ISSI Newsletter

quarterly e-newsletter of the International Society for Scientometrics and Informetrics

ISSN 1998-5460

#21 / volume 06 number 1

March 2010

CONTENTS

EDITORIAL

NEW VOLUME, NEW LAYOUT

Twenty ISSI Newsletter issues have been presented in the past five years. Which is, on the one hand, an impressive figure. On the other hand, *twenty*, as a Hungarian idiom says, *is way too many, even of* [such a delicacy as] *dumplings filled with plum*. Empirical studies has shown that the above phrase is undoubtedly true: more than 5 dumplings is indeed a tremendous (though not impossible) amount to consume in one go. On the basis of these strictly scientific tests, and taken into account the Newsletter's similarly delicious nature, we came to the conclusion that it was high time to think about something new. As the content has stood the test of time, it was the layout that provided us the most obvious playground for changes. And voilà! You are keeping the very first Newsletter issue with a new, improved design in your hands (or if you are environmentally conscious, on your screen). The changes incorporate a new body font easier to read, together with a little bit more conservative layout that is supposed to reflect the Newsletter's grown-up status as well as its serious intentions and ambitions. We hope you'll enjoy it and keep on reading the Newsletter in the future, too.

Balázs Schlemmer, technical editor

ISSI e-Newsletter (ISSN 1998-5460) is published by ISSI (http://www.issi-society.info/).
Contributors to the newsletter should contact the editorial board by e-mail.

- **Wolfgang Glänzel,** *Editor-in-Chief:* wolfgang.glanzel[at]econ.kuleuven.be
- **Balázs Schlemmer,** *Technical Editor*: balazs.schlemmer[at]gmail.com

PAVING THE ROAD FOR SCIENTOMETRIC EXPERTISE IN EUROPE:

THE EUROPEAN SUMMER SCHOOL FOR SCIENTOMETRICS (ESSS)

esss european summer school for scientometrics

1. IMPORTANCE OF BIBLIOMETRICS AND SCIENTOMETRICS:

Bibliometrics and scientometrics are beyond the shadow of a doubt important for many stakeholders.

Philosophers of science increasingly use scientometric procedures to analyse developments and trends in science and technology.

Scientists find themselves caught in the journal impact factor dominated "publish or perish" game, and librarians at least partly base their collection management decisions on bibliometric indicators.

In addition to the peer review system, bibliometrics and scientometrics are now heavily used for research assessment. Science and technology are strongly based on competitiveness and innovation. To survey and assess them, excellence clusters are required to highlight the leading position of departments, universities or countries, to manage better investments and to transfer inventions into innovation. Emerging technologies have to be identified to ensure and increase international competitiveness.

Government offices, university directors and research funders need to decide which fields of research and institutions are to be supported or built up, and how research budget and grants are to be allocated. Such decisions are increasingly based on or at least influenced by bibliometrics and scientometrics.

2. LACK OF AVAILABLE SCIENTOMETRICS EXPERT KNOWLEDGE

Data handling, indicator construction and interpretation require competent expert knowledge, which is currently only avail-

able to a limited extent for all stakeholders in Central Europe (especially in German speaking countries) not the least due to lacking training opportunities.

This is probably most evident for research quality managers. A majority of them have never had any bibliometric or scientometric training at all, however, they manipulate and interpret data and use bibliometric databases and tools on a daily basis. Their taken decisions often have severe implications for individual researchers as well as for whole institutions.

3. THE FORMATION OF ESSS AS CONSEQUENT RESPONSE

Responding to the lack of a pertinent scientometrics education (especially in German speaking countries) and to the increasing demand (particularly of research quality managers), the University of Vienna (Austria), the Humboldt University of Berlin (Germany), the Institute for Research Information and Quality Assurance – iFQ - (Germany) and the Katholieke Universiteit Leuven (Belgium) joined cooperatively to establish the European Summer School for Scientometrics (esss) in 2010.

4. ESSS MISSION

esss offers training covering major aspects of quantitative analysis of science and technology. Its courses are especially designed to cater for the needs of science policy makers, research quality managers, scientists, information specialists and librarians. ESSS is open to interested parties from all over Europe and beyond.

Attendees can expect a sound overview of state-of-the-art scientometric methods and the opportunity to familiarize themselves with the most commonly used data bases, to learn how to construct relevant indicators and how to interpret the data.

Theoretically imparted knowledge will be consolidated in hands-on trainings whenever suitable in order to guarantee a sustainable learning experience.

Participants will challenge themselves academically, gain crucial experience, ad-

Archway of the main building, University of Vienna / Photo: University of Vienna

esss website: http://www.scientometrics-school.eu/

vance their careers and experience knowledge sharing and exchange of ideas with esss staff, lecturers and other participants.

esss has an international orientation using English as language of instruction, and its venue will rotate among the organizing institutions annually.

Cooperations with other institutions are welcome in order to bring esss to other stimulating and exciting locations all over Europe.

5. ESSS SCHEDULE 2010

The inauguration of ESSS will take place at Humboldt University, Berlin, Germany on June 16th, 2010.

The Humboldt-University is located right in the very centre of Berlin close to the city's most well-known places of interest. The famous Brandenburger Tor, the Reichstag, the chancellery and the Museumsinsel (Museum Island) are within walking distance.

The first ESSS course will take place in conjunction with this year's iFQ annual conference on research evaluation, which will be held from June 14th-15th at the Wissenschaftszentrum Berlin (WZB).

In 2010 a three-day programme (16–18 June) will be offered as a premium foretaste of what attendees can expect from 2011 onwards.

The first day of the ESSS, June 16th, will feature keynotes and presentations by internationally recognised scientometrics experts. On day 2 and 3 we will present the first two interactive modules: module one will focus on journal impact measures and composite indicators while module two will be concerned with Hirsch-type indices.

On both days lectures introducing the theoretical background will be combined with practical exercises. The lectures will be delivered by experts from the three organizing institutions, and internationally recognised scientometricians will be invited as trainers to reinforce the lecturers.

6. ESSS SCHEDULE 2011 ONWARDS

In the following years the coursework will be extended to cover five modules. Each

module will be tailored to the needs of a predefined audience. Modules can then be attended independently from other modules to allow for maximum flexibility of esss participants.

esss will be integrated into the postgraduate course offerings of the University of Vienna and aims to award attendees with ECTS points for successfully completed course modules.

The University of Vienna will also be the host for esss in 2011.

7. ESSS ORGANIZATION

All organizing institutions are represented in the steering committee which will be responsible for the esss activities.
The esss steering committee members are

- Wolfgang Glänzel and Koenraad Debackere (K.U. Leuven),
- Stefan Hornbostel and Sybille Hinze (iFQ and Humboldt Universität zu Berlin),
- Juan Gorraiz and Christian Gumpenberger (University of Vienna).

The annual organization of esss will be coordinated by a local and alternating Programme Committee.

In 2010 the esss Steering Committee will also serve the Programme Committee's tasks.

8. ESSS WEBSITE

Please also visit our esss website at http://www.scientometrics-school.eu. Once registration is open further details will be communicated.

Wolfgang Glänzel

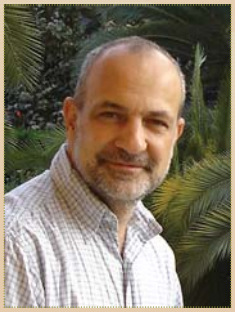
Koenraad Debackere

(Katholieke Universiteit Leuven)

Stefan Hornbostel

Sybille Hinze

(iFQ and Humboldt Universität zu Berlin)



Juan Gorraiz

Christian Gumpenberger

(University of Vienna)

The esss steering committee

We are looking forward to seeing you in Berlin!

THE UNIVERSITY LIBRARY AT THE UNIVERSITY OF BERGEN ANNOUNCES THE

15TH NORDIC WORKSHOP ON BIBLIOMETRICS AND RESEARCH POLICY.

28-29 SEPTEMBER 2010, THE UNIVERSITY OF BERGEN, NORWAY

Bibliometric researchers in the Nordic countries have arranged annual Nordic workshops on bibliometrics since 1996. The general idea of the workshop is to present recent bibliometric research in the Nordic countries and to create better linkages between bibliometric research groups and their PhD students.

The workshop language is English and the workshop is open to participants from any nation. The workshop is also open to participants who wish to take part without presenting.

More information will soon be available at the workshop website (www.uib.no/ub).

Deadline for registration, presentations and abstract submission: August 15th, 2010

Further questions can be addressed to the workshop coordinator:
Dag W. Aksnes, dag.aksnes[at]ub.uib.no

Hope to see you in Bergen in September.

Dag W. Aksnes
University Library Bergen / NIFU STEP – Norwegian Institute for Studies in Innovation, Research and Education

The view of Bergen, Norway. Both photos were taken from the Ulriken mountain resturant, where the conference dinner will take place. (Photos: courtesy of S.M. Tunli – tunliweb.no)

PHD COURSE IN INFORMETRICS AND RELATED QUANTITATIVE RESEARCH METHODS

UMEÅ UNIVERSITY, 14-18 JUNE, 2010

PURPOSE OF THE COURSE

Over the last decade or so, informetrics has become a hot topic. It is an interdisciplinary research field that attracts researchers and students from a number of related research areas, e.g. sociology of science, history and philosophy of science, library and information science, knowledge management, research policy, regional economics. Informetrics has two strong traditions: bibliometrics and webometrics. During the last years there has been a strong development of research techniques as well as theoretical models of communication processes and networks. There are also connections to areas of application. Search engine technology is today highly dependent on informetric theories. The focus on applying informetric indicators for research evaluation and allocation of research funds has never been stronger, especially in the public sphere, where heated debate takes place between the evaluated researcher, administrators, politicians and informetricians.

The course will present various perspectives on informetrics as research practice, in terms of its application for research evaluation and for mapping research fields, as well as the relation between informetric analyses and theories on the social and intellectual organization of research fields. Further, the course will focus on a fundamental theoretical issue in informetrics, indicators of citation. The course will present and discuss several pertinent quantitative research ımethods, for example social network analysis and statistical modeling.

In addition to this, doctoral research courses and workshops have proven to be excellent means for sharing the Nordic mentoring expertise and creating contacts between doctoral students and senior researchers. Thus, an important goal of the research course is to provide the doctoral students with a forum in which to present their research projects and to discuss informetric research issues with senior researchers and fellow students.

The research course will contain:

- A few general lectures given by senior researchers on
  - → introducing aspects of informetrics research
  - → presenting advanced informetric issues by examples from research
- Specific tutorials on methods and tools where students work actively in the lab with informetric problems
- Group discussions focusing on central methodical issues
- Presentations of research projects by the research students focusing on methodical issues

CONTENT

- Informetric theories, models and indicators
- Central issues in quantitative research evaluation
- Science mapping, techniques, problems and evaluation
- Webometrics and beyond
- Informetric research in Information Retrieval
- Informetric research in and Science studies
- Statistical modeling: building and testing models
- Large scale network analysis
- Introductions to relevant research tools

LECTURERS

- Prof. Mike Thelwall (Wolverhampton)
- Prof. Olle Persson (Umeå)
- Assoc. Prof. Jesper W Schneider (Aalborg)
- Asst. Prof. Rickard Danell (Umeå)

More lecturers may be included at a later stage. The lecturers involved have recent experiences of informetric research. A required reading list will be sent to participants with the note of acceptance.

LOCATION & OTHER DETAILS

Location: Dept of Sociology, Umeå University, Sweden (www.umu.se/soc)
Time: 14-18 June, 2010
Accommodation: To be announced
Application deadline: 07 May, 2010

CREDITS

We recommend 5 ECTS (European Credit Transfer System) for this course. This 5 ECTS (European Credit Transfer System) course corresponds to 5 weeks of study. It includes teaching, seminars and discussions concentrated to one intensive course meeting for 5 days. This means that students need to work on the course both in advance of and after the course meeting. They are asked to study the suggested literature and to prepare a presentation in advance of the course. In order for the students to obtain 5 ECTS they are supposed to write a conference-like paper 1-2 month later, to be send to the course management, evaluated, and to be accepted by the course management (with the course teachers as reviewers). If they do not do that or the submission is really unacceptable, we suggest that they can obtain a reduced number of ECTS. The main responsible person for the course will sign a certificate per participant stating that they actually participated and also give a recommendation about the number of ECTS that the student has achieved. Of course, it is up to the responsible persons of the local PhD programme to finally decide how credits should be assigned.

TARGET AUDIENCE

The course is intended for all PhD students with an interest in informetric research methods. Provided there is room for more participants, the course will also be open for other researchers.

HOW TO APPLY

Research students are required to submit a five-page abstract (approx. 2500 words) of their PhD research topic, including research questions and the methodical issues of the research project. The application should also include a brief CV. Please, send the application as word or PDF file by e-mail to Jesper W. Schneider (jws[at]db.dk)

For NorsLIS doctoral students: The application must contain the following information:

- Name of workshop
- Last name
- First name
- Birth date
- Male/female
- E-mail address
- Name of supervisor
- Supervisor's e-mail address

NEW JOURNAL IMPACT INDICATORS TAKE REFERENCES INTO ACCOUNT:

A COMPARISON

RAF GUNS
University of Antwerp, IBW, Antwerpen, Belgium

RONALD ROUSSEAU
University of Antwerp, IBW, Antwerpen, Belgium;
KHBO, Industrial Sciences and Technology, Oostende, Belgium;
K.U.Leuven, Dept. Mathematics, Leuven, Belgium

ABSTRACT: The definitions of the reference return ratio, the audience factor and the source normalized impact per paper (SNIP) are explained, without going too much in mathematical detail. Comparisons between the three are made. It is observed that the main reason for their introduction is to correct for citation behaviour, and this on a classification-free basis.

INTRODUCTION

Recently a number of new journal impact indicators have been proposed: Zitt and Small's *audience factor* and Moed's *SNIP* (Zitt & Small, 2008; Moed, 2010). To these two we add Nicolaisen and Frandsen's *reference return ratio* (3R) (Nicolaisen & Frandsen, 2008). These indicators have in common that they are based on or are similar to the classic impact factor and that numbers of references play a role in their definition. In this contribution we will compare the structure of the mathematical formulae used to define these indicators. We will further discuss the reasons why they are introduced.

In order to unify the notations and fix the ideas, we will consider a citation period of one year, the year Y, a publication period of three years, namely the years [Y-3, Y-1] and a certain reference period which we do not have to specify for the moment. Journal impact indicators will be calculated for the journal J_o in the year Y. We further assume that for all purposes (counting publications, references, citations) the same type of documents, referred to as 'articles' will be used. We first recall the definition of the standard synchronous impact factor. The

standard 3-year impact factor for journal J_0 in year Y, denoted as $IF_3(J_0,Y)$, is defined as (Garfield, 2006; Rousseau, 1988):

$$IF_3(J_0,Y) = \frac{Cit(J_0,Y)}{Pub(J_0,Y)}$$

where $Pub(J_0,Y)$ denotes the number of articles published by J_0 during the corresponding publication period [Y-3, Y-1] and $Cit(J_0,Y)$ denotes the number of citations received by journal J_0 in the year Y to articles published during the period [Y-3, Y-1]. As J_0, Y and the publication period [Y-3, Y-1] are fixed, we will simply write:

$$IF_3 = \frac{Cit}{Pub}$$

This notation will be used when we introduce the new journal impact indicators.

NEW JOURNAL IMPACT FACTORS TAKING REFERENCES INTO ACCOUNT

In this section we introduce the audience factor, the source normalized impact per paper and the reference return ratio.

1. THE AUDIENCE FACTOR (ZITT & SMALL, 2008)

The audience factor has been introduced in (Zitt & Small, 2008). We adapt the original definition to our notation and to the specific periods used here as an example.

The audience factor of journal J_0 in the year Y, denoted as AF, is defined as

$$AF = \frac{\sum_J w_J \cdot Cit_{J,J_o}}{Pub}$$

where Cit_{J,J_0} denotes the number of citations received by journal J_0 from journal J (during the citation period, here the year Y) to articles published in J_0 during the publication period. The factor w_J is a weighting factor defined as $m_S / m_T(J)$, where $m_T(J)$ is the average number of references in articles published in journal J during the reference period and m_S is the average number of references in a reference set of articles, also published during the reference period. Note that Zitt and Small take the reference period equal to the publication period. As reference set they use all articles published in journals covered by the used database or all articles published in journals belonging to the same JCR category as J_0. Other reference sets are of course feasible. In general, the factor w_J gives greater weight to citations from journals that have short reference lists.

The formula for the audience factor can also be written as

$$AF = m_S \frac{\sum_J \frac{Cit_{J,J_0}}{m_T(J)}}{Pub}$$

This illustrates more clearly that AF not only depends on the number of references in journals citing J_0, but also on the average number of references in the database as a whole.

2. MOED'S SOURCE NORMALIZED IMPACT PER PAPER: SNIP (MOED, 2010)

Recently, Henk Moed proposed another indicator, called the Source Normalized Impact per Paper or SNIP. An important parameter leading to the definition of SNIP is the citation potential. It is defined using the notion of a journal's field. Journal J's field is the set of articles published in the year Y that cite at least one article published in journal J during the reference period. The average number of references in the articles belonging to the journal's field is then called its citation potential. The formal definition is as follows. Let $\{a_J^1, ..., a_J^m\}$ be the set of m articles belonging to J's field and let r_J^k be the number of references of article a_J^k then the citation potential of journal J, denoted as R_J, is defined as:

$$R_J = \frac{\sum_{k=1}^{m} r_J^k}{m}$$

Moed refines the citation potential by taking only references into account if the cor-

responding item belongs to the database used. This refined citation potential is denoted as R_J^{db}. The raw impact of journal J in year Y is then defined as its 3-year synchronous impact factor (denoted by Moed as $RIP_J(Y) = IF_3$). Finally, the SNIP is a normalized version of this impact factor. Normalization is performed as follows. The 3-year impact factor is divided by the ratio of the journal's citation potential and the median value, MED, of all citation potentials of all journals in the database. Hence journal J_0's SNIP in the year Y is:

$$SNIP_{J_0}^{Y} = \frac{RIP_{J_0}(Y)}{R_{J_0}^{db} / MED}$$

where MED denotes the median value of $\left\{ R_{J_k}^{db} ; k = 1, ..., N \right\}$, N being the number of journals in the database. The ratio

$$\frac{R_J^{db}}{MED}$$

is a relative citation potential (taking the used database into account).

Concretely, Moed proposes a reference period of 10 years, but observes that values of SNIP are only weakly dependent on the length of this period.

3. THE REFERENCE RETURN RATIO (NICOLAISEN & FRANDSEN, 2008)

In 2008 Nicolaisen and Frandsen introduced another variation on the impact factor idea, taking references into account as did Zitt, Small and Moed. Their impact indicator is called the Reference Return Ratio, or the 3R in short. This indicator is defined as:

$$3R = \frac{Cit}{Ref}$$

where *Ref* denotes the total number of references in journal J_0 of articles published during the publication period referring to articles published (in any journal) during the reference period.

Concretely, Nicolaisen and Frandsen allow more flexibility by the use of the following six parameters:

- n_p denoting the length of the publication period (in our example this is 3);
- n_c enoting the length of the citation period (in our example this is 1);
- n_r denoting the length of the reference period;
- Y_p denoting the first (oldest) year of the publication period (in our example this is Y-3);
- Y_c enoting the first (oldest) year in the citation period (in our example this is Y);
- Y_r denoting the first (oldest) year in the reference period.

We observe that the authors allow for a citation period which is longer than one year, in line with the generalized impact factors as discussed in (Frandsen & Rousseau, 2005). In principle, the same parameters, allowing for different publication, citation and reference periods, could also be used for the other two indicators.

COMMENTS

The inverse of the audience factor's weighting factor $1 / w_J = m_T(J) / m_S$ can be considered a kind of relative citation potential playing a somewhat similar role as Moed's R_J^{db} / MED. One difference is that the audience factor's w_J directly influences the weight of each citing journal, whereas SNIP's relative database citation potential gives equal weight to each citing article or journal. Furthermore, while Zitt and Small use an average in the denominator, Moed uses a median, which ensures that for one half of the journals in the database SNIP is smaller than the impact factor and for the other half the impact factor is smaller than SNIP.

Zitt, Small and Moed mention explicitly that their approach uses normalization on the citing side as opposed to normalization "on the cited side". The difference between the two types of normalizations is that in normalization "on the cited side" one compares with the number of citations received by comparable journals, e.g. journals be-

longing to the same JCR category. In normalization "on the citing side" one considers a citation potential by taking the number of references into account. This distinction was already mentioned in (Zitt et al., 2005). In both cases a "neighbourhood" in the journal citation network plays a role. Such a neighbourhood consists of journals that cite or are cited by the journal under study. Recently Zitt announced a new version of the audience factor (referred to as v0.2) which makes this idea of a "neighbourhood" more concrete (Zitt, 2010). The reference return ratio is different in this respect: in the calculation of 3R, one does not directly take characteristics of the citing side into account, only of the referencing behaviour of the journal under study.

By refining the citation potential to only those references that are in the database, Moed explicitly takes the database content into account. The used database is indeed an important element in calculating actual values of journal impact indicators. Recall that this aspect is an essential element in the "conglomerate" notion introduced in (Rousseau, 2005), where e.g. citations are explicitly taken from a pool (a database or part thereof).

ADVANTAGES OF THESE NEW JOURNAL INDICATORS

These recent additions to the set of journal indicators do have quite some advantages compared to the basic impact factor definition. These advantages are discussed here, based on the introducers' own lists.

The three proposals take the citation environment into account. Indeed, the audience factor is defined exactly as the standard impact factor, except for the fact that a weighting factor for received citations is applied. We recall that this weighting factor is the ratio $m_S / m_T(J)$, which gives greater weight to those journals that have on average shorter reference lists. This is the 'audience' element of the proposal. If one takes all weights equal to one, one gets the standard 3- (or n)-year synchronous impact factor. SNIP divides the impact factor by the journal's citation potential, thus also accounting for the audience. Finally, 3R takes only part of the citation environment (the journal's self-citations) into account.

Moed's SNIP is the standard 3-year synchronous impact factor adapted to the content of the database (and hence less biased against the humanities, as the humanities are generally less covered by the international databases).

In principle these proposals do not depend on an external classification scheme such as the JCR categories. Of course one may choose such a category as the reference set of journals for the audience factor, but this is a choice made by the implementer and not an element in the definition.

The audience factor as well as SNIP correct for differences in propensity to cite and field-dependent response (speed).

An important difference between these proposals is the fact that Moed defines the citation neighbourhood (a journal's field) based on articles; the audience factor uses journals; and 3R uses the journal under study to determine the citation neighbourhood. This aspect makes SNIP better adapted to incorporating multidisciplinary journals, such as *Nature* or *Science*.

3R seems to be best adapted to correct for the special characteristics of review journals. Indeed, such journals publish articles with long reference lists and hence their denominator is larger than that of ordinary journals, leading to a reduction of their 3R value.

DISADVANTAGES AND PROBLEMS

Since all these new indicators are in some way based on the standard impact factor, they inherit some well-known problems of the impact factor and of journal evaluation by citation analysis in general.

All three proposals include journal self-citations, but of course, self-citations can be removed in the same way as they are sometimes removed for the classical impact factor.

Moed, Zitt and Small note that their proposals (SNIP and the audience factor) do not distinguish between citing journals: it does not make a difference if a citation comes from a top journal (such as *Nature* or *Science*) or from a "lesser" journal. This observation also holds for the reference return ratio. Impact measures based on Pinski-Narin-type influence weights such as the Eigenfactor score, or the SCImago Journal Rank (Pinski & Narin, 1976; Bollen et al., 2006; Bergstrom, 2007, Grupo Scimago, 2007) do take this aspect into account. Recently, Waltman & Van Eck (2010) showed that indicators such as the audience factor are insensitive to field differences, whereas Pinski-Narin-type influence indicators are insensitive to insignificant journals. They also showed that the Article Influence Score (based on the Eigenfactor score) represents a trade-off between the two insensitivity properties: by varying a parameter one can give more or less weight to one of either properties. We add that, as far as we understand it, Pinski-Narin type approaches tend to enhance the Matthew Effect for journals.

The authors of these new measures also note that their indicators do not control for imbalance in citation transactions (exporters vs. importers of knowledge). Nicolaisen and Frandsen, however, explicitly mention that 3R reflects investments and returns on investments of a journal. They also note that the denominator of 3R depends on the editorial practices of just one journal, whereas the other impact indicators depend on editorial practices of many journals. Hence, for a journal editor it would be easier to manipulate 3R. The idea of investments and returns on investments is also proposed in (Liang & Rousseau, 2008), where the yield period of a journal is defined as the time needed to accumulate the same number of citations as the number of references included during the period of study.

It has recently been observed that there is an inflation of impact factors (Althouse et al., 2009; Stock, 2009). Over the period 1994-2005 this 'scientific' inflation was about 2.6% per year. Is this due to the growth of science, or to reference lists that become longer and longer? It was shown by Althouse et al. (2009) that the growth of the scientific literature as a whole has very little influence on this inflation. They found that an increasing number of references per article is the greatest contributor to impact factor inflation over time. Increase in international collaboration seems to be an important factor leading to longer reference lists (Persson et al., 2004). Moed mentions explicitly that SNIP does not take the growth of the literature into account. However, by the results obtained by Althouse et al. (2009) this aspect seems to be of minor importance.

CONCLUSION

Already in 2002 Glänzel and Moed (2002) announced new and exciting challenges for citation analysis. As the ISI journal impact factor was described as having several severe methodological shortcomings, they were convinced that appropriate new methodological approaches might help overcoming the limitations of the original measures. Hence, the authors of this contribution welcome the useful additions to our profession's arsenal of tools described in this *Newsletter*. We leave it to the future to determine which proposal(s) will turn out to be most useful. It goes without saying that any journal impact measure, including the new proposals, cannot be used as a proxy for the impact of single articles, or as a measure for the research quality of a scientist.

ACKNOWLEDGEMENT

The authors thank Michel Zitt for some clarifications. We also thank him and Wolfgang Glänzel for suggestions improving the contents of this article.

REFERENCES

Althouse, B.M., West, J.D., Bergstrom, C.T. 7 Bergstrom, T. (2009). Differences in impact factor across fields and over time. *Journal of the American Society for Information Science & Technology*, 60(1), 27-34.

Bergstrom, C.T. (2007). Eigenfactor: measuring the value and prestige of scholarly journals. *College & Research Libraries News*, 68(5).

Bollen, J., Rodriguez, M.A. & Van de Sompel, H. (2006). Journal status. *Scientometrics*, 69(3), 669-687.

Frandsen, T.F. & Rousseau, R. (2005). Article impact calculated over arbitrary periods. *Journal of the American Society for Information Science & Technology*, 56(1), 58-62.

Garfield, E. (2006). The history and meaning of the journal impact factor. *Journal of the American Medical Association*, 295(1), 90-93.

Glänzel, W. & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics,* 53(2), 171-193.

Grupo Scimago (2007). SCImago journal & country rank: un Nuevo portal, dos nuevos rankings. *El Profesional de la Información*, 16(6), 645-646.

Liang, LM. & Rousseau, R. (2008). Yield sequences as journal attractivity indicators: payback times for *Science* and *Nature*. *Journal of Documentation*, 64(2), 229-245.

Moed, H.F. (2010). Measuring contextual citation impact of scientific journals. *Journal of Informetrics*, doi:10.1016/j.joi.2010.01.002. Available from http://arxiv.org/abs/0911.2632.

Nicolaisen, J. & Frandsen, T.F. (2008). The reference return ratio. *Journal of Informetrics*, 2(2), 128-135.

Persson, O., Glänzel, W. & Danell, R. (2004). Inflationary bibliometric values: the role of scientific collaboration and the need for relative indicators in evaluative studies. Scientometrics, 60(3), 421-432.

Pinski G. & Narin F. (1976). Citation influence for journal aggregates of scientific publications: theory, with application to the literature of physics. Information Processing and Management, 12, 297–312

Rousseau R. (1988). Citation distribution of pure mathematics journals. In: *Informetrics 87/88* (Egghe L. & Rousseau R., eds) Elsevier, Amsterdam: 249-262.

Rousseau, R. (2005). Conglomerates as a general framework for informetric research. *Information Processing & Management*, 41(6), 1360-1368.

Stock, W.G. (2009). The inflation of impact factors of scientific journals. *ChemPhysChem*, 10, 2193-2196.

Waltman, L. & Van Eck, N.J. (2010). The relation between eigenfactor, audience factor, and influence weight. Available from http://arxiv.org/abs/1003.2198.

Zitt, M. (2010). Experimental measure of the audience factor of scientific journals. Available from: http://www.obs-ost.fr/fileadmin/medias/tx_ostdocuments/PresentationAF_UK_BAT.pdf

Zitt, M. , Ramanana-Rahary, S. & Bassecoulard, E. (2005). Relativity of citation performance and excellence measures: from cross-field to cross-scale effects of field-normalisation. *Scientometrics*, 63(2), 373-401.

Zitt, M. & Small, H. (2008). Modifying the journal impact factor by fractional citation weighting: the audience factor. *Journal of the American Society for Information Science and Technology,* 59(11), 1856-1860.

INDIAN CITATION INDEX (ICI):

METHOD, MATERIAL AND CONSTRUCTION APPROACH

PRAKASH CHAND
Head ICI Project & Ex-Scientist,
NISCAIR (CSIR), New Delhi, India
pchand.delhi[at]gmail.com

ABSTRACT: This paper presents in brief the strategy, methods, material used for constructing Indian Citation Index (ICI). It also describes its scope, objectives, contents coverage policy, document types' policy, data inflow mechanism for processing, content processing steps, process model, and method of assigning indexing terms, technology used, and future plan of its development. The database is now in the pre-beta stage and accessible from www.indiancitationindex.com.

1.0 INTRODUCTION

India contributes significantly to world spectrum of R&D literature by publishing over 1500 peer reviewed journals. However, coverage of their contents in international bibliographic/citation indexes is significantly low. It is estimated that over 50% of India's R&D knowledge contents published in local national journals are not covered by international databases. Argued reasons thereof are low accessibility, availability, poor quality of contents, late run of publications, lack of sound review system, editorial policy, and over all quality standards etc. The non-inclusion of this significant proportion of journals in International databases has serious consequences. The research addressed by them does not inform the research community, and the research findings remain oblivious to the scholars. One of the strong arguments (for their non-inclusion) is that these journals address specific local issues in agriculture, food, health, environmental impact etc. However, the research reflected in these journals can have important and useful bearing for developing and improvised economies. Many local problems have interdependence with the global system and local solutions can be adopted in wider setting. The climate change debate epitomizes this argument. Another consequence is the 'poor' estimation/evaluation of Indian research. It lacks realistic representation as it is based on Indian knowledge contents covered in and re-

trieved from WoS or SCOPUS database(s). Some of the above cited reasons strongly convinced the author to construct Indian Citation Index (ICI) for true representation and reflection of India's scholarly literature. It is expected that ICI would provide access to the scholarly contents of Indian journals and also help in realistic evaluation reports. This communication briefly outlines the development of 'Indian Citation Index (ICI)'

2.0. ABOUT INDIAN CITATION INDEX (ICI)

Indian Citation Index (ICI) is a home grown abstract and citation database, with multi-disciplinary information/knowledge contents from about 1000 top Indian scholarly journals. This database is intended to bridge up the long awaited gap of Indian scholarly information for access and evaluation. The database in its first phase would provide access to over .3 million source and 5 million cited records with five year's back files of about 1000 top scholarly publications. The ICI database aims to provide objective content and powerful tools to search, track, and measure and collaborate in the fields of sciences, social sciences, arts, and humanities; to turn raw data/information into the powerful knowledge you need. Building Indian Citation Index (ICI) hopes to serve as necessary supplementary tool to existing international citation indexes

2.1.1 SCOPE

Indian R&D literature across all disciplines i.e. sciences, technology, medicine, agriculture, social sciences and humanities published in journals/ serials or in other documents that originate locally from India.

2.1.2 OBJECTIVES

- To ensure access to articles published in local Indian R&D literature at national & global level
- To reflect and represent true picture of locally published Indian scholarly contributions at national and global level
- To have an authentic tool/ground for effective, rigorous evaluation of Indian scholarly works

3.0. ICI COVERAGE POLICY

Currently, ICI has planned to cover about 1000 Indian titles of R&D nature with five year's depth i.e.2004 onwards, provided all records in a journal comply with the document type policy set out for ICI database. The titles that have no ISSN or have irregular publication schedules are not selected for ICI. Also, the item types, e.g. trade publications, where not all articles fit the Document Type Policy or items do not carry references are not included. After completing five year's depth of coverage in first phase, the rest of the archive (complete back files) of selected journals will be added in the database.

3.1. METHODOLOGY:

Following steps are being followed for contents coverage of ICI:

Identification of locally published Indian journals: In initial selection of journals a fairly liberal criteria is applied. Final selection based on quality is left to the impact factor to be generated latter after completion of first phase of ICI. Currently in selection of journals following parameters are followed:

1. Indian journals
2. R&D nature
3. Timeliness
4. Editorial board consisting of international repute experts
5. Coverage – all disciplines
6. Documents/journals with ISSN
7. English language documents, at least it carries an English language title and abstract of all research articles. The full text article may be in any language

SN	DOCUMENT TYPES	DEFINITION
1	Research article	Article of original research carried out by author(s)
2	Review article	Significant review of original research literature
3	Short communication	Brief research communication/ news/views/ article
4	Editorial	Editorial item summarizing several articles or providing editorial opinions or news. Letter to or correspondence with the editor
5	Research note	Original research communication in discussion or commentary form
6	Case study	Analytical findings of an individual case and communicated in journal publication
7	Research method	Reporting of new research methods
8	Opinion papers	Article provide opinion of individual or group
9	Observations (R&D)	Expression of observations on R&D method/experiment/test and findings etc
10	Special articles	Invited article, memorial lecture/working paper/special paper/expert views/ comments
11	Proceedings paper	The paper published in a Conference/Symposium/Seminar/Workshops etc or summarization of all papers from conference proceedings.

Table 1 *Types of scholarly material/documents selected for ICI database*

8. Over all high quality journals
9. Documents or journals with minimum one issue per year
10. Quality criteria of journal is assessed/judged on:
    - → Authority – reputation of publisher, diversity in affiliation of authors, editorial board of international repute etc
    - → Peer reviewed
    - → Availability
    - → Popularity

Sources consulted for preparation of Journals list

- WoS covered Indian journals
- SCOPUS covered Indian journals
- Open Access Indian journals
- Indian Science Abstracts (ISA) covered journals
- Other left out Indian journals particularly from Agriculture and Health sciences
- Indian ISSN data
- Directory of Indian periodicals

4.0. DOCUMENT TYPES POLICY

Considering general time lag in Indian publications, ICI provide an opportunity initially for two years to such publications/publishers for bringing their publications on time. See Table 1 for types of scholarly material/documents that are being selected for ICI database.

The following types of material (documents) is NOT the subject matter of ICI coverage

- Advertisement; News and Views
- Abstracts of dissertations or articles
- Announcements etc
- Bibliographies
- From the desk
- Digests; Popular article; Obituaries; Memorial lecture
- Book reviews

5.0. DATA FOR ICI CONSTRUCTION

Two pronged strategy for data identification, collection and insertion is followed and both are to run concurrently:

- **First:** The Indian journals available in e-form with IndianJournals.com and OA are taken together and their contents are examined, and processed based on defined parameters for decision to include in the ICI database
- **Second:** The remaining journals contents which are not available in e-form are identified and collected in hard copy through photocopy mode i.e. content page of each issue of a journal and first and last page of each identified article of that journal's issue from different libraries like National Science Library (NSL) and Indian Agricultural Research Institute (IARI) Library. Data entry of such collected documents is done through a defined worksheet.

5.1. WORKFLOW DIAGRAM

See Figure 1 for the Workflow diagram.

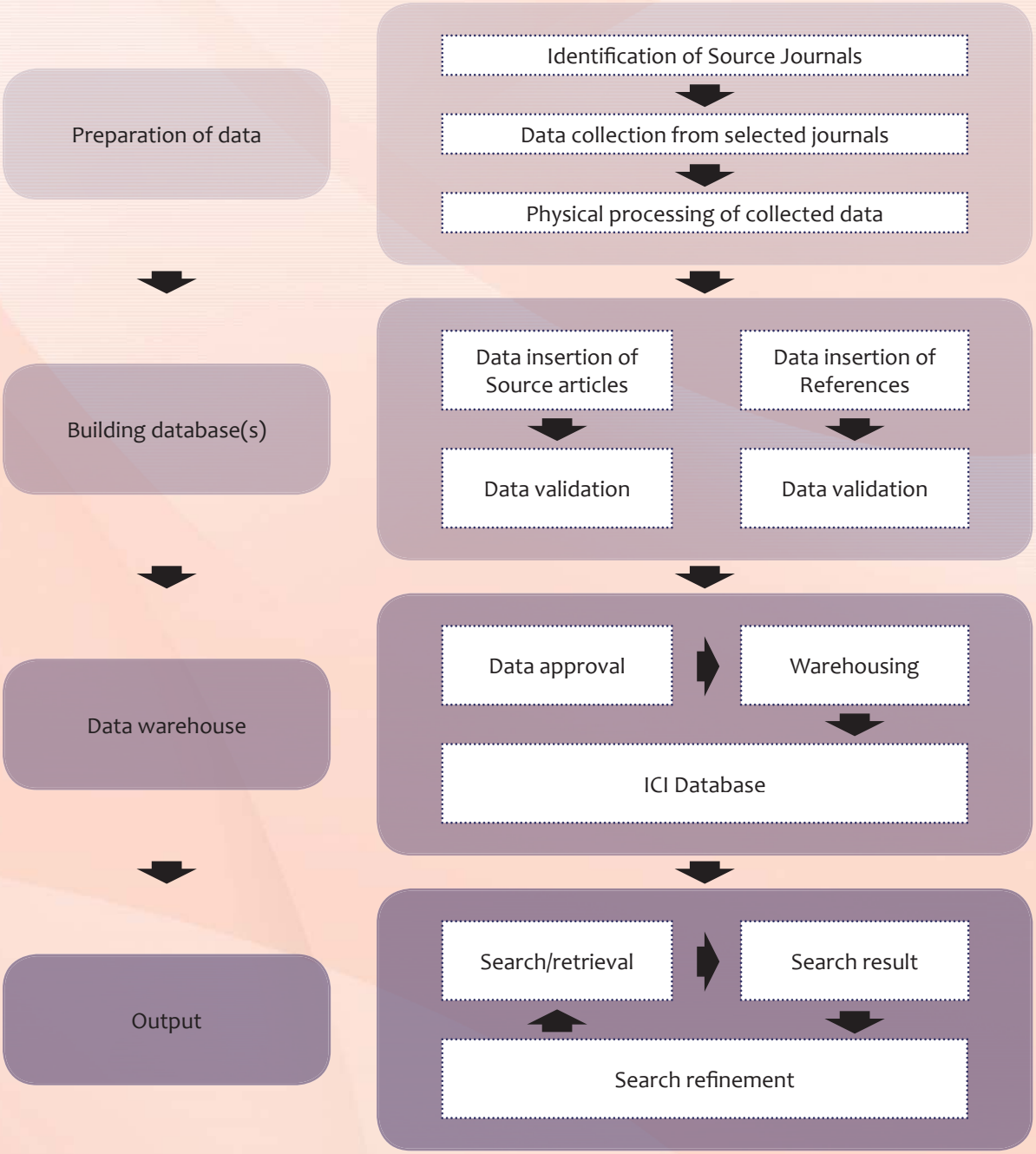


Figure 1 *Workflow diagram*

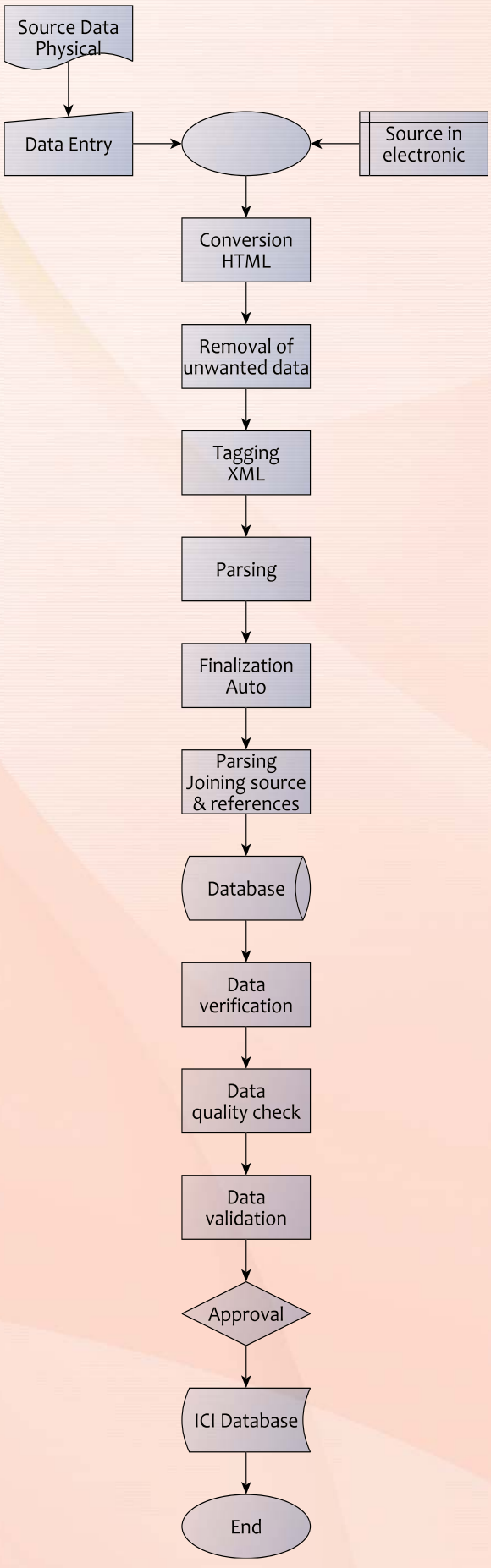


Figure 2 *Data insertion process*

6.0. INDEXING ON ICI

ICI team manually adds Keywords/Index terms derived from title of the article/abstract and supplement by thesauri in order to improve recall while search is done. A team of professional indexers assigns Keywords/ Index terms to each record(s) using free terms and controlled vocabularies/thesauri.

7.0. AUTHOR (S) AFFILIATION

As a policy, ICI records affiliation of all authors of an article (s); however, there are also few records in which only the correspondence address is given as the affiliation information. For documents where the correspondence address is not marked as such, the affiliation information of the first author is presumed to be the correspondence address.

8.0. CONTENT PROCESSING

Processing of content in ICI involves four main steps that add quality and context to records:

STEP 1: CAPTURING BIBLIOGRAPHICAL AND CITATION INFORMATION

Bibliographical & citation information, such as article title, author(s), affiliation(s), abstracts, and references, are captured from the source as collected by the ICI input team. During processing, Keywords/Index terms are added to the data based on consulting various relevant thesauri and also free terms taken from title and abstract of source article and added manually by subject specialists.

STEP 2: PRODUCTION DATABASE

These records are then loaded into a production database. This is done through our in-house Data Entry system, when the data is entered there are in built functions to validate the data as initial quality checks. Once data is entered there is a separate team to

Figure 3 *Screenshots of the Indian Citation Index (ICI)*

verify the data against the original source materials and make necessary corrections for quality control and data integrity. Also, ICI team of professional experts assigning subject headings and incorporate necessary modification in already entered Keywords.

STEP 3: ICI WAREHOUSE

The records are then loaded into the ICI Warehouse. The main function of the warehouse is to link references and articles, and in this way to determine how many times an article is cited. Along with this all the records go through an approval process where our data is verified, then approved before available for search/view through indiancitationindex.com.

STEP 4: INDIANCITATIONINDEX.COM

All records are sent to indiancitationindex.com, where the database records are made searchable and retrievable.

Currently processing of content in ICI takes approximately two to three weeks time to ensure the quality of ICI content and establish the right context and linking.

9.0. DATA INSERTION PROCESS

See the flowchart in Figure 2.

10.0. SCREEN (HOME PAGE)

The users may access ICI clicking on www.indiancitationindex.com and can see all search features (*Fig 3*). The product is in pre beta stage therefore all envisaged features and data validations etc. are yet to be built.

11.0 FUTURE PLAN

ICI database like other similar databases is ever growing. At this stage it is in developing phase and few more months will be invested to equip it with full features available in similar citation databases. ICI with its strong R&D team is working hard to launch it as a commercial product by middle of the year. Also, ICI is making comparative analysis of the search and analytic features of other citation databases in view of users and utility perspective and accordingly taking decision to build all such identified functionalities in the database. The ICI team is also experimenting with novel features that can enhance the search and retrieval features and provide scholars to uncover further insights of Indian research activity.

ACKNOWLEDGEMENT

The development of Indian Citation Database is supported by Divan Enterprises Ltd.

BIBLIOGRAPHY

1. Bharat Bhushan and Banwari Lal. Indian Science Citation Index: A strategy. Program, 1991, 25(1), 59 - 67

2. Yishan Wu. Yuntao Pan, Yuhua Zhang, Zheng Ma, Jhingan Pang, Hong Guo, Bo Xu, Zhioing Yang. China Scientific and Technical Papers and Citations (CSTPC): History, impact and outlook. Scientometrics, 2004, 60(3), 385 - 387

3. Kuang - hua Chen. The construction of the Taiwan Humanities Citation Index. Online Information Review, 2004, 28(6),, 410 - 419

4. Pero Sipka. The Serbian Citation Index: Context and content. The 10th International Conference of the Information Society for Scientometrics and Informatrics, Stockholm, Sweden, July 24 - 28, 2005, Vol. 2, pages 710 – 711

5. Sujit Bhattacharya and B.M. Gupta. Towards developing a national bibliographic database in Science and Technology in India. University News, Vol. 42, 2004.

A PERFORMANCE INDICATOR BASED ON COMPLETE DATA FOR THE SCIENTIFIC PUBLICATION OUTPUT AT RESEARCH INSTITUTIONS

GUNNAR SIVERTSEN
Norwegian Institute for Studies in Innovation, Research and Education (NIFU STEP)
gunnar.sivertsen[at]nifustep.no

In 2005, Norway implemented a performance based funding model for the Higher Education Sector with an indicator based on complete data (in journals, series and books) for the scientific publication output at the level of institutions (Sivertsen 2006, 2008). The same indicator was implemented in Denmark in 2009 (Schneider 2009). Belgium (Flanders) aims at adding a similar indicator with data from the social sciences and humanities to the so called BOF-key in 2011. Norway has recently made all publicly funded research institutions outside the Higher Education Sector use the indicator as well.

The data for the indicator are produced by the institutions themselves in a shared quality assurance system that creates an open and transparent national database with references to all scientific publications from all institutions. The data are not only used by the Government for funding, but also by the institutions themselves for internal purposes at various levels, e.g. in redistribution of funds at local levels, in annual reports and statistics, in CV's and applications, in open Current Research Information Systems, and as bibliographical references that lead on to the available full text of publications.

The creation of a national database of this kind starts with acknowledging that not only the Government, but also the research institutions themselves, need continous and structured information about their research activities for internal and external purposes. In bibliometrics, we are used to serve the institutions from the outside with professionally processed data and analysis of their activities. But institutions also produce their own data. At the department level, there is a tradition for listing the staff's scientific production as bibliographic references in annual reports. For some years now, this tradition has evolved into creating information systems with databases that record all scientific publications at the level of institutions. Since these databases demand input from the researchers themselves, there is often a problem with incompleteness, and the data may lack the necessary structure and coherence that can be seen in professional bibliographic data sources like the *Web of Science* (WoS). On the other hand, institutional databases may include scientific and scholarly publications in books, series or journals that are not covered by WoS or similar data sources. Institutional databases may also solve the problem with attributing author names and addresses to unique persons and institutions, and they may add routines of quality assurance and validation to this process.

The design of the indicator for performance based funding in Norway started with an ambition to get the best from both types of data sources, creating complete and quality-controlled structured bibliographic data at the institutional level, not only for each of the institutions, but in one and the same national database for all institutions. An agreement on this ambition was reached between the Ministry of Research and Education and the institutions in 2003-2004 on the basis of a design of the indicator and its database that was developed by the Norwegian Association of Higher Education Institutions (2004). This organization – a parallel to the Rectors' conferences in other countries – has since then had the responsibility for the maintenance and further development of the indicator and its database. This responsibility is carried out by a National Publishing Board with representatives at the level of deans from all types of institutions and major research areas. My own role has been to contribute as an expert in the design phase and as a consultant to the National Publishing Board after the implementation. I have had a similar role in Denmark since 2007.

In both countries, the transformation of the publication data into the indicator for the performance based funding model is based on routines and mechanisms on three levels:

- *Delimitation and documentation*. A definition defines, and the publication database records and validates, structured and standardized bibliographic references to different types of scholarly publications in different fields of research. Author names and addresses are at the same time connected to persons and institutions.
- *Comparable measurement*. A system of "weights" takes into consideration different publishing traditions and make them comparable in the same measurement of "Publication points" at the level of institutions.
- *Incentives and funding*. In the overall budget each year, a certain percentage of the basic annual funding from the Government is redistributed between the institutions according to their shares in the total Publication points.

These three components constitute what has been called "the Norwegian model" abroad. I will explain them in further detail below. At the end, I will discuss the effects of the model so far in Norway and answer a couple of typical questions that have been raised among bibliometricians abroad.

DELIMITATION AND DOCUMENTATION

Institutional databases are potentially endless in their coverage and at the same time

incomplete because they depend on input from all researchers. These databases cannot support performance based funding before the data are structured, validated, and complete within a definition. With regard to the completeness of the data, our experience is that this aim was not fulfilled at an acceptable level before a year after the database was connected to a funding model. The structure and quality of the data was attained in the following way.

We decided to cover only the scientific and scholarly publications completely. They are not endless if they are defined in the following way: A scientific or scholarly publication must:

1. present new insight
2. in a form that allows the research findings to be verified and/or used in new research activity
3. in a language and with a distribution that makes the publication accessible for a relevant audience of researchers
4. in a publication channel (journal, series, book publisher) with peer review

In order to control the third and fourth requirement of this definition and to standardize the data, dynamic lists of so far 18,000 ISSN-titles and 1,000 publishers of books were created. These lists can be downloaded at http://dbh.nsd.uib.no/kanaler/. Suggestions for additions are received through the same web page. In addition to requirement 3 and 4, a publication channel must publish on behalf of authors from more than one institution to be included.

The first two requirements in the definition point at the *publication type* as a restriction. Not all articles in *Nature* are scientific, neither are all books on Oxford University Press. It depends on the publication type, which must be scientific or scholarly according to requirement 1 and 2 in the definition. This solution can be recognized as coming from bibliometric research where a certain selection of journals and publication types (articles, reviews, etc.) often defines a dataset. The same solution is needed, but not sufficient in a database with input from researchers. A quality assurance system is needed as well. Let me introduce it by mentioning costs.

The largest costs of a database of this kind are connected to the maintenance of the institutional electronic information system for internal and external use in which the bibliography of the institution's scientific output is a component. These major costs have already been covered in institutions running their own bibliographical databases. The costs can be *shared and reduced* in one common national information system. But this is only the first step forward. The next step is a quality assurance system working at all levels, and it demands additional costs. The system is not only necessary for a legitimate funding model. Without it, the institutions would not be able to use the data for research statistics and other internal and comparative purposes. Here follows a short description of the quality assurance system that we have found necessary in order to produce data of sufficient completeness and quality. There are four levels of data recording and control:

1. Input by researcher and/or assistant
2. Support and annual quality control at department level
3. Cases of doubt are resolved at faculty/university level
4. Questions of general importance are resolved by the National Publishing Board

As a principle, each institution is responsible for the quality of its own data. Uniform guidelines giving further explanations of the definition above have been written in collaboration between the Ministry and National Publishing Board. People involved at the second and third level meet regularly within and between universities to discuss guidelines, routines and problems. It is important that all institutions share the same database, because the transparency of the database within and between the institutions stimulates quality assurance. Everybody can see each other's data at the end.

As an aid to data recording and quality assurance, references to a large part of the pub-

lications are imported from certain parallel professional bibliographic data sources, among them *Web of Science* by agreement with Thomson Reuters, before they are validated by the researchers or their assistants and transformed into the format of the database.

Although several people and levels are involved in quality assurance, the work load should not be overestimated. A publication is recorded only once as one unique publication that may have several authors, not one time each for all of the authors. With unique (fractionalized) counts one will find that the average number of scientific or scholarly publications per researcher per year is somewhere between one and two.

An important principle in the uniform guidelines is how to attribute publications to institutions. This can either be done on the basis of the person's present institutional affiliation(s) or on the basis of the affiliation(s) that the author chose to indicate as address(es) in the publication itself. We chose the last alternative, taking into account the need for publicly available verification of the relation, the consequences of mobility, and the ethical requirements to authorship that are used for example in biomedicine (The Vancouver Recommendations). The chosen alternative means that it does not affect funding if an institution decides to "hire a CV". On the institutional level – in the funding model – a publication remains at the institutions that were originally credited in the publication itself. At the individual level – in a personal CV – publications still, of course, follow the authors.

I have now described the principles of the publication database as such. I will conclude this section by giving the three main reasons for the choice of the national database as a platform for the bibliometric indicator for the funding model:

- Completeness: All scientific and scholarly publications are included. The limit of the data is *defined* by the institutions in collaboration, not decided by a certain commercial data source.
- Transparency: Every institution can see and check all other institutions' data. The institutions own their data. The database is also online and open to society at large.
- Multiple use of the data in CV's, applications, evaluations, annual reports, internal administration, bibliography for Open Archives, links to full text, etc.

COMPARABLE MEASUREMENT

In the measurement for the funding formula by the end of each year, the publications are *weighted* as they are counted. Thereby, and for this purpose only, the bibliometric indicator has been constructed. It is meant to balance between field specific publishing patterns, thereby making these patterns comparable and giving a balanced representation of the research activity at institutions that may have different research profiles (e.g. a technical versus a general university). In one dimension, three main publication types are given different weights: articles in ISSN-titles, articles in books (ISBN) and books (ISBN). In another dimension, publication channels are divided into two levels in order to stimulate publishing in the most prestigious and demanding publication channels within each field of research. The highest level is named "Level 2". It includes only the leading and most selective international journals, series and book publishers, and they may not account for more than about 20 per cent of the world's publications in each field of research. The weighting of publications by type and channel is shown in *table 1*.

	CHANNELS AT (THE NORMAL) LEVEL 1	CHANNELS AT (THE HIGH) LEVEL 2
Articles in ISSN-titles	1	3
Articles in ISBN-titles	0,7	1
Books (ISBN-titles)	5	8

Table 1 Publication points in Norway

These publication points are given to the publications at the level of institutions, not to the authors. The points of publications with more than one institutional address are *fractionalized* among the participating institutions according to their number of participating authors. In publications with high numbers of authors, each institution is still credited a minimum of one tenth of the publication's points.

The high "Level 2" is meant to stimulate the institutions to encourage their researchers to seek the most competent and critical peer review and a wide relevant readership. We also want to avoid a tendency to increase publishing in less significant channels. The different weights given to publications on the two levels are *not* meant to express a quality judgement of each publication or an expectation of its potential citation impact.

The division of the publication channels into two levels is difficult and controversial. The highest "Level 2" is revised annually in collaboration between the national councils in each discipline or field of research and the National Publishing Board. Bibliometric statistics (world production versus national production in channels on both levels, and citation statistics for publication channels) are used as an aid in this process, but not as criteria by themselves. Since it has been misunderstood in a recent publication that the Norwegian model is based on "Journal Impact Factors" (Sandström and Sandström 2009), I emphasize here that the final decisions about the highest level are made by qualitative judgment and consensus among peers in a process organized on a national level by the universities themselves.

INCENTIVES AND FUNDING

Like most other countries, Norway and Denmark have a so called "dual funding system" for research with direct block grants to the institutions on the one side and additional, competitive funding from research councils and similar organizations on the national or international level on other side. The indicator we have in focus here, only belongs to the direct funding on the institutional level. It is not used by the research councils or other organizations responsible for indirect competition-based funding, neither is it used to evaluate performances or applications for funds or positions at the individual level. The publication indicator is meant to measure and stimulate the research activity at the level of *institutions* and to enhance the focus and priority they give to research as *organizations*.

Performance based funding is only a smaller part of the total direct funding of research at Norwegian institutions. The performance indicators also include external funding and doctoral dissertations. The publication indicator only reallocates 2 per cent of the total expenses of the Higher Education Sector. In Norway, one publication point represents no more than 5,000 Euro, and even less in Denmark, at least initially. Still, the publication indicator receives a lot of attention by almost all researchers. In Norway, it is now possible to talk of effects six years after implementation. As we shall see, the indicator has stimulated research activity in such a degree that the change can be detected by indicators on the international level.

EFFECTS AND EXPERIENCES

One effect of the indicator in Norway is a substantial growth in the publication output in the Higher Education Sector. The growth has taken place both in the higher and lower strata of the hierarchy of publication channels, as shown in *figure 1*. Since 2005 (the first year with relatively complete data), the publication activity on level 1 (normal) has increased by 39 per cent, while publication activity on level 2 (internationally leading publication channels) has increased by 55 per cent. So far, Norway seems to have avoided that the increase mainly occurs in the less significant publication channels, as has been reported in Australia (Butler 2004).

As seen in *figure 2*, Norway's share in the world's article production as measured

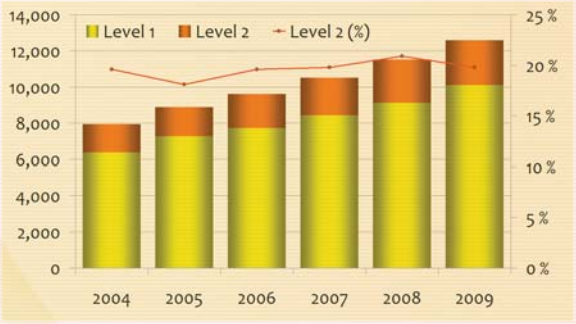


Figure 1 *Publication points in the Norwegian Higher Education Sector 2004-2009. Level 2 represents internationally leading publication channels expected to publish around 20 per cent of the total. The red line and the axis on the right side represent the observed percentages on Level 2.*

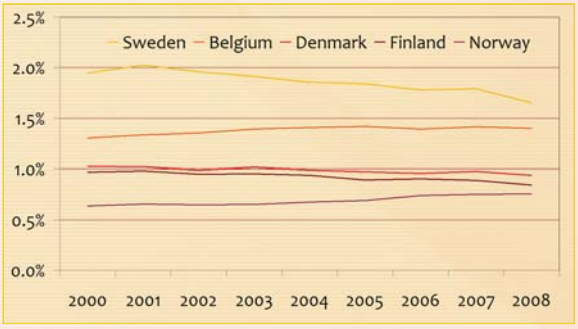


Figure 2 *Shares in the world's scientific output in Web of Science 2000-2008. Source: National Science Indicators (NSI), Thomson Reuters.*

within *Web of Science* has increased while there have been decreases in the neighbouring Scandinavian countries. Most of the EU and OECD countries have decreases in the same period. Belgium's positive trend is similar to Norway's and can possibly explained by the introduction of bibliometric indicators in the BOF-key for the Flemish universities in 2004 (Debackere and Glänzel 2004).

As mentioned above, the aim of the performance indicator is not an increase in scientific output itself, but a stimulation of the institutions as organizers of research. Here are some important qualitative changes that are reported within the networks and meetings of The Norwegian Association of Higher Education Institutions:

- The institutions have stronger incentives to facilitate research for their researchers
- Research is now perceived as a common and institutional responsibility, not only as an individual task
- New publications receive attention, not only from external peers, but also internally from the institution
- Research management improves with the aid of complete bibliometric information about the research activities

Apart from this, there are of course debates and disagreements on the indicator among the researchers themselves. Protagonists and antagonists are found in all disciplines and major research areas. There is no sign that the disciplines or research areas as such are divided in the question of whether the indicator – or performance based funding in itself – does good or harm to research.

DISCUSSION

Before I end, I will clarify two questions that have seemed important so far in the reception of the model among experts in other countries. I will also say a few words about how this model might affect bibliometric research.

The model has been discussed as a possible solution to the limitations of other bibliometric data sources in the humanities and social sciences (Dolan 2006, Hicks and Wang 2010). But, as seen above, the model is in fact meant to give a balanced representation of research activities in *all* fields. Still, since it includes scholarly publications in journals, series and books that are not covered by the main international indexing services, it can become *part of* a solution for the humanities and social sciences if other countries create similar national databases.

A common reaction to the model among bibliometricians is that its focus on "quantity of outputs" rather than "citation impact" is a serious limitation in comparison with other approaches like WoS bibliometrics. I have three answers to this. One is that the effect of the model on Norwegian research is now an empirical question. Norway's citation rate relative to the world has remained 25 per cent over the average during the last six years, while it was down at 10 per cent over the average ten years ago. And as seen above,

the scientific output in publications has increased in both the higher and lower strata of publication channels.

My second answer is that the model is not opposed to using citations as an additional indicator in the performance based funding model. Indeed, it can be combined with such attempts, and this combination will probably be introduced in Belgium (Flanders) within a year or two. Also, before a citation analysis is made for a performance based funding model, we need the institutions to agree on how the publications are allocated between them, especially in the cases where these publications are shared between two or more institutions. The model provides for this procedure.

My third answer is that making all institutions in a country agree on how the bibliometric data should be produced – and then make them produce the data – is a little bit different from desk top bibliometrics with bought or downloaded datasets. The first procedure allows for complete data within a definition that the institutions agree on. The other procedure does none of this, but it may include citations. If we do not insist that only one of these procedures is valid, we might in the end get a better balance and collaboration between the commercial suppliers of citation data on the one side, and the research institutions, their authorities and their international organizations on the other.

Another possible influence on bibliometrics is that the model produces complete data for a whole nation and in all fields of research down to the level of individuals and in a setting (the institution) where other types of information about the resources and employees may be available for analysis. This may open up for new combinations of bibliometrics and social studies of science in general.

REFERENCES

Butler, L. (2004), What happens when funding is linked to publication counts? In: Moed, H.F.; Glänzel, W; Schmoch, U. (Eds.), Handbook of Quantitative Science and Technology, Kluwer Academic Publishers, 389-340.

Debackere, K. and Glänzel, W. (2004), Using a bibliometric approach to support research policy making: The case of the Flemish BOF-key, Scientometrics 59 (2): 253-276.

Dolan, C. (2007), Feasibility study: The evaluation and benchmarking of humanities research in Europe, HERA, Humanities in the European Research Area (www.heranet.info).

Hicks, D. and J. Wang (2010), Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project (in press).

Norwegian Association of Higher Education Institutions (2004): A bibliometric Model for Performance-based budgeting of Research Institutions, (English version), http://www.uhr.no/documents/Rapport_fra_UHR_prosjektet_4_11_engCJS_endelig_versjon_av_hele_oversettelsen.pdf

Sandström, U. and. E. Sandström (2009), The field factor: towards a metric for academic institutions, Research Evaluation, 18(1), 243-250.

Schneider, J. (2009), An outline of the bibliometric indicator used for performance-based funding of research institutions in Norway, European Political Science, 8 (3), 364-378.

Sivertsen, G. (2006), A bibliometric model for performance based budgeting of research institutions, pp. 133-135 in: Proceedings of the 9th International Conference on Science & Technology Indicators, 7-9 September 2006, Leuven, Belgium. Leuven: Katholieke Universiteit Leuven – Steunpunt O&O Statistieken.

Sivertsen, G. (2008). Experiences with a bibliometric model for performance based funding of research institutions. In: Gorraiz J. and Schiebel E. (eds). *Excellence and Emergence. A New Challenge for the Combination of Quantitative and Qualitative Approaches.* Book of Abstracts. 10th International Conference on Science & Technology Indicators. Vienna, Austria. 17-20 September 2008. Vienna (AUT): Austrian Research Centers GmbH, pp. 126-128.