Merit, Expertise and Measurement

Monitoring knowledge flows, new performance indicators, and evaluation cultures in 21st century science, technology and society

CWTS Research Programme 2012 - 2015

Table of Contents

Summary

This new research programme of the Centre for Science and Technology Studies (CWTS), Leiden University, presents our research goals and priorities for the next four years. Our main concerns are monitoring and analyzing knowledge flows, and research evaluation. The programme introduces new approaches to these well-established goals of scientometric research. First, we will move from a longstanding tradition of data-centric methods justified by ad-hoc reasoning towards a systematic theory-based framework for developing bibliometric and scientometric indicators. We will also explore the potential of alternative forms of metrics ("alt-metrics"). Second, in interpreting and applying performance indicators we will increasingly base ourselves on the systematic analysis of current scientific and scholarly practices rather than only on general statistical arguments. Specific attention will be paid to the humanities and social sciences. We will also analyze the impact of research assessment exercises, and the performance criteria applied, on the primary process of knowledge production. Third, we will explore the possibilities and problems in assessing the societal impact of research ("social quality"). Increasingly, this dimension is becoming the second pillar of research evaluation next to scientific impact.

On the basis of our research programme, we will also start to systematically explore our possibilities to inform all actors in the scientific and scholarly system about new developments in science and scholarship, such as the emergence of new interdisciplinary fields (eg. the neurosocial sciences), science based innovation (eg. patent regimes), e-research and new applications of information and communication technologies (eg. virtual research environments), and new paradigms of scientific publication (eg. open access).

Our broadened research agenda can be seen as a response to the widespread use of bibliometrics in performance based research management, as a step to help prevent abuse of performance measures, and as a contribution to the systematic development of good evaluation practices.

Introduction: a new role for research information?

Scientists and scholars are confronted by complex changes in the way they create and communicate new knowledge and technologies (Borgman 2007). These transformations do not amount to a single-issue revolution, in spite of the hype about the world wide web, new social media or e-science (The Virtual Knowledge Studio et al. 2008). Also, they are not completely unique. Changes of comparable magnitude took place earlier, such as the "scientific revolution" of the 17th century and the emergence of the research university in Germany. Yet, they are fundamentally affecting all relevant dimensions of the practice of knowledge creation in the natural and life sciences, as well as in the technical and social sciences and the humanities and arts (Dutton, Jeffreys, and Goldin 2010; Bulger et al., 2011). To be sure, the world of research is not unique in being confronted with radical changes. Other social areas are also transformed by the explosive mix of globalization, commercialization, and technological change. This may sound like a consolation, but it actually adds to the process of complex changes because science and scholarship are increasingly caught up in the transformation of these other realms of society. Science and scholarship are shaped by these additional dynamics as well, because the way the scientific and scholarly system functions is fundamentally affected (Wouters, Beaulieu, Scharnhorst, & Wyatt, 2013). Consequently, researchers are in the epicentre of nothing less than a tornado. They need to adapt their ways of working while they strive to maintain proven standards of quality in their work. This also requires institutional innovation and radical improvements of the way universities and research institutes are being managed. Lastly, it puts tremendous pressures on national and international science policies that somehow need to accommodate the tension between serving an increasingly sharp international economic competition on the one hand, and creating shared standards and open international data, information and knowledge infrastructures on the other hand.

These challenging transformations have induced a new need for *information about research.* This has stimulated the scientific analysis of the knowledge production process, that can produce this information at all levels for researchers, managers and science policy makers. These needs can only be met by advanced applications of information science. This research programme aims to contribute to these applications by developing new ways of:

1. monitoring and analyzing the scientific and scholarly system, knowledge flows, communication networks in science and scholarship, and interactions between research and society;
2. evaluating research including the assessment of the societal impact and interactions of research.

The way research and research performance are being measured, the kinds of scientific and scholarly work that visibly make a difference in monitors and assessments, and the criteria applied in evaluation and management will shape the ways the creation of new knowledge is organized. This will determine the key research agendas of the near future. Research evaluation is relevant in contributing to these new agendas. Although not often used in this way, research assessments can proactively support innovative practices in all fields of scholarship and research.

This research programme aims to discover the main characteristics of what may be possible in these newly emerging evaluation practices and thereby lay the groundworks for good evaluation practices and for new frameworks for knowledge and technology monitors. This programme is based on the mission of the Centre for Science and Technology Studies CWTS at Leiden University.

Mission CWTS

The Centre for Science and Technology Studies (CWTS) investigates the development of science and technology (S&T) using large-scale databases of scientific and technical publications. CWTS is a leading provider of science &

technology indicators and of performance and benchmark studies of scientific groups and institutes, including bibliometric mapping and network analysis of science. The CWTS Leiden Ranking is recognized as one of the most reliable citation based rankings of large universities. CWTS has developed longstanding relationships with the most prestigious universities in the world and is a recognized leader in the field of bibliometrics, scientometrics and informetrics. Currently, CWTS is exploring new venues in the fields of webometrics and science & technology studies.

Our ambition for the next 5 years is to be an international leader in:

1. quantitative science, technology & innovation studies, among which indicator development and research evaluation;
2. integrating knowledge in scientometrics and in combining knowledge from scientometrics with insights from science and technology studies, demography, economics, and the social sciences more generally;
3. the analysis of the role of indicators in the sciences (including the social sciences and humanities).

CWTS has played a crucial role in the development and application of citation based indicators for research evaluations and assessments at all levels of aggregation (international and national monitors and studies, university-wide evaluations, assessments of research groups, benchmark studies, and bibliometrics at the individual level) (Van Raan 1988; Moed 2005a; Moed, Glänzel, and Schmoch 2005).

A key asset of the centre is its high quality processing of citation databases and the data expertise in managing large citation databases of its staff. As a result, the citation analyses produced by CWTS are based on corrected and verified data of high quality, different from the public citation indexes commonly available on the web. We take this combination of expertise and data infrastructure as point of departure to meet the challenge of the complex changes in the scientific system that we are now witnessing. CWTS is certainly not the only scientometric centre in this endeavor. But as one of the oldest, most experienced, and largest scientometric centres in the world it has the responsibility to take the initiative in the required methodological and theoretical advances in the field. With this new research programme, we aim to bring scientometrics to a new level of quality, in close collaboration with our colleagues in the field. This should lead to new international standards of quality for assessments and science & technology indicators.

Transformations in knowledge flows and evaluation

We already mentioned the combination of globalization, commercialization, and technological development that can be held responsible for many of the changes in the practices of researchers in most disciplines. The disadvantage of speaking in these container terms is that they tend to make invisible the intricate interactions at the level of daily practices that actually produce these world encompassing trends. They tend to create a metaphor of transformations happening upon us rather than analyzing the active role of social actors. In the case of scientific developments this is perhaps even more disadvantageous given the active role of scientific results and instruments in these global trends. In the following, we focus on those changes that directly influence the various processes of knowledge creation and evaluation and on those changes in which researchers play an active role.

Globalization

Research is becoming a global enterprise in a more fundamental sense than before (Suresh 2011). Although scientific research has been internationally oriented from its very beginning (Collins 1998), the scale has increased and the characteristics of research funding and collaboration are shifting (Sonnenwald 2007; Shrum, Genuth, and Chompalov 2007; Wagner 2006; Olson et al. 2005; Olson, Zimmerman, and Bos 2008). In many fields, albeit not in all, international networks have emerged as the "natural habitat" for researchers and scholars (Royal Society 2011). The nature of these networks may vary, a feature which is often overlooked in the literature on collaboration networks. Although co-authorships are the most frequently used indicator of scientific collaboration, these concepts should not be seen as identical (Laudel 2002; Melin and Persson 1996; Shrum, Genuth, and Chompalov 2007). For example, in philosophy of science where many scholars are focused on the production of single authored monographs, the network of their colleagues is visible in the conferences where they meet. In physics, these networks take the form of co-authorship networks, some of them very large. In clinical research, the network may predominantly be an international meeting place between research and industry where resources are allocated. These international networks do not imply that science does not have a national dimension anymore. The national role of scientific research may actually even become more urgent, for example in the creation of a homegrown research base in upcoming scientific powers such as China, Brazil or Iran. This may be particularly relevant for the design and engineering disciplines and for the social sciences and humanities (ACLS 2006; KNAW 2010). But these national dimensions should now be interpreted in the context of the interlocking international networks.

A potentially new trend is the internationalization of the institutional arrangements of the universities. During the 19th and 20th century, the nation state was the linchpin of the institutional configuration of academic research. In the 19th century, the research university was created in Germany and the concept of the "discipline" emerged in France. The new institutions were strongly influenced by the creation of European national states. In the 20th century the state has only become more important in the funding and organization of research. But now we see the first glimpses of research funding and management going truly global, for example in the co-funding schemes organized by the EU, the ESF and recently the NSF.

If we are indeed witnessing a new stage in globalization of research, we can expect two effects. First, new national interests will start to play a role in research on a global scale (for example those of the BRICs countries). This may also lead to new forms of secrecy and classified research: globalization may not always promote openness. Second, a new pool of human talent may enter the stage and innovate scientific research and scholarship, provided that the poverty trap in developing countries can be overcome. This may also affect the current gender and ethnic imbalances in the scientific and scholarly workforce.

This new form of globalization has an immediate effect on research evaluation because it increases global competition for expertise. The ramifications of this development are not yet clear. It may enable the best research institutes to tap into the pool of worldwide talent and dominate their fields even more than they already do. In this sense, global university rankings and journal impact factors may become even more attractive as a management tool notwithstanding their evident shortcomings (Weingart 2005, Moed & van Leeuwen, 1996). The cultural diversity in the world of research will continue to increase. Obviously, the extent to which researchers are able to operate in these international networks and have an impact on the international research agenda in their field has become an important aspect of research assessment exercises. How these networks function has become an important research topic in science, technology and innovation studies (Waltman, Tijssen, and van Eck 2011; Royal Society 2011; Börner 2010; Olson, Finholt, and Teasley 2000).

Commercialization and valorization

Scientific and scholarly research has become a key asset in the creation of wealth and in global economic competition (Slaughter and Leslie 1997; Mirowski and Sent 2007; Blume 1992). As a result, scientists and scholars have been asked to contribute with their research to the making of profit, in particular since the 1980s (in Europe and North America). Since the early 1990s this requirement has been expanded to one of being socially relevant, a concept originally developed in the early 1970s (in Canada and the Netherlands in particular) (Blume 1986). The analytical concepts used to analyze the interactions between the scientific system and society at large differ markedly depending on whether for profit companies or social movements are put central. Science and technology policies tend to focus on the former. This is not only relevant to the technical and social sciences: the emergence of the new "creative industries" has put the interactions between the humanities and society more central. Science and technology studies research since the 1970s has resulted in a vast number of case studies of interactions between academia, industry and society and has produced an interesting set of competing, sometimes compatible, theoretical approaches with which to analyze "valorization", "social relevance", and "social quality" (Etzkowitz & Leydesdorff, 2000; Nowotny, Scott, & Gibbons, 2001).

If researchers are being asked to add the social and economic valorization and relevance of their work to their agenda, it goes without saying that this has led to new criteria for research evaluations and assessments. Both in peer review of research projects, annual appraisal interviews, and institutional research assessment exercises, the question about the social relevance and economic impact is being asked. This has created a serious problem for both researchers and evaluators because these wider impacts of the outcomes of research (different from the more narrow research output in the form of publications) are very difficult to prove and evaluate. This is mainly caused by the complex nature of the interactions between academia, industry, and the public sector. The way particular research influences professional practices or is included in the development of a new product can take many forms. These are not always easy to detect. The time scale of these interaction processes varies enormously. Because of the multidirectional interactions, causality is almost never provable or even obvious. In addition, it is not yet clear what the relationship is between the social and economic quality of research and its scientific and scholarly quality. Recent studies have shown that these two dimensions require different kinds of work from researchers. Articulation work is of a different nature than the often technical tasks at the lab bench or in the field. Yet, they are not completely unrelated. It can be expected that these interactions will vary depending on research mission, internationalization, style of research and the structure of the specific economic or cultural sector involved.

In recent years, researchers in the UK and the Netherlands have pioneered different approaches to enable the accurate evaluation of valorization and social quality of research. Methodologically, these approaches are mainly based on surveys, interviews and qualitative appraisals of the interactions between researchers and other societal actors, sometimes enriched with available statistical material. As a result, more often than not the evaluation methods focus on the perception of social quality or valorization as proxy for the impact of research. To what extent this can be systematically extended to find more robust evidential material is still an open but increasingly urgent question.

Technological interactions

Science and technology are involved in a complex spiral of interaction. In many fields, albeit not in all, research results lead to new forms of technology. At the same time, technical research in most areas has increasingly been built on scientific first principles, although professional tacit knowledge remains important. These new technologies in their turn are finding applications outside of their original context. In the form of new research instruments, this may lead to completely new research problems, the redefinition of older problems or the confluence of formerly disconnected specialties. For example, new imaging technologies have had a huge variety of practical implementations that have strongly influenced medical research, the cognitive sciences and recently also fields such as economics.

As a result, the practices at the laboratory bench, in the field or in the literary archive are increasingly integrated using advanced inscription, imaging and database technologies. These changes in practices are not represented very well in the formal scientific literature. The scientific literature has always been a formalized and stylized representation of selected scientific results and was never a realistic portrait of research practices (Bazerman 1988), but this relationship is now becoming even more strained. This raises a number of questions about the validity of research evaluations that are mainly based on the analysis of the formal scientific literature. For example, to what extent is database work visible in bibliometric assessments? Which new forms of output are emerging as potential sources for evaluation? In what ways is the role of the author shifting? Many of these questions are also tackled by journal editors in their management of the peer review process. But these experiences are not yet systematically mined for lessons for research assessments at the level of research institutes. This also holds for the experiences of journal and book publishers who are responding to changing research practices with experiments of new formats for scientific journals, such as web based "live journals" or the presentation of live annotated data streams. These developments will lead to demands for new types of metrics that are able to analyze the contributions of researchers in these new formats.

This tendency to partly undermine the validity of traditional bibliometrics is exacerbated by the development of new media and the re-emergence of visual forms of knowledge. Youtube videos, blogs, and wikipedia are examples of new formats used by scientists and scholars in a search for more engaging and socially relevant forms of communication. They are also important platforms for sharing knowledge in peer communities. Some of these communications are regulated by strict forms of peer review, others are new forms of self-publishing. To what extent they should play a role in research evaluations is still an open question. To address these questions, it is imperative that bibliometrics teams prioritize collaboration with webometrics experts.

The latter is the more important since the transformation of the media landscape has also affected the methodologies of research evaluation and assessments. Whereas the field of scientometrics was initially mainly defined by the constraints of the Science Citation Index databases, publishing on the web has opened up a whole new set of methodologies from other fields (computer science, data science, visualization). The field of scientometrics has been slow to adopt these new possibilities, partly as a result of a lock-in to the Web of Science based paradigms. This also holds for CWTS itself. This research programme aims to redress the balance between tradition and innovation in our research.

Growth and change of evaluation practices

The process of research evaluation is not only shaped by the dynamics in scientific knowledge production and communication, but also by policy cycles in the management of the university systems and related configurations of research institutes. Indeed, we are not merely speaking about applying knowledge of the research system to policy issues, but about the rise of scientific governance as a new paradigm (Irwin 2007). Research evaluation is not a uniform system across different national research and innovation systems. On the contrary, there is a large variety of evaluation practices across countries, disciplines, and types of universities that each have varying rhythms in which they respond to political pressures. The short-term consequences of evaluation for the further development of research may also vary strongly. This has to do with the overall level and organization of the allocation of resources for universities and academic research. In some parts of the world, the relative share of national product invested in scientific research is decreasing, whereas other countries are investing huge sums in order to enter the global scientific competition. The relative size of disciplines varies by country, as are their societal appreciation. An important dimension of variation among different research systems is the way quality control mechanisms have been coupled to funding mechanisms. For example, in some countries, fairly direct translations from number of citations into resources for research exist, whereas more stable forms of block funding still dominate the allocation of funds in other countries.

The different practices and regulations of institutional research assessments have not yet been studied sufficiently (Lamont 2009; Langfeldt 2004; Martin, 2011), in contrast with journal and grant proposal peer review which has been studied extensively (Daniel 1993; S. Cole, Cole, and Simon 1981; Bornmann 2011; Chubin and Hackett 1990; Rinia et al. 1998). As a result, we do not yet know enough about the varieties of "evaluation cultures" that link up to "epistemic cultures" (Knorr-Cetina 1999) in which knowledge is being produced. It is difficult to distinguish generic trends in research evaluation from singular cases. Yet, we can tentatively see some common patterns in developing evaluation practices:

1. Research evaluation is now encompassing the universities as whole. There are virtually no areas where researchers do not - in one form or other - have to respond to regular assessment exercises. This has gone together with the application of quantitative performance indicators in all areas of science and scholarship, including the humanities.
2. These indicators have been applied far outside their original scope or context. Although indicators of course do not develop autonomously, they have in a certain way started to live a life of their own. The concept of scientific "quality" or "impact" has been deeply transformed by the emergence of the Science Citation Index. This can be expected to happen again as a result of new indicators. Performance indicators are social indicators and as such they construct social reality as much as they measure it.
3. The ambitions of the evaluation systems keep developing from assessing past performance, to picking the winners, to detecting emerging trends in funded research, to monitoring the research process as a value creating process. These different levels of ambition feed upon each other, and require quite different types of indicators and increasing levels of reliability and granularity. Indicators for scientometrics of individual researchers are being applied increasingly.
4. The relevant dimensions of evaluation have steadily increased. Societal impact and valorization have been added to the dimension of scientific quality and impact. Moreover, the extent of international collaboration has emerged as an independent dimension related to scientific quality, as have different measures of vitality and feasibility of research agendas.

These developments in the evaluation systems and practices have important implications for the requirements of formal performance indicators. Here we can see shifts from descriptive to analytical indicators, from simple counts to complex normalized weighting, from ex-post to ex-ante assessments, and from one-dimensional to multi-dimensional analysis. Although we focus here on the scientific system, these trends in evaluating are not unique for science. They are part of a more generic trend in accountability practices in a variety of societal sectors and in the way strategic intelligence is managed in business processes. How these different regimes of accountability relate to each other is an interesting question in itself (Woolgar 2002).

Research questions

How can we improve our understanding of the dynamics of science, technology, and innovation by the measurement and assessment of the scientific and scholarly system, in particular of scientific products, communication processes and scholarly performance? This is the overarching question of this research program. To answer this question, two specific research questions are central:

1. How do scientific and scholarly practices interact with the "social technology" of research evaluation and monitoring knowledge systems?
2. What are the characteristics, possibilities and limitations of advanced metrics and indicators of science, technology and innovation?

These questions will be pursued in all research themes of this program. Taken together they shape the analytical framework for the applied research and contract research projects that are performed by CWTS and its company CWTS BV.

This program has a modular structure. It consists of two main parts (a theoretical/methodological and a thematic part) and three sections on the research chairs. Each module has a relatively autonomous position in that only a limited number of staff (organized in one or more working groups per theme) are responsible for its execution and further development. Yet, the program as a whole functions as one program because each module develops synergies with other parts of the program in the form of shared projects, common input into service development by the CWTS BV, and shared discussions and presentations at the weekly research seminars.

Methodological and theoretical advances

In the past, methodological research at CWTS focused on two topics: the development of bibliometric indicators for research assessment and bibliometric mapping of science. Both topics are highly relevant from the point of view of evaluative bibliometrics. We plan to maintain the focus on these two topics, but there will be some important shifts in emphasis. Also, the topic of bibliometric mapping of science will be broadened into bibliometric network analysis. Moreover, we will gradually extend the scope of our methodological work by studying the possibility of non-bibliometric scientometric indicators, that are not based on the current scientific literature. Our methodological research is closely linked to our theoretical assumptions about science and technology indicators in general and citation analysis in particular. CWTS will renew its interest in citation and communication theories, combining theoretical work in science and technology studies and the history of science with new models of the citation and publication process.

In the longer term, we expect that our work on evaluation and indicators will lead to fundamentally novel approaches to measuring scientific quality. We see quality as a relational attribute which is shaped and certified in communication processes in the scientific community and in interactions between science and society. The current citation based indicators are only able to capture a limited set of dimensions of quality, mainly the impact of a publication on the scientific community as visible in explicit formal citations. Many important dimensions of quality cannot be measured in this way. We do not expect that all dimensions of quality will be measurable. However, it should be possible to support peer judgements of quality with more informed empirical evidence, including quantitative data, than is currently possible. In this respect, the fundamental question of what scientific and social quality of research is and how it can be measured and represented (a foundational question of the field) needs to be revisited.

Bibliometric indicators for research assessment

Both at CWTS and elsewhere, the development of bibliometric indicators for research assessment has long been done in a somewhat informal way. Indicators were developed without explicitly incorporating them in a broader mathematical or statistical framework, and indicators were justified mainly using empirical arguments. This resulted in a data-centric approach where the rationale for, and interpretation of, the chosen indicators were often developed in an ad-hoc fashion. The theoretical work that was performed at scientometric centres was not always strongly connected to empirical studies. In this work we will build further on previous theoretical work at CWTS (Nederhof and van Raan 1987; Noyons and van Raan 1998; Peters and van Raan 1994; van Raan 1990, 2000, 2001a, 2001b; Waltman et al. 2011; Waltman and van Eck 2011) to create new theoretical and methodological models for indicator development and testing.

At CWTS, we will move towards a more theoretically-oriented approach to the development of bibliometric indicators. In this approach, indicator development will become more and more based on explicit theoretical models of the scientific publication and citation process. In this framework, the indicators will be judged mainly based on their mathematical and statistical properties. These models will for instance allow us to distinguish between observable and non-observable features of the publication and citation process (e.g., between the observable concept of citation impact and non-observable concepts such as scientific influence or quality). Model-based indicator development has the advantage of making an explicit distinction between what one intends to measure and what one is in fact measuring. This will help us to study the properties of bibliometric indicators (e.g., validity and reliability or bias and variance) in a more formalized way. The limitations of the indicators should be made explicit as well. For example, a complex concept such as scientific impact cannot be measured by one indicator (Bollen et al. 2009; Moed 2005b). This is the reason we have moved more systematically towards a portfolio approach to performance indicators.

In the development of bibliometric indicators, special attention will be paid to indicators for the humanities and the social sciences and to indicators for evaluation at the level of individual researchers. In both cases, the challenge will be to develop indicators that provide meaningful results based on relatively small amounts of publication or citation data. For the development of meaningful indicators for the humanities and social sciences, more research is needed on the specific characteristics of research and communication practices in these fields. For example, are citation patterns between books comparable to those between journals? Are indicators based on highly cited items in citation databases meaningful for the humanities and social sciences? How robust can indicators be at the level of the individual research group? Will we need a new concept of an indicator at the individual level?

The main technical issues to be addressed in the near future are:

1) Field normalization. The current field normalization method of CWTS relies on the assumption that Web of Science subject categories are more or less homogeneous in terms of citation characteristics. This assumption is often violated, which can lead to quite significant biases in bibliometric indicators. New field normalization methods need to be developed that rely on more realistic assumptions.
2) Database coverage. Bibliometric indicators are sensitive to the coverage of the database that is being used. Because the coverage of databases such as Web of Science changes over time, the effect of database coverage on bibliometric indicators changes over time as well. An important objective is to minimize the effect of database coverage on bibliometric indicators (van Leeuwen et al. 2001).
3) Stability. In the past, CWTS paid relatively little attention to the issue of the stability or robustness of bibliometric indicators. The issue of the stability of indicators has a high practical relevance and therefore needs serious attention. To deal with the issue of stability, our indicators will be complemented with confidence or stability intervals. Also, average-based indicators (e.g., MNCS) may be complemented with indicators based on counting highly cited publications (Tijssen et al. 2002).
4) Fractional counting. This relates to the many publications that are co-authored and the way these publications are aggregated to the levels of groups, institutes, disciplines and nations. Sometimes it makes sense to count every publication by every author as one. Sometimes it makes more sense to fraction the publication over the number of authors. This problem merits a more systematic approach than is currently customary in the field of bibliometrics.
5) Document types. This technical problem relates to the question what types of documents in the citation databases should contribute to the citation scores and to what extent. This depends on the role of different documents in the process of knowledge creation and dissemination and on the way they are recorded in the database.

We acknowledge that the above issues are not new. In fact, some of the issues already have a long history in the field of bibliometrics. Nevertheless, we feel that a serious research investment in these issues is again needed. Until recently, many of the above issues have not received sufficient attention, or the way in which they were treated has not been sufficiently rigorous. This is a shortcoming of the field of bibliometrics as a whole. The current state of affairs is clearly unsatisfactory, and therefore the whole indicator framework of CWTS is currently undergoing a fundamental rethinking. Weaknesses in the framework need to be fixed, and the strengths of the framework need to be made more explicit by means of solid mathematical and statistical arguments.

Given the service activities of CWTS, an important element of our methodological work on bibliometric indicators will be the interplay between on the one hand our theoretically oriented indicators research and on the other hand the practical application of indicators for research assessment and science policy purposes. We will pay special attention to the translation of the results of our theoretically oriented research into clear recommendations and guidelines for the practical application of bibliometric indicators. Examples of highly relevant applied topics include the *h*-index, journal impact indicators, and university rankings.

To further this work and to stimulate collaboration among the scientometric centres in the world, we will publish a manual on the most commonly used CWTS indicators with a detailed explanation of their possibilities and shortcomings, their role with respect to peer review judgement of quality, and the precise way they are being calculated. We have also renewed the Leiden Ranking and provided a detailed explanation of the ranking methodology. Global university rankings have become ever more important in recent years. Most of these rankings are not transparent, their limitations are not always clear, and they are difficult to compare. An important objective of the Leiden Ranking is to stimulate transparency and accountability of university rankings.

Bibliometric network analysis

Bibliometric networks are networks of, for instance, publications, journals, researchers, or keywords. These entities are linked to each other based on citations, co-citations, bibliographic coupling, keyword co-occurrences, co-authorship, etc. Bibliometric network analysis is concerned with the analysis of such networks. Instead of focusing on the properties of individual entities in a network, bibliometric network analysis concentrates on the way in which relations between entities give rise to larger structures, such as clusters of related publications or keywords. In this sense, bibliometric network analysis is closely related to the analysis of complex systems.

The main objective of our research into bibliometric network analysis will be to provide content and context for research assessment purposes. Providing content means providing insight into the type of research an organization or an individual is doing. For instance, on which topics is an organization focusing its research activities? How do these topics relate to each other? Providing context means providing insight into the ‘scientific environment’ in which an organization or an individual is operating. This includes the identification of competing organizations and the detection of emerging research areas.

Our research into bibliometric network analysis has already resulted in a new classification system of scientific fields (Waltman & van Eck, 2012). In the short term it will also support the development of an integrated computer system for large-scale analysis and monitoring of bibliometric data. This system will provide easy access to a large variety of bibliometric statistics and will offer extensive support for selecting and analyzing relevant scientific literature and for identifying and assessing key players in this literature. The system relies on various types of visualizations to present bibliometric statistics in an attractive and easy to understand way. Bibliometric network analysis techniques are used to offer a large degree of flexibility in selecting and analyzing the relevant literature. The system should help the user to understand both the power and the limitations of the indicators.

Given the above objectives of our research into bibliometric network analysis, there are three important technical problems which we will tackle in the next few years:

- *Delineation*: How to delineate a scientific field, or how to delineate the literature on a certain topic? How to identify the oeuvre of an author? We have already developed some basic methods for delineating scientific literature, but these methods need to be refined in various ways.
- *Partitioning*: How to identify topics or themes in a selection of literature? How to deal with overlap of themes or with hierarchically organized structures (e.g., a theme that consists of a number of sub-themes)? And when a partitioning has been found, how to determine the stability or robustness of this partitioning? We have already developed a clustering technique that successfully handles some of these issues in the VOSViewer software (van Eck and Waltman 2010), but this technique needs to be extended in order to properly deal with overlap and hierarchy. The issue of stability or robustness also needs further investigation.
- *Mapping and visualization*: We make a distinction between mapping and visualization. Mapping refers to the traditional topic of bibliometric mapping of science. Visualization is a broader topic, dealing with any type of visual presentation of research assessment results. Visualization includes mapping as a special case. Visualization may be interactive, with computer software that allows users to interactively explore the research assessment results that are presented. Our mapping research will deal with developing improved mapping techniques and measuring the stability of mapping results. Our visualization research will deal with topics such as labeling and coloring of mapping and clustering results. More generally, we will experiment with new ways of presenting and exploring research assessment results. We will also seek to extend our mapping and visualization techniques to enable the presentation of dynamic developments rather than the current static snapshots.

Research into bibliometric network analysis is of a methodological nature. The methods and techniques developed in this line of research can support the various research themes at CWTS as well as the service activities of the institute. Our focus will be mainly on supporting and improving the research assessment activities of CWTS. We are also interested in the use of bibliometric network analysis for analyzing the properties of scientific systems and their dynamics (e.g., how a certain field develops over time), but this will be mainly developed in the framework of our theme "Scientometrics as a Social Science".

Diverse data and altmetrics

The data available for monitoring or assessment purposes has increased substantially in recent years and we can expect this trend to continue more rapidly in the future. The existing multidisciplinary citation indexes, Web of Science and Scopus, will probably continue to extend their coverage. Disciplinary databases with uniquely enriched content facilitate citation analysis more frequently than in the past (e.g. Mathscinet, Spires). Academic repositories, preprint archives and other scientifically relevant web based content are among the other data sources that will become more accessible for bibliometric analyses in the future. This wealth of data poses both opportunities as well as challenges to scientometricians. No longer do we depend on a single static data source from a single provider. We are able to choose among several options. At the same time, however, this necessitates a thorough knowledge of the available data sources in order to make informed choices regarding the source or combination of sources that can best be used to address particular research questions.

In addition to this gradual extension of traditional citation databases, we are witnessing the emergence of web based alternatives to traditional metrics. These are based on the rise of the web as a common platform of scholarly and scientific work and social transactions. Academic blogs, twitter, Facebook pages, applications and groups, and web based bibliographic platforms such as Zotero and Mendeley all open up vast new sources of data that can be used to construct and measure proxy indicators of scholarly activity and impact: alt-metrics (http://altmetrics.org/manifesto/).

These developments in scholarly use of the web have also led to hybrid forms of traditional and alt-metrics forms of assessment. For example, one can now routinely calculate one's h-index on the basis of Google Scholar data with a simple web based service http://www.harzing.com/pop.htm (Harzing 2010). However, the quality of this data is often even less clear than the data in the available scientific repositories and citation indexes. The risks of applying this data in evaluation contexts is comparably larger and are in urgent need of systematic study. In this research, the interaction between technologies of management and control (Beniger 1986) and "technologies of narcicism"[1] will be

[1] We thank Diana Hicks for suggesting this term.

a central analytical perspective. We will also aim to generate thorough knowledge about specific data sources and about specific collection strategies of web and transactional data. We will determine for each data source and collection strategy the benefits (added value) and limitations for applications in monitoring and evaluation studies (Wouters & Costas, 2012). With respect to web based data, we will study in particular the following types of data:

- web based citation data:
  - Google Scholar and Google Book citations:
  - various forms of open citation systems;
  - the use of web based referencing systems (Mendeley, Zotero, etc.);
- usage statistics, both source-specific and aggregated;
- hyperlink data structures;
- various forms of alternative metrics which are advocated on the web;
- transactional data regarding the societal impact and valorization of research;
- developments in structured content schemes.

In addition to publishing our finding in the literature, we also hope to contribute to manuals and wikis about the use of web data for analyzing the scientific and scholarly system (Thelwall 2005).

Communication and citation theories

A large part of the bibliometric/scientometric literature is concerned with research evaluation. Two perspectives are predominant, namely the methodological perspective and the application perspective. The methodological perspective largely focuses on the development and analysis of bibliometric indicators. To a lesser extent, attention is paid to the way in which bibliometric indicators should be used and interpreted. The application perspective uses bibliometric indicators to actually perform research evaluations, for instance at the level of countries, organizations, individuals, or journals.

In the bibliometric/scientometric literature, there seems to be relatively little interest in a deeper reflection on the use of bibliometric indicators for research evaluation purposes. Theoretical questions on the mechanisms at work in bibliometric evaluation processes are largely ignored in the research agenda, although they have been discussed in a more topical way from the very beginning of the field. Examples of such questions include: What do citations measure? If citations do not always reflect intellectual influence or some similar notion, can the use of citation-based indicators for evaluation purposes still be justified? How can researchers be expected to adapt their behavior to the indicators that are used to assess their performance? If researchers adapt their behavior, how does this influence the validity of performance indicators? In what ways do performance indicators feedback into the primary process of knowledge creation? These questions receive relatively little attention in the literature. And when these questions are being addressed, informal arguments rather than solid analyses are predominant.

Our purpose is to address questions such as those mentioned above in a formal model-based framework. We will formulate explicit and precise assumptions on research evaluation processes and in particular on the way in which researchers can be expected to behave. The assumptions we make will provide us with models that we can use to study the mechanisms at work in bibliometric evaluation processes. Moreover, due to the model-based nature of our research, we can study the relevant mechanisms in a much more rigorous way than is typically being done in the literature. In this way, our research may lead to a better understanding of bibliometric evaluation processes and, in addition, it may yield suggestions on how these processes can be improved, for instance by changing the indicators that are being used.

In the past, CWTS has mainly relied on "the standard account" of citation analysis. This account readily acknowledges that authors frequently do not give credit where credit is due, but on average citations are still seen as a valid indicator of academic recognition (Nicolaisen 2007, 623). This approach has also been adopted by CWTS (Nederhof and Van Raan 1987). Although many indicators have been partly validated in this way, the main argument is grounded in empirical findings and correlations, rather than in theoretically grounded arguments. As a result, the question "what does a citation actually mean?" is usually answered in a pragmatic way, emphasizing different dimensions of "quality" or "influence" in different cases. This state of affairs is unsatisfactory, if only because it hinders the development of transparent quality standards of indicators for research evaluation. Our theoretical work in the next four years aims to address this problem and give a new impulse to the work on citation theories and models in the field. We will not aim for one grand overarching theory of citation that will provide the foundations of citation analysis once and for all. This would be a chimera. Rather, we aim for models of scientific communication that will help us establish a more precise understanding of citation patterns and other relationships between entities in scientific databases (both publications and other forms of scientific output).

This theoretical work should also clarify the way performance indicators have changed the citing behavior of scientists and scholars (Weingart 2005, Martin, 2011). State of the art knowledge in science and technology studies and cultural anthropology on accounting practices and the role of quantification and standards will be particularly useful in this respect (Lampland and Star 2009). We will draw upon both our methodological work on bibliometric indicators and our ethnographic studies of the evaluation process in research institutes and its impact on the creation of knowledge. On the basis of these analyses, we hope to contribute to the development of alternatives for assessing scientific ingenuity beyond citation and impact indicators.

Thematic approaches

Research assessment and scientific practice

Research evaluation may take different forms: annual appraisal interviews, institutional research assessment exercises, and global assessments of national science systems. In this theme, the focus is on the actual application of bibliometric techniques and instruments, and the way bibliometric techniques can be helpful in measuring research performance, solving important problems policy makers come across, etc. This theme is strongly related to our service role in evaluation practices. The contract research CWTS performs for clients has been an important incentive for innovation and will remain to do so. In this theme we will study how bibliometric and scientometric indicators are actually applied in research evaluation. Which portfolio of performance indicators is the most useful in different evaluation contexts? How are the indicators interpreted by the various parties in the evaluation exercise? Which aspects of scientific and scholarly communication and performance are foregrounded in these assessments and which are put in the background? How do the specific characteristics of the field play out in the assessment procedure and how do they shape the outcome? Here we will pay special attention to the humanities and social sciences. Also, the recent trend to apply bibliometric indicators at the level of the individual research group and researcher will be analyzed. And last but not least, how do the various parties deal with uncertainties, error and bias?

The key question this theme will tackle is that of the relationship between expert peer review based evaluation and bibliometric indicator based evaluation. In the past, CWTS has dealt with this tension in a pragmatic way. As explained above, we will move towards a more systematic, theory based, approach in which we will probe in much more detail how expertise develops in particular scientific fields in relation to the bibliometric insights of those fields. We will not assume that the two ways of evaluating the quality of scientific and scholarly work are diametrically opposed: this would amount to setting up a straw man. In practice, peer review and bibliometrics are combined in a variety of ways. But how these combinations are developed by both evaluating institutions and the researchers that are being evaluated is not self-evident. Because it is exactly this interplay where the criteria for scientific quality and impact are being developed, we will zoom in on this aspect of the practice of evaluation.

This raises the question of how we define the process of evaluation. First, evaluation will be analysed from the research group's perspective. Evaluation is a more complex interaction than simply the measurement of the performance of the researcher. It is a communication process in which both evaluators and the researcher under evaluation define what the proper evaluation criteria and materials should be. The key outcome of evaluation systems is not only the conclusion with respect to the future prospects of the researcher and her manuscripts. At least as important, and sometimes even more important, are the intermediate effects of the process of evaluation on the researcher, on the evaluator, and on the future instances of evaluation.

Second, we will pay specific attention to the constructive effects of research evaluation. Evaluation systems inevitably produce quality and relevance as much as they measure it. This holds both for indicator based evaluation and for qualitative peer review evaluation systems. Evaluation systems have these effects because they shape the career paths of researchers and because they form the quality and relevance criteria that researchers entertain. These feedback processes also produce strategic behaviour on the side of the researchers which potentially undermines the validity of the evaluation criteria. We will therefore put central how current and new forms of peer review and indicator systems as main elements of the evaluation process will define different quality and relevance criteria in research assessment, on the short term as well as on the longer term.

Third, we will analyse the diversity of current evaluation practices in a comparative research design. The existing evaluation practices and cultures vary by nation, by institution and by discipline. Although virtually all evaluations aim to ascertain excellence at the international level, how this is operationalized varies greatly. In some cases, citation analysis is very influential, in other cases evaluators and researchers tend to frown upon these quantitative indicators or claim that they are not applicable to their discipline or institution. In some countries, traditional peer review systems are still dominant at the national level, whereas in other countries these criteria have been supplanted with a large set of requirements based on the economic and social effects of research.

This research will be developed in three relatively independent research lines:

- meta-analysis of research evaluations at the level of scientific disciplines and research groups, including our own reports;
- ethnographic and historical studies of the practice of research and research evaluation (for more details see the Scientometrics Chair);
- and comparative studies of the development of scientific and scholarly careers.

The research line on careers will focus on: evaluation and selection processes of scientists, motivation of scientists & academic leadership (see also the Science Policy Studies Chair). To reach the academic top, recruiting and keeping the best staff is crucial, as a critical mass of competent highly skilled people is decisive for excellence. 'Recruiting the best scholars' is increasingly seen as the core business of universities. As the scientific labour market is increasingly global, competition for excellent academic staff is growing. In attracting excellent researchers, the reputation of universities plays an important role, as does universities' Human Resource Management , and the prevalent career system. However, empirical studies about the development of scientific and scholarly careers are hardly done. In this research line we will focus on three important, yet understudied topics, which have implications for the development of the scientific and scholarly careers: evaluation and selection of scientists; motivation of scientists; and academic leadership. In each of these topics, special attention will be given to the academic gender balance. Are there [still any?] gender differences in career support, motivation and network activities of early career scientists? Which gender differences in academic leadership can be distinguished? And what about gender differences in scientific and societal quality of scientists? (Arensbergen, van der Weijden, & Besselaar, 2012)

In the academic world, the evaluation and selection of researchers have a strong effect on the academic careers of particularly early career researchers. Generally, these evaluation processes involves a group activity [e.g. committees]. How committees reach their final evaluation decisions is mainly still a black box. Therefore we will study how group dynamics influence the decision-making of grant proposals and how these decisions are established. This understanding may contribute to an improvement of decision-making in evaluation processes within science.

In the study of scientific motivation, we will focus on the careers of the new generation of researchers: early careers scientists, as for example PhD students and postdoctoral researchers. As the number of PhD students is growing each year, only a small percentage of early career scholars have the opportunity to pursue an scientific career. In our studies we will distinguish between e.g. internal and external motivation, job satisfaction and the academic labour market. We also include the intellectual capital, cultural capital and social capital of young scientists.

The main task of scientific group leaders is to achieve scientific goals and to facilitate intellectual stimulation among the members of the group in which motivation is the keyword. To this end, the group leaders must possess professional competencies and academic leadership skills. Academic leadership is the ability to: (a) organize, acquire and combine resources; (b) manage the research process and direct the researchers and; (c) position the group in their academic and societal environment researchers (Verbree 2011). Currently, the call for academic leadership has become widespread. At the same time leadership is becoming increasingly complex by changes in the science system such as stress on research excellence, the focus on societal relevance of research and the strong competition among researchers and research organisations in order to receive funding are examples. All of these developments intensify the traditional tasks of group leaders and extend these with the need for entrepreneurial activities, which in turn ask for new skills. We will study how the new generation of academic leaders view and manifest academic leadership. Outcomes of this research may be used in training programmes offered to scientists in various career phases, and in career, mobility and evaluation policy developed in academia.

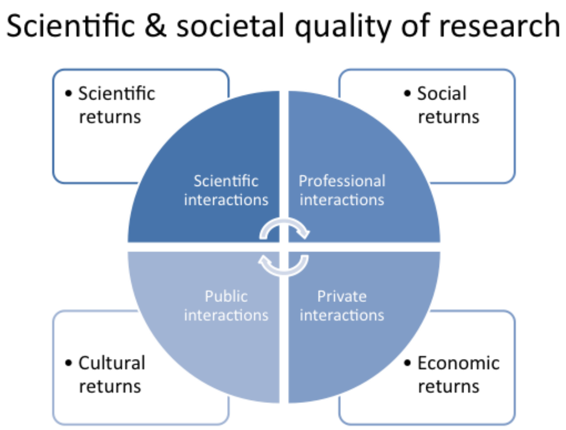
Societal impact of research

Questions regarding the socio-economic and cultural relevance of scientific research have been on the science policy agenda for decades (Bush 1980). The combined processes of globalization and commercialization (Mirowski and Sent 2002; Mirowski 2011) have created a new need for the evaluation of the social, economic, cultural and ecological impact of scientific research. This demands for other evaluation methods that do more justice to the full variety of goals and activities of researchers. This development can be seen in the context of “the new mode of knowledge production” in which the social and economic context of research has been put central (Nowotny, Scott, and Gibbons 2001; Gibbons 1994; Leydesdorff 2000).

From the perspective of the knowledge-based society, policy makers stress the importance of "knowledge valorisation". This term is used for the transfer of knowledge from one party to another with the aim of creating (economic and societal) benefits. However, it is often used in a limited interpretation: only describing the transfer of knowledge to the commercial sector. The value in other domains, for example in professional or public domains, is often not taken into account. The added value of science may be different for different stakeholders. The term valorisation is also often used to describe a one-way-interaction: the dissemination of scientific knowledge to society, while in practice we often observe more mutual, interactive processes.

In this research theme, we will therefore extend the concept of "productive interaction" in analyzing the societal impact of research (Spaapen and van Drooge 2011). This analytical perspective has several advantages. Firstly, it is not limited to economic outcomes, but it also includes social and cultural results. Secondly, it acknowledges that the assimilation of scientific knowledge in society is as much a process as it is an outcome. Thirdly, the term refers to a mutual way of learning.

"Societal quality" is described as the value that is created by connecting research to societal practice and it is based on the notion that knowledge exchange between research and its related professional, public and economic domain strengthens the research involved. This definition encompasses explicitly more than economic value creation only (often referred to as valorisation). It also entails research that connects to societal issues and interactions with users in not-for profit sectors such as health and education etc. as well as to the lay public. The result of analysing and measuring productive interaction (of scientific research) with the non-scientific stakeholders is 'societal quality' (in addition and complementary to the well defined 'scientific quality'). When societal quality entails the valuation of communication of research groups with relevant societal communities (based on productive interaction), the resulting 'value' for the target groups is (shown in the figure):



- Scientists > scientific returns
- Professionals > social returns
- Private sector > economic returns

- (Lay) Public > cultural returns

Productive interaction between scientists and non-scientist stakeholders is thus divided in four broad domains:
- the private sector > including private peers, business community, suppliers);
- the professional sector > including professionals (the 'learned profession'), policymakers, public peers and students;
- the education sector > particularly in applied research contribution to innovation of education is an important objective in itself;
- and the general, lay, public sector.

This conceptual framework coincides with the 4-D model presented by the "Landelijke Commissie Valorisatie", as well as the ERiC (Evaluating Research in Context) model of the Royal Netherlands Academy of Arts and Sciences (KNAW). It also aligns to some extent with the valorisation ranking presented by Scienceworks (Elsevier 2011). Here a distinction is made between Entrepreneurial (private/economic), Collaborating (professional) and Communicative (public).

To analyze these interactions, it is imperative to develop an overview of the type of activities aimed at societal stakeholders. Depending on the scientific discipline, activities may range from organising workshops, performing contract research to writing policy papers or taking part in a media event. Thus, information will be collected about the type of activities that have been executed to realize the socio-economic objectives, and the way the participants in the research group have interacted with the societal stakeholders.

An important problem to analyze the societal impact or quality of research is lack of data. In contrast with the measurement of impact on the scientific communities, there is a lack of:
- a clear definition of the stakeholders of a research group (as smallest unit for evaluation) and what kind of interactions with these stakeholders can be considered as productive;
- the systematic collection of data to substantiate productive interactions. Many interactions that take place are implicit.
- well-defined methods to retrieve relevant information from the internet.

We will focus on the development of robust data sets regarding productive interactions and societal impact of research, as well as the analysis of these datasets, in the context of specific pioneering projects in which the interaction between research and society can be well defined. This will create the possibility to construct, measure, and test potential indicators of societal impact.

To sum up, we will:
1. Identify stakeholders, their interaction with science, and recognition of the concept by the academic and outside world;
2. Develop an independent method to describe, quantify and visualise the societal quality of research.
3. Validate the results of the independent method, to set weighing standards

for specific science research areas and specific stakeholder domains, including reference to mission and objectives.

This research complements the research in the Chair Science and Innovation Studies (see below). The different methodologies explored at CWTS may moreover lead to synergy in the further development of the research agenda in this area. On the basis of this research, CWTS will contribute to the development of new standards for assessing societal impact and implications of academic research.

Scientometrics as a social science between the disciplines

CWTS has not only applied bibliometrics in the context of research assessment, it has also developed a strong tradition in the analysis of the scientific and technology system. This has resulted in a series of reports of the Dutch Science and Technology Observatorium (NOWT). Moreover, CWTS studies have focused on particular aspects of the scientific system based on the analysis of the Web of Science and patent data. This has contributed to the growth of quantitative science and technology studies as a field (Van Raan 1988; Moed, Glänzel, and Schmoch 2005). At the same time, the quantitative and qualitative tradition in the sociology and history of science and in science and technology studies have grown apart (Martin, Nightingale, & Yegros-Yegros, 2012). Separate conference series have developed with mostly different audiences. The citation networks among the relevant journals confirm the separation between quantitative and qualitative science and technology studies (Van den Besselaar, 2000). In the area of science policy studies, this distinction seems to be less stringent.

Although the idea of recreating a unification of the field seems not realistic, it does make sense to explore more systematically the various uses of scientometric methodologies in the context of qualitative research projects. In particular it seems promising to combine the substantive expertise of researchers in a particular area with a scientometric analysis of the same field. We expect that especially our bibliometric network analysis methodology will be useful for researchers in many scientific and scholarly fields. We will also continue our current research in monitoring the science and technology system.

On the basis of our research programme, we will also start to systematically explore our possibilities to inform all actors in the scientific and scholarly system about new developments in science and scholarship, such as the emergence of new interdisciplinary fields (eg. the neurosocial sciences), science based innovation (eg. patent regimes), e-research and new applications of information and communication technologies (eg. virtual research environments), and new paradigms of scientific publication (eg. open access). In the next two years, this will be done by responding to user driven demand, after which a more systematic research theme will be formulated. We expect that scientometrics will be able to contribute to fundamentally novel information tools for researchers and scholars. Also, we aim to stimulate a research tradition in science & technology studies which combines qualitative and quantitative research methodologies to gain a deeper understanding of the dynamics of science and technology.

Chairs at CWTS

Science and innovation studies

The research in the Science and Innovation Studies Chair consists of a group of interrelated research projects. Three of those projects revolve primarily around quantitative methods and established bibliometric approaches. They build on previous work on the development of science metrics, and related meso- or macro level empirical studies at CWTS. This first set of projects is primarily designed to push the boundaries of existing bibliometric approaches and enrich CWTS in-house information systems. The research issues will focus on major relational features of 'science/innovation ecosystems': research cooperation between public sector and private sector; geographical dimensions of research collaboration partnerships; and interactions between knowledge creation, knowledge utilization and technological innovation.

These three data-driven projects are:

1. "Joint research publications: an evidence base for assessment and benchmarking of university-industry research interactions?"
   The main research questions are the following: to which degree are public-private research co-publications a valid indicator of public-private research collaboration? And to what extent can this information source provide performance indicators across a wide range of universities and other research institutes (e.g. in worldwide university ranking systems)?
2. "Geography of science and public-private research interactions: trends and determinants of localization, regionalization and globalization"
   This research question is related to the issue of how, where and how fast scientific cooperation is 'globalizing' within contemporary knowledge production systems (measured in terms of geographical distances between collaboration researchers and research organisations). To what extent are the characteristics of research areas and collaborative relationships (networks) play a role? And to what extent can information regarding the mobility of individual researchers provide additional information with respect to knowledge flows, and 'brain gains'/ 'brain drains' issues?
3. "From basic science to applications: evolutionary flows and revolutionary changes.
   This project will focus on mapping the knowledge flows between science and technology with a special emphasis on tracking early signal of breakthrough events in knowledge production or knowledge application trajectories. The project involves the creation and maintenance of CWTS databases that link journal publications, conference proceedings papers, and patents.

The second type of project is of a more qualitative 'meta-level' nature and will be based on case studies dealing with measurement and assessment issues within policy initiatives and programmes implemented by (Dutch) government agencies.

4. "Funding schemes, incentive systems and assessment frameworks: to what degree are organisational changes, socio-economic impacts, or policy effects measurable?"

This project will be partially conducted at the Leiden University's Campus The Hague, offering a closer connection to government policy agencies in The Hague engaged in strategic decision-making making and policy assessments. Several cases studies will be conducted to identify and assess the (potential) effects of new or previously existing performance indicators, assessment frameworks and measurement methodologies on regional or national initiatives in higher education systems, or the science and innovation systems.

Science policy studies

The purpose of the CWTS Chair in Science Policy Studies is to help develop a 'science of science policy'. Science policy covers a broad range of topics, including priority setting, human resources policies in research, effectiveness of funding instruments, use of scientific results in government policy, organizational structure of national and international research systems. Science policy research at CWTS focuses on two subjects within this field:

1. The impact of various types of research and development activities on economic growth
2. The career system in research.

These topics fit well with the existing CWTS expertise and they are central issues in science policy, both at the national and at the European level.

Research and development and economic growth

In science policy debates the economic importance of science and technology is often emphasized by science advocates and put forward as an argument to increase public research funding. Sometimes this line of reasoning is accepted (for example President Obama's American Recovery Act spent a lot of extra money on research) but often it is rejected and research spending is seen as a luxury and is cut in recessions. Moreover, there is substantial debate about the economic relevance of various types of research. Often, spending on applied R&D is thought to have a greater impact on the economy than basic research; and targeted basic research, aimed at specific economic sectors, is often given priority over untargeted basic research.

Unfortunately, the scientific literature does not provide much guidance in these debates. There are three main approaches. Firstly, macro level empirical studies try to find correlations between R&D spending and GDP growth, but they are inconclusive because if correlations are found, the direction of causality is unclear. Secondly, micro studies consider a particular discovery and try to quantify its economic significance, or look at the local economic impact of a particular research institution, using indicators such as the value added produced by spin-outs and other science based businesses. These studies, however, have to use many debatable attributions and other assumptions; and it is difficult to obtain a complete picture from them. Thirdly, the theory of economic growth has been developed since the 1960s. According to this theory, capital accumulation based on saving and growth of the labor force cannot explain modern economic growth of per capita income; some 80% of this growth has to be attributed to other factors, usually dubbed 'technological progress'. From the mid-1980s work has been done on 'endogenizing' this technological progress, that is, explaining within growth models how technological progress is generated. The central mechanism is analogous to capital accumulation: technological progress is treated as cumulative growth of a stock of knowledge or 'ideas' which is then considered as a kind of production factor. The effect of R&D spending on economic growth depends on the precise assumptions made about their effects on knowledge accumulation. These effects depend on values of parameters that can neither be measured empirically nor derived from a more explicit theory. For a very specific parameter value there is a well behaved growth path, but if the parameter is only slightly different, growth becomes explosive. Then it can be 'tamed' again, but only by making equally arbitrary assumptions. Consequently, this literature is insufficiently convincing to establish a scientific consensus about the economic relevance of R&D, let alone of the various types of R&D.

Our research in this area concentrates on modifying endogenous growth theory by introducing a much more explicit modeling of the roles of the various types of R&D. We replace the concept of the stock of knowledge by that of a technology function. This function specifies that the total factor productivity at any given time is the result of a process of selection by trial and error from an underlying probability distribution. The distribution used is the type of power law often found in empirical scientometric work. In this formulation, different types of R&D lead to different types of trial and error process. As a consequence, it becomes possible to distinguish the economic effects of pure random trial and error invention, applied R&D, targeted and non-targeted basic research. A seminal paper for a closed economy (e.g. the OECD as a whole) is available (Van Bochove 2012). In 2012 and later years, this model will be expanded to open economies that are embedded in a global research infrastructure. In addition, the possibilities to make quantitative empirical estimates of the rates of return on investments in the various types of R&D will be explored.

Career policy

Human resource management is one of the hottest topics in science policy, as witnessed by government position papers in several countries and the EU, the increasing priority given to person oriented funding instruments (e.g. EURYI and Marie Curie at the European level, and the Veni-Vidi-Vici funding instrument in the Netherlands), as well as editorials in *Science* and *Nature* (Waaijer et al. 2011). Human resource management includes issues such as gender policies, international mobility, PhD supervision systems, research worker immigration policies, researcher centered funding arrangements, tenure track systems, tenure arrangements, and mandatory retirement schemes. Much research has been done on many of these topics, but one issue has largely remained below the science policy research radar: the career system as such. Though there are many international differences in career systems, they have a common structure that is taken for granted as though it were a law of nature. This common structure is the succession of positions in the order PhD → post docs → junior research staff → senior research staff → professor. This is accompanied by a gradual increase in income, job security, and research independence. The time it takes to obtain full job security and independence seems to have risen substantially over the past half century, in many cases to an age of 40 years (as is evident from data on such indicators as the age on which the NIH R01 grants are obtained).

Our research will focus on three aspects of international career systems. Firstly, we currently lack accurate knowledge about the details of the career system in the most important countries: exactly when it was introduced in various countries, what are the relative salaries in each phase and what are their the international differences; what is the time spent in each phase (trends and international differences); what are the selection rates (which part of PhD's becomes post docs, etc.), what is the percentage of non-natives of any country in each phase, etc. Secondly, though there is an extensive literature on the best way to supervise PhDs, there is little literature on the rationale of the system as such, such as the length of the PhD trajectory, the need for temporary positions, etc. The current system has been taken for granted too easily. Thirdly, we will study the consequences of the current career system. Do they lead to a competitive disadvantage for the research sector in the labor market? And if so, can the shortfall of students in research intensive, mathematically inclined, disciplines be explained by this competitive disadvantage? Is the relatively low share of females in senior research jobs a consequence of the career system? How do the long selection period and the long time to research independence affect the quantitative and qualitative output of the total research system? How well do researcher-centered funding arrangements work in rectifying some of the problems of the system?

Scientometrics

The Chair Scientometrics will bring together several themes in the CWTS research programme with a focus on the further development of the theoretical understanding of the role and construction of science and technology indicators. The main focus will be on:

- citation and communication theories and models (see the module Communication and citation theories)
- the implications of the citation culture.

Implications of the citation culture

Understanding the citation culture focuses on the implications of evaluation processes for knowledge creation. As indicated above, the last couple of decades have seen an unprecedented growth of evaluation institutions and procedures in scientific and scholarly research. Bibliometrics has played an important role in the stimulation of the use of performance indicators and the current demand for research assessment reports can indeed be seen as a measure of success for the field. At the same time, this creates a new challenge: to counter abuse and misunderstanding of indicators by providing professional information to indicator users, and to process the various modes of critique and criticism of bibliometric and scientometric indicators in a creative way. This challenge is intimately related to the foundational debate about the meaning of the citation, a debate which has been conducted from the very beginning of citation analysis. The question of the meaning of the citation was already tackled by students of Robert Merton, and although many approaches have since than been taken, with varying results, the question of the meaning of the citation is still relevant (Wouters 1999; Nicolaisen 2007; Holton 1978; Elkana et al. 1978; de Solla Price 1978; Zuckerman 1996; Stephen Cole 1992; Cole & Cole 1973).

The goal of the research in this Chair is to open a new perspective on indicators (complementary to the other ways CWTS addresses these debates) by understanding the role of performance indicators and research assessment procedures in the primary process of knowledge creation. In short: how does evaluation influence knowledge production?

Although scholars have made isolated attempts to answer this question, firm theoretical and empirical research on this question is still lacking. A systematic study of the role of indicators is urgent given the increased impact of indicators on research assessment. This research theme will produce useful feedback for the way indicators should and should not be used in research assessments (for example, normative conclusions on empirical basis can be included in the CWTS manual). It will also tie in with and contribute to studies on the meaning of the citation and citation analysis (cf. Hicks & Potter 1991; Leydersdorf 1987; MacRoberts & MacRoberts 1989, 2010; Woolgar 1991; Wouters 1999), and the body of work in sociology of science on relationships between (science) policy, measurement, and knowledge production (cf. Butler 2007; Colwell et al. 2012; Gläser et al. 2002; Martin 2011; Porter 1995; Ritzer; Stone 2001; Verran 2001; Weingart 2005; Whitley et al. 2010). And because formal indicators are also increasingly used in other societal sectors, it is now possible to set up this theme in a comparative research design and draw upon the theories developed in the sociology of evaluation processes (e.g. in organizational sociology, science and technology studies, medical informatics, business studies, history of science). Although a number of other societal sectors could be chosen for comparative research, the role of evaluation and indicators in health care seems a very good candidate. Health care overlaps with research in the area of biomedical research providing the opportunity to better understand the generic roles of performance indicators in professional work (care work compared to research work). Both can be studied in the same overarching institutional framework (e.g. in an academic hospital). Thereby we will also be able to tease out the impact of institutional structures on evaluation processes (by comparing different hospitals).

The main research question will be split in a number of sub-questions, which will be operationalized through a combination of qualitative and quantitative methods. First, we will study how evaluation processes are experienced and treated by principal investigators in biomedical research, at the level of research practice. Second, we will probe the coping strategies of researchers (both PIs and other researchers). To what extent are their strategies aimed at directly influencing the performance measures? How do evaluators respond to these strategies? In what ways does evaluation shape the communication among researchers? Third, we aim to study the dynamics of evaluation procedures as they actually happen in a large research institute. Do quantitative performance indicators have unintended consequences? If so, which different options do research managers have to deal with these effects? And how does all this affect daily life in research?

We will answer this question by studying the actual use of scientometric indicators and other types of performance measures in one or more large research institutes. The studies will be based on systematic participant observation, interviews, an examination of relevant archival material, the study of the management structure, policy documents and procedures in the research institute, and social network analysis of the research networks, supported by quantitative and statistical studies of perceived relationships. Rather than treating scientometric indicators as *a priori* and independent of practices of knowledge production, an ethnographic approach allows for a detailed, bottom-up scrutiny of what these indicators can be constitutive of in terms of performance measurement, standards for evaluation processes, construction knowledge and of authority, etcetera. Using case studies pushes for explanations of the diversity and nondeterministic character of scientific and technological knowledge production (Beaulieu, Scharnhorst, Wouters 2007). The study will be organized in a project approach, where each project builds on the results of earlier projects.

Bibliography

ACLS. 2006. Our Cultural Commonwealth. The report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences. New York: ACLS.

Arensbergen, P., van der Weijden, I., & Besselaar, P. (2012). Gender differences in scientific productivity: a persisting phenomenon? Scientometrics. doi:10.1007/s11192-012-0712-y

Bazerman, Charles. 1988. Shaping written knowledge: The genre and activity of the experimental article in science. Madison: University of Wisconsin Press.

Beaulieu, A., Scharnhorst, A. and Wouters, P. (2007). Not another case study. A Middle-Range Interrogation of Ethnographic Case Studies in the Exploration of E-science. Science, Technology, and Human Values 32(6): 672-692.

Beniger, James R. 1986. The Control Revolution. Technological and Economic Origins of the Information Society. Cambridge. Massachusets, and London, England: Harvard University Press.

Van den Besselaar, P. (2000). Communication between Science and Technology Studies journals. Scientometrics, 47, 169-193.

Blume, Stuart. 1986. The development of Dutch science policy in international perspective 1965-1985. Den Haag: Ministerie van Onderwijs en Wetenschappen.

———. 1992. Insight and Industry: Technological Change in Medicine. Cambridge, Massachusetts: : MIT Press.

Bochove, Cornelis van. 2012 "Basic Research and Prosperity: Sampling and Selection of Technological Possibilities and of Scientific Hypotheses as an alternative Engine of Endogenous growth, CWTS working Paper Series, CWTS-WP-2012-003, http://hdl.handle.net/1887/18636

Bollen, Johan, de Sompel Van, Aric Hagberg, and Ryan Chute. 2009. "A Principal Component Analysis of 39 Scientific Impact Measures." PLoS ONE 4 (6): e6022.

Borgman, Christine. 2007. Scholarship in the Digital Age. Cambridge: MIT Press.

Börner, Katy. 2010. Atlas of Science: Visualizing What We Know. The MIT Press, October 31.

Bornmann, Lutz. 2011. Scientific Peer Review. In Annual Review of Information Science and Technology, 45:199-245.

Bulger, M., Meyer, E. T., Flor, G. de la, Terras, M., Wyatt, S., Jirotka, M., Eccles, K., et al. (2011). Reinventing research? Information practices in the humanities. Network (pp. 1-83). Retrieved from http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/information-use-case-studies-humanities

Butler, L. (2007). Assessing university research: a plea for a balanced approach. *Science and Public Policy*, *34*(8), 565–574.

Bush, Vannevar. 1980. The Endless Frontier: A Report to the President. New York: Arno Press.

Chubin, Daryl E., and Edward J. Hackett. 1990. Peerless science : peer review and U.S. science policy. Albany, N.Y.: State University of New York Press.

Cole, Stephen, Jonathan R. Cole, and Gary Simon. 1981. "Chance and Consensus in Peer Review." Science 214: 881-886.

Cole, S., and J. Cole. 1973. Social Stratification in Science. Chicago: University of Chicago Press.

Cole, Stephen. 1992. Making science: between nature and society. Harvard University Press.

Collins, Randall. 1998. The Sociology of Philosophies: A Global Theory of Intellectual Change. Belknap Press of Harvard University Press, March 4.

Colwell, R., Blouw, M., Butler, L., Cozzens, S. E., Feller, I., Gingras, Y., Hurtubise, J., et al. (2012). *Informing Research Choices: Indicators and Judgment. The Expert Panel on Science Performance and Research Funding* (p. 142). Ottawa.

Daniel, Hans-Dieter. 1993. Guardians of science : fairness and reliability of peer review. Weinheim ;;New York: VCH.

Dutton, William H., Paul W. Jeffreys, and Ian Goldin. 2010. World Wide Research: Reshaping the Sciences and Humanities. New. The MIT Press, May 21.

van Eck, N.J. and L. Waltman. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84(2):523-538, 2010.

Elkana, Yehuda, Joshua Lederberg, Robert K. Merton, Arnold Thackray, and Harriet Zuckerman. 1978. Toward a metric of science: the advent of science indicators. New York; Chichester etc.: Wiley.

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, 109-123.

Gibbons, M. 1994. The new production of knowledge: The dynamics of science and research in contemporary societies. London: Sage.

Gläser, J., Laudel, G., Hinze, S., & Butler, L. (2002). *Impact of evaluation-based funding on the production of scientific knowledge: What to worry about, and how to find out* (pp. 1–40).

Harzing, Anne-Wil. 2010. The Publish or Perish Book: Your Guide to Effective and Responsible Citation Analysis. Tarma Software Research, September 24.

Hicks, D. and Potter, J. (1991). Sociology of scientific knowledge: A reflexive analysis of science disciplines and disciplining science. Social Studies of Science 21: 459-501.

Holton, Gerald. 1978. Can Science Be Measured? In Toward a metric of science: the advent of science indicators, 53. New York;Chichester etc.: Wiley.

Irwin, Alan. 2007. STS Perspectives on Scientific Governance. In The handbook of science and technology studies, third edition, ed. Edward J. Hackett, Olga Amsterdamska, Michael Lynch, and Judy Wajcman, 583-608; 24. Cambridge, Mass: MIT Press.

KNAW. 2010. Quality assessment in the design and engineering disciplines. Royal Netherlands Academy of Arts and Sciences. http://www.knaw.nl/Pages/DEF/27/160.bGFuZz1FTkc.html.

Knorr-Cetina, Karin. 1999. Epistemic Cultures: How the Sciences Make Knowledge. Cambridge, Massachusetts: Harvard University Press.

Lamont, Prof. MichÃ¨le. 2009. How Professors Think: Inside the Curious World of Academic Judgment. 1st ed. Harvard University Press, March 31.

Lampland, Martha, and Susan Leigh Star. 2009. Standards and Their Stories: How Quantifying, Classifying, and Formalizing Practices Shape Everyday Life. Cornell University Press, January. http://www.amazon.com/gp/redirect.html?location=/exec/obidos/ASIN/0801447178&AWSAccessKeyId=1TQA6S4ATNFN6EYN0H82.

Langfeldt, Liv. 2004. "Expert panels evaluating research: decision-making and sources of bias." Research Evaluation 13 (1) (April): 51-62. doi:Article.

Laudel, Grit. 2002. "What do we measure by co-authorships?" Research Evaluation 11 (1): 3-15.

van Leeuwen TN; Moed HF; Tijssen RJW; et al. (2001). "Language biases in the coverage of the Science Citation Index and its consequences for international comparisons of national research performance" Scientometrics, Vol. 51, Issue: 1, pp. 335-346, DOI: 10.1023/A:1010549719484

Leydesdorff, Loet. 2000. "A Triple Helix of University-Industry-Government Relations." The Journal of Science & Health Policy 1: 43-48.

Leydesdorff, L. (1987). Towards a theory of citation. Scientometrics 12: 287-291.

MacRoberts, M.H. & MacRoberts, B.R. (1989). Problems of citation analysis: A critical review. Journal of the American Society for Information Science 40(5): 342-349.

MacRoberts, M.H. & MacRoberts, B.R. (2010). Problems of citation analysis: A study of uncited and seldom-cited influences. Journal of the American Society for Information Science and Technology 61(1): 1-12.

Martin, B. R. (2011). The Research Excellence Framework and the "impact agenda": are we creating a Frankenstein monster? Research Evaluation, 20(3), 247-254. doi:10.3152/095820211X13118583635693

Martin, B. R., Nightingale, P., & Yegros-Yegros, A. (2012). Science and technology studies: Exploring the knowledge base. *Research Policy*. doi:10.1016/j.respol.2012.03.010

Melin, G, and O Persson. 1996. "Studying research collaboration using co-authorships." Scientometrics 36 (3): 363-377-377.

Mirowski, and Edward J. Sent. 2007. The Commercialization of Science and the Response of STS. In The handbook of science and technology studies, third edition, ed.

Edward J. Hackett, Olga Amsterdamska, Michael Lynch, and Judy Wajcman, 635-689. Cambridge, Mass: MIT Press.
Mirowski, Philip. 2011. Science-Mart: Privatizing American Science. Harvard University Press.
Mirowski, Philip, and Esther-Mirjam Sent. 2002. Science bought and sold. The University of Chicago Press.
Moed, Henk F. 2005. Citation analysis in research evaluation. Vol. 9. Dordrecht: Springer.
Moed, Henk F., Wolfgang Glänzel, and Ulrich Schmoch, eds. 2005. Handbook of Quantitative Science and Technology Research. Dordrecht: Kluwer Academic Publishers. http://www.springerlink.com/content/mp3mj00t3n6ljj2q/.
Moed, H. F., & van Leeuwen, T. N. (1996). Impact factors can mislead. Nature, 381(6579), 186.
Nederhof, A., and A. Van Raan. 1987. "Citation theory and the Ortega hypothesis." Scientometrics 12 (5) (November 26): 325-328. doi:10.1007/BF02016674.
Nicolaisen, Jeppe. 2007. Citation Analysis. In Annual review of information science and technology. Volume 41, 2007, 609-642. New York etc.: Interscience Publishers.
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;
Nowotny, Helga, Peter Scott, and Michael Gibbons. 2001. Re-thinking Science: Knowledge and the Public in an Age of Uncertainty. Cambridge (UK): Polity Press.
Olson, Gary M., Ann Zimmerman, and Nathan Bos, eds. 2008. Scientific Collaboration on the Internet. The MIT Press, October 31.
Olson, Gary, Thomas Finholt, and Stephanie Teasley. 2000. Behavioral Aspects of Collaboratories. In , 1-14. Mahwah, NJ: Lawrence Erlbaum Associates.
Olson, Gary, Judith Olson, Nathan Bos, Ann Zimmerman, and SOC Data Group. 2005. Theory of remote collaboration (TORC). http://colab.cim3.net/file/work/Expedition_Workshop/2005-09-23_Pioneering_Governance_Mechanisms/Olson_Theory_of_Remote_Collaboration_2005_09_23.doc.
Peters, HPF; van Raan, A.F.J. (1994). On determinants of citation scores: A case study in chemical engineering. Journal of the American Society for Information Science 46 (1) 9-21;
de Solla Price, D. 1978. Toward a Model for Science Indicators. In Toward a metric of science: the advent of science indicators. New York; Chichester etc.: Wiley.
Porter, T. (1995). Trust in numbers: The pursuit of objectivity in science and public life. Princeton University Press.
van Raan, Anthony, ed. 1988. Handbook of Quantitative Studies of Science and Technology. Amsterdam: Elsevier Science Publishers.
van Raan, A.F.J. (1990). Fractal dimension of co-citations. Nature 347, 6294, 626;
van Raan, A.F.J. (2000). On growth, ageing, and fractal differentiation of science, Scientometrics 47 (2) 347-362;
van Raan, A.F.J. (2001a). Competition amongst scientists for publication status: Toward a model of scientific publication and citation distributions. Scientometrics 51 (1) 347-357;
van Raan, A.F.J. (2001b). Two-step competition process leads to quasi power-law income distributions: Application to scientific publication and citation distributions. Physica A: Statistical Mechanics and its Applications, 298(3) 530-532.
Rinia EJ; van Leeuwen TN; van Vuren HG; et al. (1998) "Comparative analysis of a set of bibliometric indicators and central peer review criteria - Evaluation of condensed matter physics in the Netherlands". Research Policy. Vol. 27, Issue: 1, pp. 95-107 DOI: 10.1016/S0048-7333(98)00026-2

Ritzer, G. (1993). The McDonaldization of society. An investigation into the changing character of contemporary social life. Sage.
Stone, D. (2001). Policy paradox. The art of political decision making, revised edition. W. W. Norton & Company.
Royal Society. 2011. Knowledge, networks and nations. Royal Society, March 28. http://royalsociety.org/uploadedFiles/Royal_Society_Content/Influencing_Policy/Reports/2011-03-28-Knowledge-networks-nations.pdf.
Shrum, Wesley, Joel Genuth, and Ivan Chompalov. 2007. Structures of Scientific Collaboration. Cambridge: MIT Press.
Slaughter, Sheila, and Larry Leslie. 1997. Academic Capitalism: Politics, policies and the entrepreneurial university. Johns Hopkins UP.
Sonnenwald, D. H. 2007. "Scientific collaboration." Annual Review of Information Science and Technology 41: 643-681.
Spaapen, Jack, and Leonie van Drooge. 2011. "Introducing 'productive interactions' in social impact assessment." Research Evaluation 20 (3): 211-218. doi:Article.
Suresh, S. 2011. "Moving Toward Global Science." Science 333 (6044) (August): 802-802. doi:10.1126/science.1210025.
The Virtual Knowledge Studio:, Paul Wouters, Katie Vann, Andrea Scharnhorst, Matt Ratto, Iina Hellsten, Jenny Fry, and Anne Beaulieu. 2008. Messy Shapes of Knowledge-STS Explores Informatization, New Media, and Academic Work. In The handbook of science and technology studies, third edition, 3:319-351; 14. Cambridge, Mass: MIT Press.
Thelwall, Mike. 2005. Link Analysis: An Information Science Approach. San Diego: Academic Press.
Tijssen RJW; Visser MS; van Leeuwen TN (2002), "Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference?", Scientometrics, Vol. 54, Issue: 3, pp 381-397. DOI: 10.1023/A:1016082432660
Verbree, M. (2011). Dynamics of academic leadership in research groups. Den Haag, Rathenau Instituut.
Verran, H. (2001). Science and African Logic. University of Chicago Press.
Waaijer, Cathelijn, Cornelis van Bochove and Nees Jan van Eck, 2011, "On the map: Nature and Science editorials", Scientometrics, Volume 86, Number 1, 99-112
Wagner, Louk de la Rive. 2006. International collaboration in science and technology: promises and pitfalls. In Science and Technology Policy for Development: Dialogues at the Interface, ed. Louk de la Rive Box and Rutger Engelhard, 165-176. 1st ed. Anthem Press, May 1.
Waltman, L. and N.J. van Eck. The inconsistency of the h-index. Journal of the American Society for Information Science and Technology, 63(2):406-415, 2012.
Waltman, L. and N.J. van Eck. A new methodology for constructing a publication-level classification system of science. Journal of the American Society for Information Science and Technology, in press.
Waltman, L., N.J. van Eck, T.N. van Leeuwen, M.S. Visser, and A.F.J. van Raan. Towards a new crown indicator: Some theoretical considerations. Journal of Informetrics, 5(1):37-47, 2011.
Waltman, Ludo, Robert J. W Tijssen, and Nees Jan van Eck. 2011. "Globalisation of science in kilometres." 1103.3648 (March 18). http://arxiv.org/abs/1103.3648.
Weingart, Peter. 2005. "Impact of bibliometrics upon the science system: Inadvertent consequences?" Scientometrics 62 (1) (January 1): 117-131. doi:10.1007/s11192-005-0007-7.
Whitley, Richard, Gläser, J., & Engwall, L. (Eds.). (2010). *Reconfiguring knowledge production : changing authority relationships in the sciences and their consequences for intellectual innovation*. Oxford: Oxford University Press.

Woolgar, Steve. 2002. The Boundaries Of Accountability: A Technographic Perspective. In Responsibility under Uncertainty: Science, Technology and Accountability. EASST 2002. Vol. York.

Woolgar, S. (1991). Beyond the citation debate: towards a sociology of measurement technologies and their use in science policy. Science and Public Policy 18: 319-326.

Wouters, P. 1999. "Beyond the Holy Grail: From citation theory to indicator theories." Scientometrics 44 (3): 561-580.

Wouters, P., Beaulieu, A., Scharnhorst, A., & Wyatt, S. (2013) (eds.). Virtual Knowledge and the Social Sciences - Experimenting in the Humanities. MIT Press.

Wouters, P., & Costas, R. (2012). Users , narcissism and control – tracking the impact of scholarly publications in the 21 century (pp. 1-50). Utrecht.

Zuckerman, Harriet. 1996. The Scientific Elite: Nobel Laureates in the United States. New Brunswick and London: Transaction Press.