportion m in the sample.

$$N \geq \left(\frac{z_{critical}}{p - m}\right)^2 p(1 - p)$$

Assuming a degree of misclassification of 50%, enforcing a lower acceptance threshold of 40% and a significance level of $\sigma = 0.05$, one obtains that the sample should be larger than approximately 70 papers (140 papers would increase the significance level to $\sigma = 0.01$). Table 1 shows the number of papers needed under different assumptions.

These results teach us the number of papers N needed in the populated categories to have some certainty. This means that the actual number of papers per overlay map depends on how narrow or wide the distribution of disciplinary categories is. The more skewed the distribution, the fewer papers are needed. Taking the example of Figure 3, one can estimate that in diverse universities such as Amsterdam or Sussex, 3,000 publications may be needed to capture precisely the five top disciplines, whereas for focused organizations such as EMBL, 1,500 publications could be enough.

In our opinion, these are rather conservative estimates, having set at $p = 0.5$ of misassignation. If one allows for "near misses" (i.e., assignation to the two nearest categories to be counted as correct), then p can be estimated in the range of 0.70 to 0.85 (Rafols & Leydesdorff, 2009). In this case, only some dozens of papers are needed to achieve $m \sim 0.5$.

TABLE 1. Approximate number of papers recommended for the reliable identification of a category in an overlay map.

p (Probability disagreement)	m (Lower tolerance)	σ (Significance level)	Minimum number of papers needed
0.5	0.4	0.10	41
0.5	0.4	0.05	68
0.5	0.4	0.01	135
0.6	0.5	0.05	65
0.4	0.3	0.05	65

DOI: 10.1002/asi