Jointly published by Elsevier Science Ltd, Oxford and Akadémiai Kiadó, Budapest

Scientometrics, Vol. 46, No. 3 (1999) 591–604

INTEGRATING RESEARCH PERFORMANCE ANALYSIS AND SCIENCE MAPPING

E. C. M. NOYONS, H. F. MOED, A. F. J. VAN RAAN

Centre for Science & Technology Studies (CWTS),Leiden University, P.O. Box 9555, 2300 RB Leiden (The Netherlands)

(Received August 31, 1999)

In this paper we present the explorations of combining the two main pillars of evaluative bibliometrics. These two pillars, performance analysis and science mapping, both have their strengths and imperfections. In this study we show how these imperfections are dealt with by an integrated analysis.

Introduction

Presently, a "standard" research performance analysis based on publication and citation data extracted from the Citation Indexes of the Institute for Sciecntific Information (ISI), is the most important application of evaluative bibliometrics. The comprehensibility of indicators based on publication and citation data is most attractive and objective, and therefore very popular as a research evaluation tool. The science mapping studies are used on a smaller scale. Complex structures of science and technology often discourage potential users. A bibliometric map of a research field is often supposed to lack direct reference to the known paradigms (*Healey* et al., 1986 and *Noyons*, 1999). In other words, the reference to the "real world" is not always clear. As a consequence, the utility becomes disputable. Moreover, both mapping and performance analysis has other drawbacks. The present paper aims at improving both the performance analysis and the mapping analysis by combining them into one integrated analysis.

The proposed method is illustrated with results from a case study of the field of neuroscience performed in a Targeted Socio-Economic Research (TSER) project for the European Commission (project code 1053, available via http://sahara.fsw.leidenuniv.nl/ed/projects.html).

0138–9130/99/US $ 15.00

Performance analysis

Bibliometric performance analysis aims at evaluating (groups of) scientific actors on the basis of bibliographic data. The most appropriate database available for this purpose is the set of ISI Citation Indexes.* Generally, "standard" performance analyses aim at assessing the activity of scientific actors (countries, universities, faculties, and departments) and the impact of their activity. The activity is measured by the number of publications in a particular period of time (in scientific journals covered by the ISI databases). The impact is measured by the number of times their publications are cited (by others). Furthermore, collaboration activity is often taken into consideration by measuring the number of times an actor publishes with others.

With respect to the impact, a normalization is used in order to compare peers with peers. Each scientific field has its own citing practices (*Schubert* et al., 1989; *Moed* et al., 1995, *Vinkler*, 1988). The average anthropologist is much less cited than the average immunologist. The field impact factor was developed to compare the impact of a publication to a "world average" of publications in the same field. For this purpose the journal impact factor and the ISI journal categories are used. In short, the impact of a publication in journal X with the journal category Y is compared to the average impact of publications in the journals in category Y.

Science mapping

Science mapping is mostly directed at monitoring a scientific field to determine its (cognitive) structure, its evolution, and main actors within. At CWTS a "standard" science mapping study identifies subdomains on the basis of co-word analysis. These subdomains are clusters of cognitively related core keywords in the studied field. The identified clusters of keywords are identified as subdomains, representing sets of publications on the covered topics (keywords). The trends in the field as a whole are studied by monitoring the interaction of subdomains and developments within each individual subdomain. To establish this monitor over time, dynamic maps are created. In a two-dimensional space the growth and interaction of the subdomains is depicted. Furthermore, by listing the most active actors by subdomain, a field actor analysis is carried out on subdomain level.

* The ISI Citation Indexes available on CD-ROM at present are: *Science Citation Index* (SCI), *Social Science Citation Index* (SSCI), *Arts and Humanities Citation Index* (AHCI), and the specialty Indexes (Compumath, Neuroscience, Biochemistry & Biophysics, Biotechnology, Chemistry, and Materials Science).

Imperfections

Both mapping studies and performance analyses have their imperfections. Apart from the fact that each research field has its own publication and citation practices, we isolated three "general" issues.

The journal-based impact factor. As described above, a normalization of impact takes place on the basis of the ISI journal categories. These categories are, however, not always specific enough to distinguish in smaller detail between specialties within a field, in order to account for the large impact differences. Furthermore, research fields tend to cross boundaries determined by classical fields (reflected by ISI journal categories) more and more. As a result, a normalization on the basis of a journal-based impact factor has its drawbacks. To illustrate this issue, some basic citation figures are given in Fig. 1 of four case journals often used by neuroscientists.

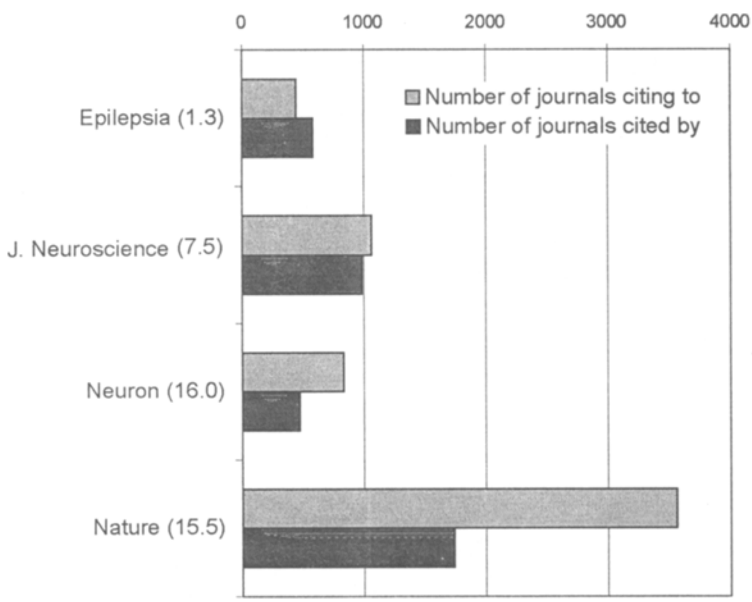

Fig. 1. Citation characteristics of three specialist journals and one multi-disciplinary journal (Numbers of journals cited by and citing to four case journals in 1996. Between parentheses is the impact factor in 1996. The world field average in 1996 was 3.0.)

We defined a specialist journal as a journal with only one category (neurosciences). We collected some general citation characteristics of three specialist journals that are used very frequently by neuroscientists. Moreover we included the data for *Nature* as an example of a typical multi-disciplinary, i.e., non-specialist journal. The chart depicts clearly that the scope differences between the specialist journals and *Nature*. The number of journals citing to and cited by *Nature* is twice to eight times the number for the specialist journals. In particular the number of journals citing *Nature* is much higher than the number of journals citing the specialist journals. On the other hand, the number of journals cited by *Nature* is not so much higher than the number of journals cited by the *Journal of Neuroscience.* The number of journals cited by the latter is, however, much higher than the number cited by the other specialist journals. Furthermore, the chart shows that the impact of the three specialist journals ranges somewhere between 1 and 16. *Neuron* has an impact around the level of *Nature*, while *Epilepsia* has a much lower impact, even below the world field average (around 3). The *Journal of Neuroscience* has an impact more than twice the world field average.

The fact that the number of journals cited by the *Journal of Neuroscience* is not so far below the number of journals cited by *Nature*, indicates that the former journal has a relatively wide scope. The numbers of citing and cited journals for the other two specialist journals indicate that their scope is much smaller. In other words, the field of neuroscience incorporates journals with different scopes. We may wonder whether the category neuroscience is appropriate to bibliometrically evaluate the field properly. The large differences between impact factors of specialist journals indicate that we should suspect the world field average as a trustworthy normalization.

Thus, the first imperfection we discern, concerns the normalization of the impact of individual publications on the basis of a world field average, used in the "standard" performance analysis. The journal scope differences indicate that the field neuroscience as defined by the journal category is too broad. Moreover, the impact of the individual neuroscience journals seems to support this. A field average based on the ISI journal categories, therefore, does not seem to be appropriate (see also *Glänzel*, 1998 and 1999; *Takahashi*, 1999).

Map validation. The second imperfection is identified in a "standard" mapping study. It concerns its reference to the "real world". An (expert) validation is of vital importance to assure this reference (*Noyons*, 1999). The recognition of identified subdomains plays an important role for mapping as a policy-supportive tool because there should be some reference to the "actual" situation. For this kind of validation, we need field experts. He should indicate whether the identified subdomains refer to actual clusters of themes within a field (see *Noyons*, 1999). To help him "understand" the structure, he has to be able to explore the structure. He has to be able to validate positions in relation to underlying data. A map of 10 subdomains would ideally need 9 dimensions for a perfect representation. The fact that these 9 dimensions have been reduced to only 2, may cause some subdomains to be in a "unexpected" position or environment.

Activity vs. impact. Thirdly, a science mapping analysis of a research field gives an overview of the structure and its evolution during a certain period of time. On top of that the contribution of actors to the field and its evolution is established by an actor analysis. An issue that has been raised more than once regarding this approach concerns the value of the actors' activity. For instance, does a strongly increasing activity of an actor in a particular subdomain indicate that this actor has a leading role in this subdomain? Maybe we are rather dealing with a "follower of fashion".

Method

In this study, the field of neuroscience, as represented by all publications covered by ISI's *Neuroscience Citation Index* (NCI), is structured on the basis of co-word analysis. We use this method to detach from a structure based on journals and their categories, because we claim they are not appropriate for that purpose, at least not in neuroscience. First, 37 subdomains were identified by clusters of cognitively related keywords, extracted from titles and abstracts. These subdomains are positioned in a two dimensional space on the basis of the cognitive relations of each individual subdomain with all the others (for details about the method, see *Noyons* and *Van Raan*, 1998). Additionally, the subdomains are analyzed with a detailed actor analysis (c.f., *Noyons* and *Van Raan*, 1996) as well as with a standard CWTS performance analysis (c.f., *Moed* et al., 1995). The latter provides per identified subdomain data on the average impact, most highly cited actors, and citing relations between individual subdomains.

The three addressed imperfections will be discussed on the basis of the integrated results.* In each case, we will suggest how the proposed method contributes to the improvements.

---

* Although the proposed method has a different objective, it bears elements of the integrated analysis proposed by *Braam*, *Moed* and *Van Raan* (1991a). They start with a co-citation analysis and use word profiles to enhance the structure.

Results

As mentioned before, the strength of an integrated bibliometric analysis is illustrated on the basis of the three identified imperfections of both composing parts.

First, the generated structure of neuroscience 1997-1998 is presented. The map of neuroscience (Fig. 2) shows a structure in which on the left-hand side the more fundamental neurochemistry and neurophysical research is represented and on the right-hand side the more applied (clinical neurology and surgery) side of neuroscience is represented. On the latter side, we find all kinds of neuro-related diseases.

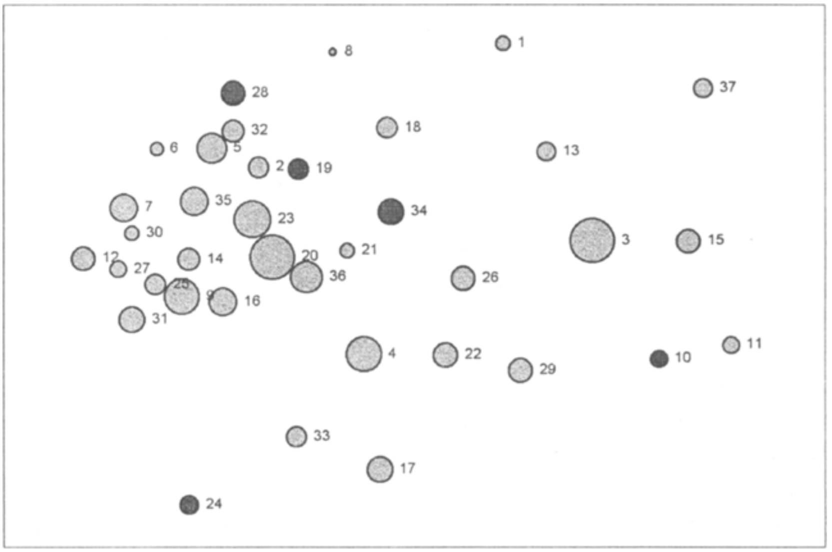

Fig. 2. General overview map of neuroscience 1997-1998
Two dimensional representation of neuroscience based on the similarities between identified clusters of keywords (subdomains). The circle size indicates the number of publications represented. The color of the circles indicates a significant increase/decrease of activity: dark grey: increase, white: decrease. The badness-of-fit criterion is 0.19, the distance correlation is 0.91 (statistics provided by SAS).

Subdomains

1	multiple sclerosis / myelin basic protein / experimental autoimmune encephalomyelitis
2	Astrocytes / Glial cell / TNF Alpha / acidic protein
3	Etiology / differential diagnosis / neurological deficit / spinal cord injury
4	Schizophrenia / Ethanol / Alcohol / normal control
5	Retina / skeletal muscle / neuronal cell / molecular mechanism
6	NGF / nerve growth / neurotrophic factor / pc12 cell
7	Ca2+ / inhibitory effect/ protein kinase
8	amyotrophic lateral sclerosis / motor neuron disease
9	h 3 / Dopamine / Antagonist / Agonist
10	Stroke / ischemic stroke / stroke patient / cerebral infarction
11	subarachnoid hemorrhage / middle cerebral artery / internal carotid artery
12	Peptide / Hormone / Secretion / Male Rat
13	CSF / HIV / AIDS
14	Glutamate / NMDA / glutamate receptor
15	MRI/ computed tomography / Functional MRI
16	Acetylcholine / Neurotransmitter / Uptake / Norepinephrine
17	Depression / Placebo / Anxiety
18	Alzheimers Disease / a beta / amyloid precursor protein / beta amyloid
19	Apoptosis / cell death / neuronal death / neurodegenerative disease
20	Animal model / electrical stimulation / Fiber / Pathophysiology
21	Ischemia / cerebral ischemia / neuronal damage / neuroprotective effect
22	Dementia / Aging / cognitive function / cognitive impairment
23	Axon / Immunoreactivity / Immunohistochemistry
24	heart rate / blood pressure / sympathetic nervous system / heart rate variability
25	Gaba / synaptic transmission / gamma aminobutyric acid / synaptic plasticity
26	PET / cerebral blood flow / white matter
27	Hypothalamus / c fos / paraventricular nucleus / locus coeruleus
28	Gene / CDNA / polymerase chain reaction / expression pattern
29	Seizure / EEG / Epilepsy / temporal lobe
30	nitric oxide synthase / l arginine / neuronal nitric oxide synthase
31	Stress / substance p / neuropeptide y / tyrosine hydroxylase
32	spinal cord / Peripheral nerve / sensory neuron / dorsal root ganglion
33	Memory / Learning / working memory / memory impairment
34	Pathogenesis / Parkinsons Disease / basal ganglion / oxidative stress
35	MRNA / rat brain / gene expression / olfactory bulb
36	Hippocampus / Cortex / Cerebellum / Striatum
37	Tumor / Brain Tumor / radiation therapy / primitive neuroectodermal tumor

From a journal based field impact factor to a topic based impact factor

We assume that a science map represents the cognitive structure of a research field properly. Moreover, we assume that the average impact figures per neuroscience subfield (i.e., "real world" themes within neuroscience) differ considerately. If so, it is to be expected that the distribution of impact average per subdomain correlates with the cognitive structure. In other words, it is to be expected that there are high and low impact subdomains, and that high impact subdomains are in each other's vicinity as well as relatively low impact subdomains. Figure 3 depicts this distribution in neuroscience.

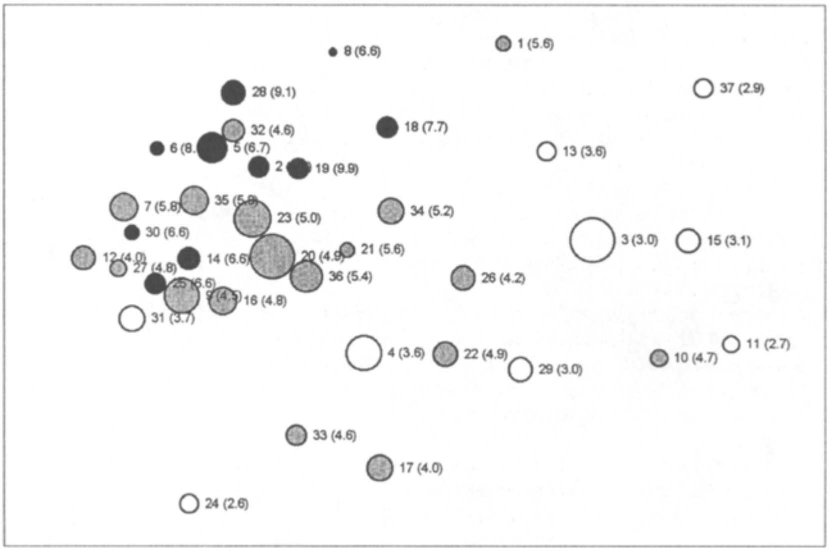

Fig. 3. Distribution of subdomain impact over neuroscience map (1997-1998)
Average impact per subdomain (between parentheses) was calculated by the short term impact of publications in 1995 and 1996 and citations received in 1995-1998. Dark grey circles reprsent subdomains with an impact which is 20% above field (neuroscience) average. White circles represent subdomains with an impact 20% below average. Further map legend: see Figure 2.

On the left-hand side the average per subdomain is considerately higher than the average impact on the right-hand side. Or even more accurately, the average impact seems to increase from the below right to upper left part of the map. This indicates that the cognitive structure (identification of subdomains) on the basis of co-word analysis, is a good alternative to disseminate the overall impact factor in order to compare peers with peers. A similar finding was reported in *Noyons* et al. (1999). Such a structuring of a research field enables us to determine a field impact average that is not based on journal averages.

Note, however, that we started the previous argument with the assumption that the cognitive map of neuroscience refers the "real world". This exact point was the second imperfection of evaluative bibliometrics, validation of science maps.

Validation of science maps

As mentioned before, we need field experts to determine the "real" validity of the science map. In order to help him validate the structure, we created interactive tools to investigate the structure in more detail and from different perspectives. An example is given in Fig. 4.

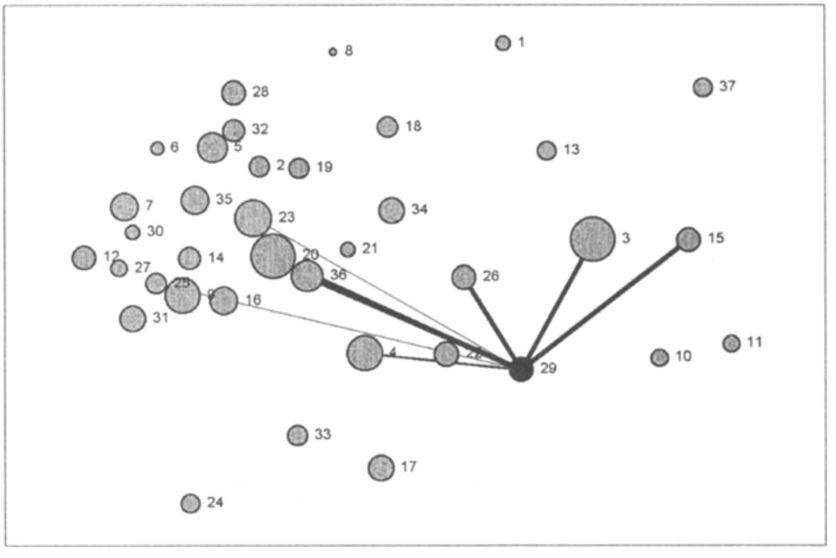

Fig. 4. Strongest pair-wise links from the perspective of subdomain 29 (Seizure/Epilepsy/EEG)
Line weight indicates the percentile (p80 is indicated by lines) of strongest publication co-occurrence relations with 29.

From the perspective of subdomain 29 (Seizure/Epilepsy/EEG), we indicated the strongest pair-wise relations with other subdomains by connecting lines. This information enhances the structure because it emphasizes the direct relations of an individual subdomain without taking into account the relations that other subdomains have with each other. In that respect, this information justifies the position that subdomain 29 has in the map regarding the direct relations it has with others. This information is important to evaluate as to how well the two available dimensions "cover" the 36 dimensions needed for a perfect representation.

These strongest directly related subdomains, as shown in the chart, accounts for an important part for the position of 29. Still, it does not account for the exact position of all surrounding subdomains. For instance, subdomain 10 (Stroke research) is in its environment but does not appear among the strongest directly related subdomains of 29. Moreover, 29 does not appear to be among the strongest directly related subdomains of 10.* The reason why they are in each others vicinity is that they share the strong relations with other subdomains (in particular 3, 15 and 26).

In a similar way, we provide this information for the citation relations of each subdomain. We enhance the structure with information that is primarily outside the map, in the sense that it does not contribute directly to the structure of the map. In Fig. 5, we indicated the most highly cited subdomains by 29.

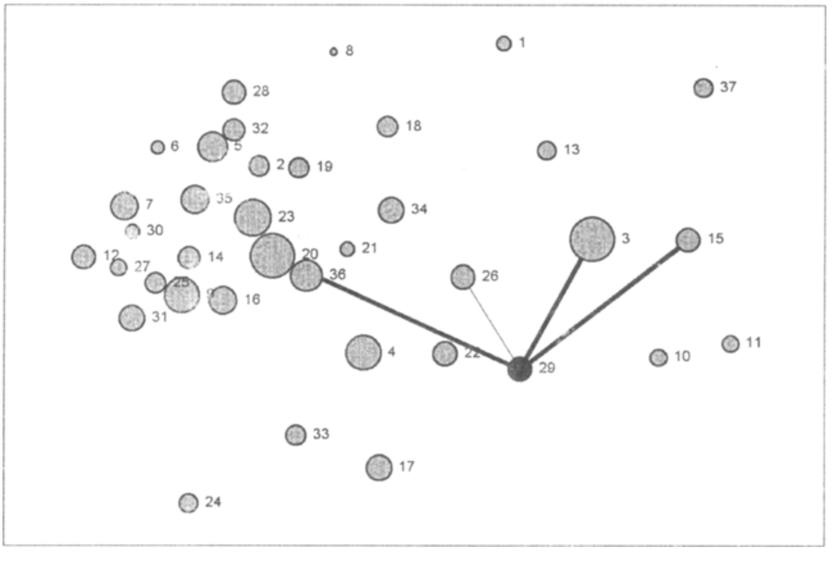

Fig. 5. Strongest citation links from the perspective of subdomain 29 (Seizure/Epilepsy/EEG) P80 of cited subdomains by 29 are indicated by lines.

Although the linkages are not exactly the same, we find a similar set of subdomains most strongly related to 29 (3, 15, 20 and 26). Again, 10 and 29 are not among each other's strongest relations, but they share related other subdomains (3 and 20).

* This information is available at the Internet site of this project (to be accessed via the WWW-projects on the CWTS page http://sahara.fsw.leidenuniv.nl).

Does activity correlate with impact?

A standard science mapping study returns a field structure and its evolution. Moreover, an actor analysis of the identified subdomains depicts the organisations most probably responsible for the evolved structure. In *Noyons* and *Van Raan* (1998), it was explained how these two parts can be integrated into one analysis. It was also mentioned that such an analysis reveals activity only. The found trends do not account for the impact of underlying publications. For instance, the growth of a certain subdomain during a particular period of time, reveals that this subdomain is "hot", or at least "heating up". The actor analysis will point out that certain organizations increased their activity during the same period. The question is: are these organizations at the research front of this subdomain or are they "dedicated followers of fashion"? With an integrated analysis of mapping and performance analysis, we claim to be able to answer this question. A first approximation of the possibilties of such an approach is demonstrated in the following example.

The next tables show the most active (Table 1) and most highly cited organizations (Table 2) for 29 (Seizure/Epilepsy/EEG).

Most of the most active organizations are also among the most highly cited organizations. For instance, Univ. Calif. LA, Harvard Univ. Boston, and Yale Univ. New Haven are in all lists in the top five. Univ. Bonn in Germany, however has increased its production with about 50% but is not among the highly cited organizations. On the other hand, Univ. London is in the top-five of highly cited organizations with a significantly increasing number of citations but are not found in the list of most active in this subdomain.

These data indicate in a first approximation that in the subdomain Seizure/Epilepsy/EEG, Univ. Bonn has improved its activity to be one of the most avtive organizations. However, their increased effort has not lead to a top-ranking in terms of citations received. With the Univ. London, for instance, the situation is exactly the other way around. They are among the top organizations in terms of received citations but they are not visible in terms of most active organizations. The former may on the basis of these data be called a "follower of fashion", whereas the latter may be called the intellectuale base of this subdomain.

Table 1
Most active organizations in Seizure/Epilepsy/EEG

95/96	97/98	Institute
85	91	UNIV CALIF LOS ANGELES, LOS ANGELES, USA
57	84	HARVARD UNIV, BOSTON, USA
42	68	MCGILL UNIV, MONTREAL, CANADA
42	65	UNIV BONN, BONN, GERMANY
48	60	YALE UNIV, NEW HAVEN, USA
26	54	UNIV CALIF SAN FRANCISCO, SAN FRANCISCO, USA
36	48	UNIV PITTSBURGH, PITTSBURGH, USA
22	48	UNIV TORONTO, TORONTO, CANADA
26	45	UNIV WASHINGTON, SEATTLE, USA
30	39	DUKE UNIV, DURHAM, USA
22	39	NYU, NEW YORK, USA
22	39	UNIV ALABAMA, BIRMINGHAM, USA
33	38	UNIV PENN, PHILADELPHIA, USA
36	37	MAYO CLIN & MAYO FDN, ROCHESTER, USA
48	36	NATL HOSP NEUROL & NEUROSURG, LONDON, ENGLAND
42	36	INST NEUROL, LONDON, ENGLAND
37	36	CLEVELAND CLIN FDN, CLEVELAND, USA
22	35	JOHNS HOPKINS UNIV, BALTIMORE, USA
20	35	UNIV ZURICH, ZURICH, SWITZERLAND
32	34	STANFORD UNIV, STANFORD, USA
30	32	RUSSIAN ACAD SCI, MOSCOW, RUSSIA
24	31	UNIV WISCONSIN, MADISON, USA
21	31	UNIV VIRGINIA, CHARLOTTESVILLE, USA
31	30	COLUMBIA UNIV, NEW YORK, USA
29	30	UNIV VIENNA, VIENNA, AUSTRIA
22	30	UNIV MICHIGAN, ANN ARBOR, USA
19	30	UNIV TEXAS, HOUSTON, USA
11	30	UNIV HELSINKI, HELSINKI, FINLAND
8	30	CHILDRENS HOSP, BOSTON, USA

Table 2
Most highly cited organizations in Seizure/Epilepsy/EEG

95/96	97/98	Organisation
53	69	YALE UNIV, NEW HAVEN, USA
43	54	UNIV CALIF LOS ANGELES, LOS ANGELES, USA
46	47	HARVARD UNIV, BOSTON, USA
30	47	UNIV LONDON, LONDON, ENGLAND
17	44	UNIV PENN, PHILADELPHIA, USA
30	42	NATL HOSP NEUROL & NEUROSURG, LONDON, ENGLAND
46	39	MAYO CLIN & MAYO FDN, ROCHESTER, USA
27	38	UNIV MINNESOTA, MINNEAPOLIS, USA
42	37	UNIV CALIF SAN FRANCISCO, SAN FRANCISCO, USA
28	35	UNIV ALABAMA, BIRMINGHAM, USA
41	34	MCGILL UNIV, MONTREAL, CANADA
34	34	DUKE UNIV, DURHAM, USA
20	31	COLUMBIA UNIV, NEW YORK, USA
35	30	MONTREAL NEUROL HOSP & INST, MONTREAL, CANADA
12	30	UNIV PITTSBURGH, PITTSBURGH, USA
19	29	UNIV WASHINGTON, SEATTLE, USA
23	28	NYU, NEW YORK, USA
36	27	INST NEUROL, LONDON, ENGLAND
27	27	UNIV MICHIGAN, ANN ARBOR, USA
20	26	JOHNS HOPKINS UNIV, BALTIMORE, USA
17	24	UNIV TEXAS, HOUSTON, USA
10	24	CLEVELAND CLIN FDN, CLEVELAND, USA
19	23	INST CHILD HLTH, LONDON, ENGLAND
17	23	MASSACHUSETTS GEN HOSP, BOSTON, USA
13	23	UNIV TEXAS, DALLAS, USA
11	23	UNIV MIAMI, MIAMI, USA
9	23	UNIV MELBOURNE, MELBOURNE, AUSTRALIA
12	22	STANFORD UNIV, STANFORD, USA
17	21	UNIV TORONTO, TORONTO, CANADA
19	20	AUSTIN HOSP, HEIDELBERG, AUSTRALIA
19	20	HOP ST VINCENT DE PAUL, PARIS, FRANCE
12	20	ERASMUS UNIV ROTTERDAM, ROTTERDAM, NETHERLANDS

Conclusion

This study presents an integration of "standard" performance analysis and science mapping for evaluative purposes. Three described imperfections of the composing parts are addressed by the integrated approach. This approach proves to be useful to address the imperfections. Furthermore, the integrated approach reveals new opportunities for evaluative bibliometrics. At this point, the full range of opportunities still has to be explored. One of the possibilities could be to enhance the basic publication data for science mapping with impact data in order to give high impact publications more weight to determine the field structure. Thus, a traditional criticism on science mapping may be tackled.

Finally, we would like to stress that the interactive character of the maps, which are published in HTML format, provides opportunities for the "standard" performance as well. The "clickability" of results enables a user to investigate aggregated indicators in more detail by disseminating these indicators and to make them accessible through an interface.

References

BRAAM, R.R., H.F. MOED, A.F.J. VAN RAAN (1991), Mapping of science by combined co-citation and co-word analysis, I: Structural aspects. *Journal of the American Society for Information Science,* 42, 233-251.

GLÄNZEL, W., A. SCHUBERT, U. SCHOEPFLIN, H.J. CZERWON (1999), An item-by-item subject classification of papers published in journals covered by the SSCI database using reference analysis. *Scientometrics,* 46(3) 431–441.

GLÄNZEL, W., A. SCHUBERT, H.J. CZERWON (1998), An item-by-item subject classification of papers published in multidisciplinary and genaral journals using reference analysis. *Scientometrics,* 44, 427-439.

HEALEY, P., H. ROTHMAN, P.K. HOCH (1986), An experiment in science mapping for research planning. *Research Policy,* 15, 233-251.

MOED, H.F., R.E. DE BRUIN, TH.N. VAN LEEUWEN (1995), New bibliometric tools for the assessment of national research performance: Database description, overview of indicators and first applications. *Scientometrics,* 33, 381-442.

NOYONS, E.C.M. (1999), *Bibliometric Mapping as a Science Policy and Research Management Tool.* DSWO Press, Leiden University.

NOYONS, E.C.M., A.F.J. VAN RAAN (1996), Actor analysis in neural network research: The position of Germany. *Research Evaluation,* 6, 133-142.

NOYONS, E.C.M., A.F.J. VAN RAAN (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, 68-81.

NOYONS, E.C.M., M. LUWEL, H.F. MOED (1999), Combining mapping and citation analysis for evaluative bibliometric purposes. *Journal of the American Society for Information Science,* 50, 115-131.

SCHUBERT, A., W. GLÄNZEL, T. BRAUN (1989), Scientometric Datafiles – A comprehensive set of indicators on 2649 journals and 96 countries in all major science fields and subfields 1981-1985. *Scientometrics,* 16, 3-478.

TAKAHASHI, K., T.C. AW, D. KOH (1999), An alternative to journal-based impact factors. *Occupational Medicine (Oxford),* 49, 57-58.

VINKLER, P. (1988), Bibliometric features of some scientific subfields and the scientometric consequences therefrom. *Scientometrics,* 14, 453-474.