The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments

Anthony F. J. van Raan
Centre for Science and Technology Studies
University of Leiden
Wassenaarseweg 52
P.O. Box 9555
2300 RB Leiden, The Netherlands
vanraan@cwts.leidenuniv.nl

Abstract

This paper[1] presents an overview of advanced bibliometric methods for (1) objective and transparent assessment of strengths and weaknesses in research performance, and (2) monitoring interdisciplinary scientific developments. In the first application, we focus on the detailed analysis of research performance in an international comparative perspective. We demonstrate that advanced bibliometric methods are, particularly at the level of research groups, university departments and institutes, an indispensable element next to peer review in research evaluation procedures. We address specific problems for the social sciences.
In the second application, monitoring of scientific (basic and applied) developments, recent advances in bibliometric mapping techniques are promising. They are unique instruments to discover patterns in the structure of scientific fields, to identify processes of knowledge dissemination, and to visualize the dynamics of scientific developments. We discuss briefly their potential for unraveling interdisciplinary developments and interfaces between science and technology.

1. Introduction: Why Bibliometric Analysis?

Science is a driving force of our modern society. Particularly excellent scientific work is the cradle of breakthroughs in our knowledge of the world. Therefore, evaluation of scientific research is crucial. Review by colleague-scientists, 'peers', is applied to judge research proposals, appointments of research staff and evaluation of research groups or programs. Peer review is typically a *qualitative* assessment of research performance. Bibliometric indicators discussed here represent the *quantitative* side. But quantitative elements are clearly also present in peer review, e.g., number of publications in high prestige scientific journals. Conversely, citations given to research work can be seen as judgements, 'votes' of colleague-scientists in favour of the work cited.

---

[1] This paper is a revised and extended version of earlier recent papers by the author (van Raan 1998; 2000a).

In this paper we discuss an advanced bibliometric method for research performance assessment. Bibliometric assessment of research performance is based on one central assumption: scientists who have to say something important do publish their findings vigorously in the open, international journal ('serial') literature.

Why bibliometric analysis of research performance? Peer review undoubtedly has to remain the principal procedure of quality judgement. But peer review and other related expert-based judgements have serious shortcomings and disadvantages (Horrobin 1990; Moxham and Anderson 1992). Opinions of experts may be influenced by subjective elements, narrow-mindedness and limited cognitive horizons. Subjectivity, i.e., dependence of the outcomes on the choice of individual committee members, is a major problem. This dependence may result in conflicts of interests, unawareness of quality, or a negative bias against younger people or newcomers to the field.

We absolutely do not plead for a replacement of peer review by bibliometric analysis. Subjective aspects are not merely negative. In any judgement there must be room for the intuitive insights of experts. We claim however that for a substantial improvement of decision-making our bibliometric method has to be used in parallel to a peer-based evaluation procedure (Rinia *et al* 1998).

The most crucial parameter in the assessment of research performance is *international scientific influence*. We consider international influence as an important, measurable aspect of scientific quality and therefore we developed standardized, bibliometric procedures to assess research performance within the framework of international influence or impact. Undoubtedly, the bibliometric approach is not an ideal instrument, working perfectly in all fields under all circumstances. But our approach works very well in the large majority of the natural, the medical, the applied sciences, and in several fields within the social and behavioral sciences. One the most important features of our method is that it provides more than just 'nice additional data'. It forces the experts to re-think their judgements and it provides challenging new insights. Thus they form, particularly at the level of research programs, an indispensable tool for decision-making in science policy, particularly in priority setting.

Bibliometric analyses performed at the *macro*-level (e.g., a whole country) yield at best general assessments of fields as a whole, for instance, how good a country's performance is in physics, chemistry, psychology or immunology, *without* a reliable breakdown to the individual research groups or programs. Therefore, research performance should be analyzed systematically on the *meso*-level of larger institutions, such as universities or major parts of universities, like faculties or institutes. After an overall assessment of these larger institutions, performance analysis can be narrowed down to the most important level: the *micro*-level, that is, the real 'workfloor' of research practice: departments, research groups and programs within universities and large institutes.

On the meso- and micro-level, all necessary information, particularly data on personnel and on the composition of groups and programs, is only available within the university or institute concerned. Such institutional infrastructure data are *never* available in general publication databases and must *always* be collected separately in relation the institutions concerned.

2. Basic Principles of Bibliometric Indicators

The core of our bibliometric approach can be described as follows. Communication, i.e., exchange of research results, is the driving force in science. Publications are not the only, but certainly very important elements in this knowledge exchange process. Work of high quality provokes reactions of colleague-scientists. They are the international forum, the 'invisible college', by which research results are discussed. In most cases, these colleague-scientists play their role as a member of the invisible college by referring in their own work to earlier work of other scientists.

We all know that the process of citation is a complex one, and that it certainly not provides an 'ideal' monitor on scientific performance. This is particularly the case on a statistically low aggregation level, for instance, an individual researcher. But the application of citation-analysis to the work, the 'oeuvre', of *a group as a whole over a longer period of time*, does yield in many situations a strong indicator of scientific performance, and in particular of scientific quality. An important, absolutely necessary condition is that applied citation-analysis is part of an advanced, technically highly developed bibliometric method.

Research output is defined as the number of articles of the institute, as far as covered by the Science Citation Index (SCI), the Social Science Citation Index (SSCI), or the Arts & Humanities Citation Index (AHCI). As 'article' we consider the following publication-types: normal articles (including proceedings papers published in journals), letters, notes, and reviews (but not meeting abstracts, obituaries, corrections, editorials, etc.). We developed software to calculate a set of standardized, basic indicators.

To discuss this set of indicators, we take the results of our recent analysis of a German medical research institute as an example (time period 1992 – 2000). Table 1 shows in the first column the number of papers published, *P*, which is also a first but good indication of the size of an institute. This number is about 250 per year. In the second column we find the total number of citations, *C*, received by *P* in the indicated time period, *and corrected for self-citations*.

The analytic scheme is as follows. We take the last sub-period 1996-2000 as an example. For papers published in 1996, citations are counted during the period 1996-2000, for 1997 papers citations in 1997-2000, and so on. There is ample empirical evidence that in the natural and life sciences -basic as well as applied- the average 'peak' in the number of citations is in the third or fourth year after publication (Moed et al 1995). Therefore a ('moving' and partially overlapping) five-year analysis period is appropriate for impact assessment.

The third and fourth indicators are the average number of citations per publication *(CPP)*, again without self-citations, and the percentage of not-cited papers, *% Pnc*. We stress that this percentage of non-cited papers concerns, like all other indicators, the given time period. It is very well possible that publications not cited within such a block will be cited after a longer time. This is clearly visible when comparing this indicator for the five-year periods (e.g., 1996-2000: 30%) with that of the whole (that is, longer) period (1992-2000: 21%). The values found for this medical research institute are quite normal.

How do we know that a certain volume of citations, or a certain citation-per-publication value is low or high? Therefore it is crucial to make a comparison with (or normalization to) a well-chosen international reference value, and to establish a reliable measure of *relative, internationally field-normalized impact*. Furthermore, as overall, worldwide citation rates are increasing, it is also necessary to normalize the measured impact of an institute (**CPP**) to international reference values.

Table 1: Bibliometric analysis of medical research institute 1992 – 2000

Institute	*P*	*C*	*CPP*	*%Pnc*	*CPP/ JCSm*	*CPP/ FCSm*	*CPP/ D-FCSm*	*JCSm/ FCSm*	*%SCit*
1992 - 00	**2,245**	**43,665**	**19.45**	**21**	**1.26**	**1.95**	**1.85**	**1.55**	**18**
1992 - 96	1,080	11,151	10.33	36	1.27	2.02	1.95	1.58	22
1993 - 97	1,198	12,794	10.68	34	1.24	2.03	1.92	1.63	21
1994 - 98	1,261	12,217	9.69	32	1.19	1.85	1.72	1.55	22
1995 - 99	1,350	13,709	10.15	31	1.21	1.89	1.76	1.56	21
1996 - 00	1,410	14,815	10.51	30	1.20	1.91	1.76	1.59	21

First, we calculate the average citation rate of all papers (world-wide) in the journals in which the institute has published (**JCSm**, the **m**ean **J**ournal **C**itation **S**core of the institute's 'journal set'). Thus, this indicator **JCSm** defines a worldwide reference level for the citation rate of the institute. It is calculated in the same way as **CPP**, but now for all publications in a set of journals (see van Raan 1996). With help of the ratio **CPP/JCSm** (5th indicator) we observe whether the measured impact is *above* or *below* international average.

Comparison of the institute's citation rate *(CPP)* with the average citation rate of its journal set (**JCSm**) introduces a specific problem related to journal status. For instance, if the institute publishes in prestigious (high impact) journals, and another institute in rather mediocre journals, the citation rate of articles published by both groups may be equal *relative to* the average citation rate of their respective journal sets. But the first group evidently performs better than the second. Therefore, we developed a second international reference level, a *field*-based world average **FCSm**. This indicator is based on the citation rate of *all* papers (world-wide) published in *all* journals of the field(s)[2] in which the institute is active, and not only the journals in which the institute's researchers publish their papers. For a publication in a less prestigious journal one may have a (relatively) high **CPP/JCSm** but a lower **CPP/FCSm,** and for a publication in a more prestigious journal one may expect a higher **CPP/FCSm**, as publications in a prestigious journal will have generally have an impact above the field-specific average.

---

[2] We use the definition of fields based on a classification of scientific journals into *categories* developed by ISI. Although this classification is not perfect, it is at present the most suitable classification available to us in terms of an automated procedure within our data-system.

We use the same procedure as the one we applied in the calculation of *JCSm*. A novel and unique aspect of our comparison with both worldwide reference indicators is that we take into account the type of paper (e.g., letters, normal article, review) *as well as* the specific years in which the papers were published. This is absolutely necessary, as the average impact of journals may have considerable annual fluctuations and large differences per article type, see Moed and Van Leeuwen 1995, 1996).

Often an institute is active in more than one field. In such cases we calculate a weighted average value, the weights being determined by the total number of papers published by the institute in each field. For instance, if the institute publishes in journals belonging to genetics and heredity, as well as to cell biology, then the *FCSm* of this institute will be based on both field averages. Thus, indicator *FCSm* represents a *world average*[3] in a specific (combination of) field(s). It is also possible to calculate *FCSm* for a specific country or for the European Union. The example discussed in this paper concerns a German medical research institute and for this institute we calculated the Germany-specific *FCSm*-value, *D-FCSm*.

As in the case of *CPP/JCSm*, if the ratio *CPP/FCSm* (6th indicator) is above 1.0, the impact of the institute's papers exceeds the field-based (i.e., *all* journals in the field) world average. We observe in Table 1 that the *CPP/JCSm* is 1.20, *CPP/FCSm* 1.91 and *CPP/D-FCSm* (7th indicator) is 1.76 in the last period 1996 – 2000. These results show that the institute is performing well above international average. The ratio *JCSm/FCSm* (8th indicator) is also an interesting indicator. Is it above 1.0, the mean citation score of the institute's journal set exceeds the mean citation score of all papers published in the field(s) to which the journals belong. For the institute this ratio is around 1.59. This means that the institute publishes in journals with, generally, a high impact. The last (9th) indicator shows the percentages of self-citations (*%Scit*). About thirty percent is normal, so the self-citation rates for this institute are certainly not high (about 20%).

We regard the *internationally standardized impact indicator CPP/FCSm* as our 'crown' indicator. This indicator enables us to observe immediately whether the performance of a research group or institute is significantly far below (indicator value < 0.5), below (indicator value 0.5 - 0.8), around (0.8 - 1.2), above (1.2 – 1.5), or far above (>1.5) the international (western world dominated) impact standard of the field. We stress that in the measurement of scientific impact one has to take into account the *aggregation level of the entity* under study. The higher the aggregation level, the larger the volume in publications and the more difficult it is to have an impact significantly above the international level. Based on our long-standing experiences, we can say the following. At the 'meso-level' (e.g., a large institute), a *CPP/FCSm* value above 1.2 means that the institute's impact as a whole is significantly above (western-) world average.

Particularly with a *CPP/FCSm* value above 1.5, such as in our example, the institute can be considered as scientifically strong, with a high probability to find very good to excellent groups. Thus, the next step in a research performance analysis is a breakdown of the institution into smaller units, i.e., research groups and/or programs. Therefore the bibliometric analysis has to be applied on the basis of institutional input data on personnel and composition of groups.

---

[3] About 80 percent of all SCI-covered papers is authored by scientists from the United States, Western Europe, Japan, Canada, and Australia. Therefore, our 'world average' is dominated by the Western world.

Then, the bibliometric algorithms can be repeated efficiently on the lowest but most important aggregation level, that of the research group or research program. In most cases the volume of publications at this level is between 10 and 20 per year. At the group level a *CPP/FCSm* value above 2 indicates a very strong group, and above 3 the groups can be, generally, considered as excellent and comparable to top-groups at the best US universities. If the threshold value for the *CPP/FCSm* indicator is set at 3.0, we filter out the excellent groups with high probability.
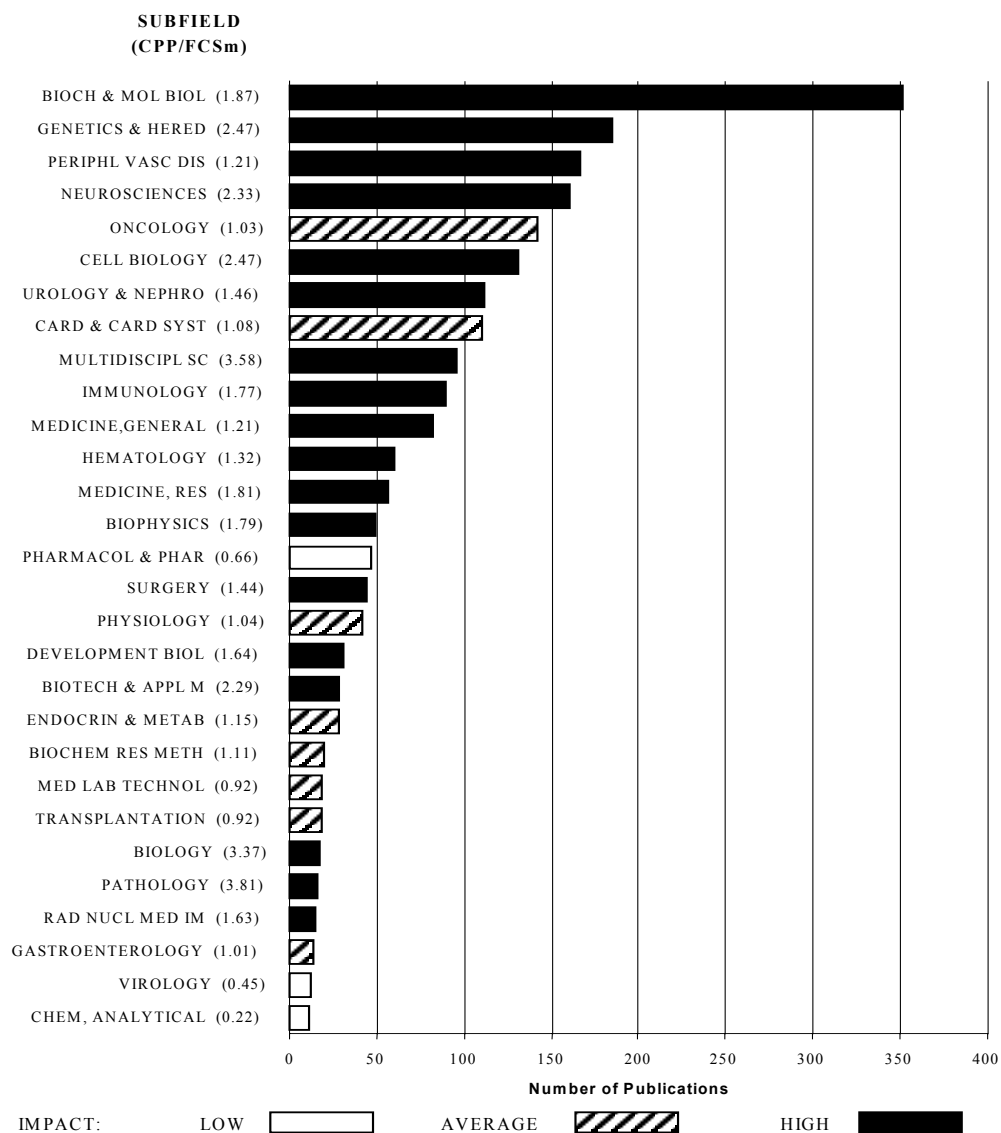
3. Bibliometric spectroscopy: measuring interdisciplinarity

A further important part of our bibliometric methodology is the *breakdown* of the institute's output *into* research fields. This provides a clear impression of the research scope or 'profile' of the institute. Such a *spectral analysis* of the output is based on the simple fact that the researchers publish in journals of many different fields. Our example, the German medical research institute, is a center for broad, medical science oriented, molecular research. The researchers of this institute are working in a typical interdisciplinary environment. The institute's publications are published in a wide range of fields: biochemistry and molecular biology, genetics and heredity, oncology, cell biology, and so on.

By ranking fields according to their size (in terms of numbers of publications) in a graphical display, we construct the research profile of the institute. Furthermore, we provide the *impact* of the institute's research in these different fields with help of *CPP/FCSm* as impact indicator normalized for *each the fields separately*. Figure 1 shows the results of this *bibliometric spectroscopy*. Thus it becomes immediately visible in which fields within its interdisciplinary research profile the institute has a high (or lower) performance (van Raan 2000b).

In Fig. 1 we observe the scientific strength of the target institute: its performance in the top-four fields is high to very high. If we find a smaller field with a relatively low impact (i.e., a field in the lower part, the 'tail' of the profile), this does not necessarily means that the (few) publications of the institute in this particular field are 'bad'. Often these small fields in a profile are those that are quite 'remote' from the institute's core fields. They are, so to say, peripheral fields. In such a case, the group's researchers may not belong to the dominating international research community of those fields, and as the consequence their work will be not be cited as frequently as the work of these dominating ('card holding') community members.

Figure 1: Research profile of medical research institute, 1992-2000

SUBFIELD
(CPP/FCSm)

Subfield	Number of Publications
BIOCH & MOL BIOL (1.87)	~353 (high)
GENETICS & HERED (2.47)	~185 (high)
PERIPHL VASC DIS (1.21)	~165 (high)
NEUROSCIENCES (2.33)	~160 (high)
ONCOLOGY (1.03)	~143 (average)
CELL BIOLOGY (2.47)	~130 (high)
UROLOGY & NEPHRO (1.46)	~110 (high)
CARD & CARD SYST (1.08)	~108 (average)
MULTIDISCIPL SC (3.58)	~95 (high)
IMMUNOLOGY (1.77)	~90 (high)
MEDICINE,GENERAL (1.21)	~82 (high)
HEMATOLOGY (1.32)	~60 (high)
MEDICINE, RES (1.81)	~57 (high)
BIOPHYSICS (1.79)	~50 (high)
PHARMACOL & PHAR (0.66)	~47 (low)
SURGERY (1.44)	~45 (high)
PHYSIOLOGY (1.04)	~43 (average)
DEVELOPMENT BIOL (1.64)	~33 (high)
BIOTECH & APPL M (2.29)	~30 (high)
ENDOCRIN & METAB (1.15)	~28 (average)
BIOCHEM RES METH (1.11)	~22 (average)
MED LAB TECHNOL (0.92)	~18 (average)
TRANSPLANTATION (0.92)	~18 (average)
BIOLOGY (3.37)	~18 (high)
PATHOLOGY (3.81)	~16 (high)
RAD NUCL MED IM (1.63)	~16 (high)
GASTROENTEROLOGY (1.01)	~12 (average)
VIROLOGY (0.45)	~12 (low)
CHEM, ANALYTICAL (0.22)	~10 (low)

Number of Publications

IMPACT: LOW ☐ AVERAGE ▨ HIGH ■

4. Implications for the social sciences

The increasing use of bibliometric indicators is a matter of achieving a more balanced and thus more objective assessment. Particularly in the social sciences, where more than in the natural and medical sciences, 'local' and 'national' orientations (Nederhof and Zwaan 1991; Kyvik and Larsen 1994) -and with that possibly 'provincial' attitudes- are present, and where also less consensus exists on what successful scientific approaches are, a reinforcement of a more international, 'cosmopolitan' and a more objective view on scientific performance is desirable.

We already noticed that bibliometric assessment of research performance is based on one important assumption: the work to be evaluated must be published in the open, international journal literature. This means that bibliometric indicators are highly applicable in the natural and life sciences. However, in the applied and engineering sciences as well as in the social and behavioural sciences (and even more in the humanities) international journals are often not the primary communication channel. Then, no doubt, bibliometric assessment becomes problematic. Nevertheless, we caution against an all too easily acceptance of the persistent characterization of the social sciences (and the humanities) as being 'bibliometrically inaccessible'.

The idea that the above features such as the less important role of journals, the 'local' orientation of many research fields, and also the dominant role of older literature, are general characteristics of all social sciences and humanities, is refuted by recent empirical work. For instance, nowadays linguistics and experimental psychology are more and more approaching the publication behaviour of the 'hard' sciences: the dominant role of international 'core' journals, and the strongly increasing citation of recent work (Nederhof and Zwaan 1991).

Bibliometric analysis has proven to be essential in the evaluation of social science research performance, as can be seen from earlier studies, for instance, concerning psychology (Nederhof and Zwaan 1991; Nederhof and Noyons 1992). Furthermore, recent experience in The Netherlands shows that bibliometric analysis can be applied successfully in the social sciences (Nederhof *et al* 2000). This has seriously questioned the findings of a peer review committee. It has also become clear that peer-review evaluation of fields where no bibliometric analysis has been applied, would have been of better quality if it had been (VSNU 1994, 1995; Kroonenberg and Van der Veer 1996). However, we maintain that bibliometric analysis is a support tool for peer review. Only in this situation will other measures of quality and esteem also be available, as part of common peer review.

Alongside technical problems, many *methodological* problems with respect to design, construction and calculation of appropriate indicators must be solved by advanced automated algorithms, enabling the choice of different indicator options. The major methodological problems are mostly general for all fields of science, but, for social sciences, several are particularly important. First of all, there are (very!) different publication and citation characteristics in the different fields of science. This is particularly the case for the social sciences. For instance, the difference in publication behaviour of the strongly internationally oriented experimental psychologists is in contrast to the much more 'locally' oriented sociologists. These differences must be known and taken into account: research fields should never be compared on the basis of absolute numbers of citations. Field-dependent normalization is absolutely necessary.

Field-dependent characteristics may change over time during the period of analysis. Even after field-dependent normalization of citation numbers, it is not clear whether a specific normalized score is high or low for that specific field. Thus, comparison with other, similar groups or with an international (world-wide or European) reference value for that specific research field is also necessary to get meaningful results. A 'European Union' comparison standard is an effective means of coping with possible Anglo-Saxon biases in the SSCI, as shown by recent work in the assessment of social psychology. Such a European reference standard can be based on a selected group of European journals, covered by the SSCI (Nederhof 1997). In other words, 'bare numbers of citations' has to be translated into a field-normalized, reference standard related impact.

The 'size of the object to be evaluated', that is, the aggregation level, must be sufficiently high. Application of bibliometric indicators at a level too low, for instance, individual scientists, will be statistically problematic, especially in the social sciences where the number of citations is often, roughly speaking, an order of magnitude lower than in the natural and medical sciences (Van Raan 1993). For research groups the situation is much better. A major methodological problem, again particularly in the social sciences, concerns the time dimension. Citations are given after publication. So, how long must we wait, in other words: what is an acceptable length for the 'citation window'? For the social sciences this window should be longer than in the natural sciences, and around five to six years. This unavoidable time lag (impact is mainly received *after* the work has been published), is often 'misused' by critics (even in the natural sciences where it is about two to three years) as a general objection against bibliometric analysis. Yet even peers generally need time to see whether research results will 'take root'!

Furthermore, trend analysis reveals striking features, such as the influence of break-through work, the effects of departure or appointment of key personnel. For instance, Nederhof and Van Raan (1993) found a strong influence of key-scientists ('star effect') in their bibliometric assessment of six British economics top-groups.

There are several important further indicators. We mention the relation between publication output and impact with type of collaboration (for instance, international) and the breakdown of output and impact according to the spectrum of research fields covered by the publications of the group or institute. There are also important media *not covered* by the SSCI. For instance, Meertens et al (1992) found in social psychology an important role of journals not SSCI-covered. They established that books and book-chapters constitute about one third of all Dutch social psychology publications. These 'non-SSCI media', however, can be cited quite considerably in SSCI-covered articles. Thus, with appropriate analytical routines, their impact can be assessed.

An important general observation in the application of bibliometric methods is that performance measurement, particularly in the social sciences, must cover a wider range of years. Bibliometric 'snapshots' are useless, even periods of five years are too short. So an important lesson is learned from bibliometric analysis: research groups need time to establish their position; it is incorrect to judge research performance on the basis of just a few years.

5. Mapping the structure of interdisciplinary research

Each year about a million scientific articles are published. How to keep track of all these developments, particularly the relations with other fields? Are there specific patterns 'hidden' in this mass of published knowledge, at a 'meta-level', and if so, how can these patterns be interpreted?

A research field can be defined by various approaches: on the basis of classification codes and/or selected keywords in a specific database, selected sets of journals, a database of field-specific publications, or any combination of these approaches. In this paper we take *micro-electronics* as an example. Along the above lines, we collected the titles plus abstracts of all relevant publications, for a series of successive years, thus operating on many ten thousands of publications. With a specific computer-linguistic algorithm we parsed the titles plus abstracts *of all these publications*. This automated grammatical procedure yields all nouns and noun-phrases (standardized) that are present in the entire set of collected publications.

An additional algorithm creates a frequency-list of these many thousands of parsed nouns and noun-phrases while filtering out general, trivial words. We consider the most frequent nouns/noun phrases as the most characteristic concepts of the field (this can be 100 to 1,000 concepts, say N concepts). The next step is to *encode* each of the publications with these concepts. In fact this code is a binary string (yes/no) indicating which of the N concepts is present in title or abstract. This encoding is as it were the 'genetic code' of a publication. Like in genetic algorithms, we now compare the encoding of each publication with that of any other publication. So we calculate 'genetic code similarity' (here: *concept-similarity*) of all micro-electronics publications pair-wise. The more concepts two publications have in common, the more these publications are related on the basis of concept-similarity and thus can be regarded as belonging to the same subfield, research theme or research specialty. In a biological metaphor: the more specific DNA-elements two living being have in common, the more they are related. Above a certain similarity threshold, they will belong to a particular species.

The above procedure allows clustering of *information carriers* -the publications- on the basis of similarity in *information elements* - the concepts ('co-publication' analysis). Alternatively, the more specific concepts are mentioned together in different publications, the more these concepts are related. Thus, information elements are clustered ('co-concept' analysis). Both approaches, the co-publication and the co-concept analysis are related by simple matrix algebra rules. In practice, the co-concept approach (Noyons and Van Raan 1998) is most suited for science mapping, i.e., the 'organization of science according to concepts'.

Intermezzo: For a super market 'client similarity' on the basis of shopping lists can be translated into a clustering of either the clients (information carriers, where the information elements are the products on their shopping lists) or of the products. Both approaches are important: the first gives insight into groups of clients (young, old, male, female, different ethnic groups, etc.), and the second is important in the organization of the super market.

In main lines the clustering procedure is as follows. We first construct a matrix composed by co-occurrences of the N concepts in the set of publications for a specific period of time. We normalize this 'raw co-occurrence' matrix in such a way that the similarity of concepts is no longer based on the pair-wise co-occurrences, but on the co-occurrence 'profiles' of the two concepts in relation to all other concepts. This similarity matrix is input for a cluster analysis.

In most cases, we use a standard hierarchical cluster algorithm including statistical criteria to find an optimal number of clusters. The identified clusters of concepts represent in most cases recognizable 'sub-fields'. Each sub-field represents a sub-set of publications on the basis of the discussed concept-similarity profiles. If any of the concepts is in a publication, this publication will be attached to the relevant sub-field. Thus, publications may be attached to more than one sub-field. The overlap between sub-fields in terms of joint publications is used to calculate a further co-occurrence matrix, now based on sub-field publication similarity.

To construct a map of the field, the sub-fields (clusters) are positioned by multidimensional scaling. Thus, sub-fields with a high similarity are positioned in each other's vicinity, and sub-fields with low similarity are distant from each other. The size of a sub-field (represented by the surface of a circle) indicates the share of publications in relation to the field as a whole. Particularly strong relations between two individual subfields are indicated by a connecting line.

In Fig. 2, the result for micro-electronics research is shown. The map clearly shows 18 sub-fields, represented by these clusters. Major sub-fields such as general micro-electronics, circuits and design, materials, circuit theory, mathematical techniques, liquids, and structure of solids can be observed. Meanwhile, we further developed our mapping procedure so that very recent updates of maps can be constructed.

Figure 2: *Bibliometric map of micro-electronics research*
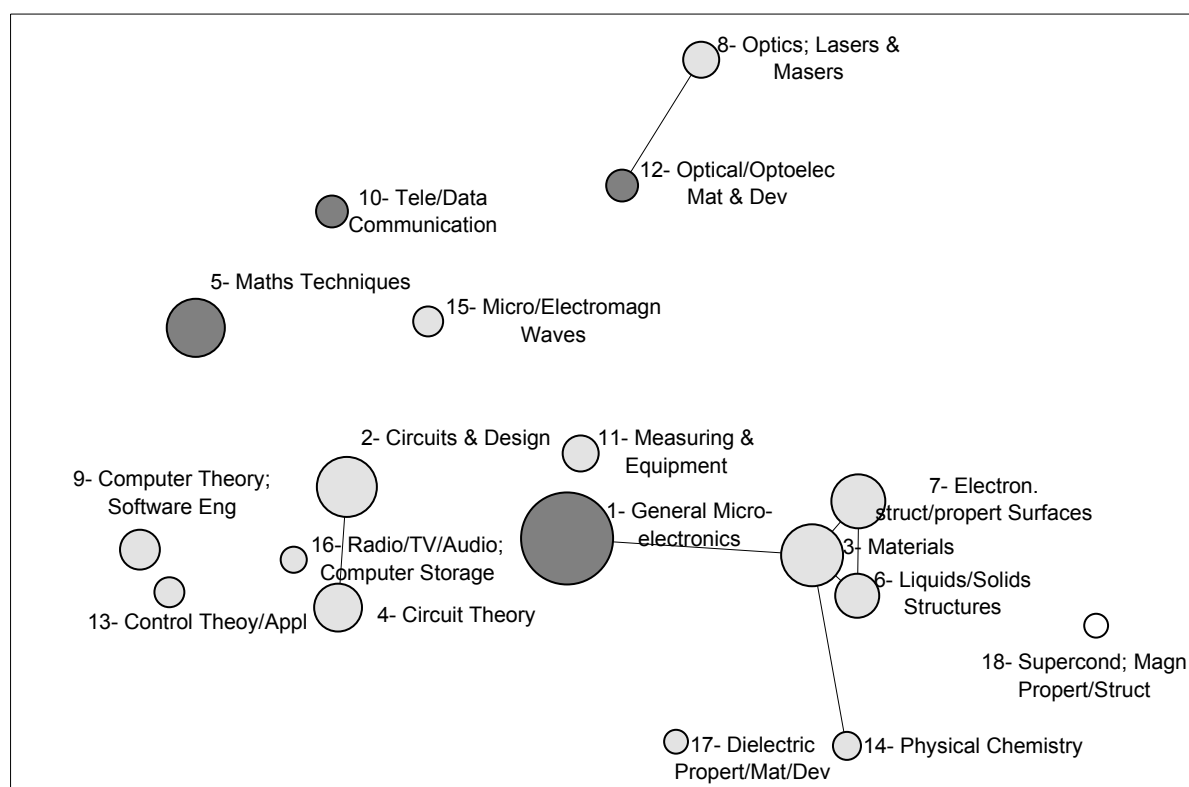
The map essentially represents a relational structure of clusters of publications, based on cluster-similarity measures. The clusters can be identified as research fields. The closer the clusters are, the more related the fields concerned. 'White' clusters (here only Cluster 18) are characterized by decreasing publication activity (worldwide), dark gray clusters (for instance Cluster 1) by increasing activity.

A next step (Noyons et al 1999) is the integration of both bibliometric methods we just described in this paper: mapping *and* performance assessment. It enables us to position actors (such as universities, institutes, R&D divisions of companies, research groups) on the world-wide map of their field, and to measure their influence in relation to the impact-level of the different sub-fields and themes. Thus a strategic map is created: who is where in science, and how strong? This 'next generation' bibliometric analysis includes a cinematographic representation of a series of maps of successive time periods. Recent developments can be found via our website[4]. This dynamic approach reveals trends and changes in structure, and even may allow 'prediction' of near-future developments by extrapolation.

Changes in maps over time (field structure, position of actors) may indicate the impact of R&D programs, particularly with respect to sub-fields characterized by research around social and economic problems. In this way, our mapping methodology is also applicable in the study of the socio-economic impact of R&D (Airaghi *et al* 1999). A similar mapping procedure can be applied to documents other than publications, for instance patents. Thus, maps of technology can be constructed.

6. Concluding remarks

We presented an overview of advanced bibliometric methods for the objective and transparent assessment of strengths and weaknesses in research performance, and for monitoring scientific, particularly interdisciplinary developments. First, we focused on the detailed analysis of research performance in an international comparative perspective. We applied our approach at the institutional level and showed that this level is the crucial starting point of the 'search for excellence'.

We demonstrated that our recently developed indicators are very informative and we concluded that advanced bibliometric methods are, particularly at the level of research groups, university departments and institutes, an indispensable method in evaluation studies. Not to replace peer review, but to support it. A number of specific problems   -and opportunities as well- for the application of bibliometric analysis in the social sciences were addressed.

In the second application, monitoring of scientific (basic and applied) developments, we showed that recent advances in bibliometric mapping techniques are promising. They are unique instruments to discover patterns in the structure of a research field. By adding 'communication linkages' based on the extent to which publications in a specific sub-field cite publications in other sub-fields, we are able to identify processes of knowledge dissemination (van Raan and Noyons 2002). Time-dependent analysis reveals the dynamics of scientific developments, with the possibility to focus on interdisciplinary developments. This is important, as we know that interdisciplinary cross-roads of basic and applied scientific fields are often the loci of discovery and technological innovation.

---

[4] www.cwts.leidenuniv.nl

References

Airaghi, A., J. Viana Baptista, N.E. Bush, L. Georghiov, M.J. Ledoux, A.F.J. van Raan, and S. Kuhlmann. 'Options and Limits for Assessing the Socio-Economic Impact of European RTD Programmes', ETAN Working Paper, EUR 18884, Luxembourg: Office for Official Publications of the European Communities, 1999. (ISBN 92-828-3721-1).

Horrobin, D.F. (1990). The philosophical basis of peer review and the suppression of innovation. *Journal of the American Medical Association (JAMA)* 263, 1438-1441.

Kroonenberg, P.M. and R. Van der Veer (1996). Internationale publicaties en kwaliteit. Een onderzoek naar het publicatiegedrag van Nederlandse pedagogen en onderwijskundigen (International publications and quality. A study of publication patterns of educational psychologists in the Netherlands). *Pedagogische Studiën 73*, 225-238.

Kyvik, S. and I.M. Larsen (1994). International contact and research performance. *Scientometrics* 29, 161-172.

Meertens, R.W., A.J. Nederhof, and H.A.M. Wilke (1992). Social Psychological Research in The Netherlands, 1980-1988. *European J. of Social Psychology,* 22, 93-100.

Moed, H.F. and Th.N. van Leeuwen (1995). Improving the accuracy of the Institute for Scientific Information's Journal Impact Factors. *Journal of the American Society for Information Science* (JASIS) 46, 461-467.

Moed, H.F. and Th.N. van Leeuwen (1996). Impact Factors Can Mislead. *Nature* 381, 186.

Moed, H.F., De Bruin, R.E., and 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, 381-422.

Moxham, H. and J. Anderson (1992). Peer review. A view from the inside. *Science and Technology Policy* 7-15.

Nederhof, A.J. and R.A. Zwaan (1991). Quality judgements of journals as indicators of research performance in the humanities and the social and behavioural sciences. *Journal of the American Society for information Science (JASIS)* 42, 332-340.

Nederhof, A.J. and E.C.M. Noyons (1992). Assessment of the international standing of university departments' research: a comparison of bibliometric methods. *Scientometrics* 24, 393-404.

Nederhof, A.J. and A.F.J. van Raan (1993). A bibliometric analysis of six economics research groups: a comparison with peer review. *Research Policy,* 22, 353-368.

Nederhof, A.J., Th.N. van Leeuwen and M.S. Visser (1997). A comparative bibliometric study of the Kurt Lewin Institute. Leiden: CWTS-report 1997-10.

Nederhof, A.J., Th.N. van Leeuwen and M.S. Visser (2000). Bibliometric profiles of academic psychology in the Netherlands. Leiden: CWTS-report 2000-05.

Noyons, E.C.M. and 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 and Technology (JASIST),* 49, 68-81.

Noyons, E.C.M., M. Luwel and H.F. Moed (1999). Combining Mapping and Citation Analysis for Evaluative Bibliometric Purpose. A Bibliometric Study on Recent Development in Micro-Electronics. *Journal of the American Society for Information Science and Technology (JASIST),* 50, 115-131.

van Raan, A.F.J. (1993). Advanced Bibliometric Methods to Assess Research Performance and Scientific Development: Basic Principles and Recent Practical Applications. *Research Evaluation*, 3,151-166.

van Raan, A.F.J. (1996). Advanced Bibliometric Methods as Quantitative Core of Peer Review Based Evaluation and Foresight Exercises. *Scientometrics* 36, 397-420.

van Raan, A.F.J. (1998). Assessing the social sciences: the use of advanced bibliometric methods as a necessary complement of peer review. *Research Evaluation* 7, 2-6.

van Raan, A.F.J. (2000a). The Pandora's Box of Citation Analysis: Measuring Scientific Excellence, the Last Evil? In: B. Cronin and H. Barskt Atkins (eds.). *The Web of Knowledge. A Festschrift in Honor of Eugene Garfield*. Ch. 15, p. 301-319. Medford (New Jersey): ASIS Monograph Series, 2000. ISBN 1-57387-099-4

van Raan, A.F.J. (2000b). The Interdisciplinary Nature of Science. Theoretical Framework and Bibliometric-Empirical Approach. In: P. Weingart and N. Stehr (Eds.). *Practicing Interdisciplinarity*. Toronto: University of Toronto Press, ISBN 0-8020-8139-8.

Van Raan, A.F.J and E.C.M. Noyons (2002). Discovery of patterns of scientific and technological development and knowledge transfer. In: W. Adamczak and A. Nase (eds.): *Gaining Insight from Research Information*. Proceedings of the 6th International Conference on Current Research Information Systems, University of Kassel, August 29-31, 2002. Kassel: University Press. Page 105- 112. (ISBN 3-933146-844).

Rinia, E.J. Th.N. van Leeuwen, H.G. van Vuren, and A.F.J. van Raan (1998). Comparative analysis of a set of bibliometric indicators and central peer review criteria. Evaluation of condensed matter physics in the Netherlands. *Research Policy* 27, 95-107.

VSNU (1994). *Quality Assessment of Research. Psychological Research in the Netherlands: Past Performance and Future Perspectives.* Utrecht: VSNU (ISBN: 90-801015-9-1).

VSNU (1995). *Quality Assessment of Research. Pedagogical and Educational Sciences. Past Performance and Future Aspects.* Utrecht: VSNU (ISBN: 90-5588-020-5