

Managing your assets in the publication economy

Ulf Kronman

The issue this article aims to address is the fact that publications may nowadays be used to assess impact and quality of research in ways academics may not be fully aware of. During recent years, scholarly publications have gained in importance, not primarily as the traditional vehicle for the dissemination of new scientific findings, but as a foundation for assessing the production and impact of organizations, research groups and individual researchers. This means that publications as artefacts *per se* are starting to play a new important role in the scientific community and that researchers need to be aware of how publication and citation counts are being used to assess their research and the outreach, impact and reputation of their mother organization. University rankings, for instance, often have some parameters based on the publishing of the ranked institution. This article is thus not about scientific *writing* as such; it focuses on what happens to your publication *after* the publishing has taken place and on aspects to take into account while planning the publishing of your article, report or book.

The need to assess research seems to be ever growing and in the urgent need for some hard numbers the evaluators turn to counting publications and citations. For academics, his or her research is thus not always assessed by peers that understand what is written, but by people that do not have the time to read the publications or even would understand the content if they had the time to read. Instead, these evaluators have to resort to using metrics from impact *proxies*, as the renownance of the journals you publish in, or counting citations from people that have read your publications and hopefully understood what was said in them, and then cited your work.

There are a lot of nitty-gritty details involved in the production of bibliometrics, and this article will describe some of them. By outlining these aspects of bibliometrics you can consider how they might affect the metric outcome of your publications when producing and managing them in the future. Before continuing with the publication management advices, a short disclaimer: Even if metrics and statistical aspects of publications are gaining importance for assessment and funding, it is still the *quality* of the research behind the publications and the *dissemination* of research findings to peers and general public that should be the primary goal for your publishing. But, on the other hand, there is no contradiction between doing high quality research and establishing a good communication with fellow peers, and to consider some means for making the research results more visible and influential, utilizing some of the considerations pointed out in this article.

A few words about the disposition of the article: firstly, I will focus on the situations where counts and impact of publications matter. Secondly, I will report on how publications are being

measured in bibliometric studies and how this affects their impact in the results of the studies. Thirdly, I will suggest some methods for improving the results in bibliometric studies based on the details mentioned in the previous part. Lastly, I will round off the article with a discussion on why we see this economy of publications emerging.

Publications as measures of production and impact

As mentioned in the introduction, publication measures are increasingly being used as tools in the race for funding in a world of tightening competition for shares of constrained budgets. For you as a researcher this means that you have to keep a good record of your publishing and see to it that all your publications are being visible and attributed to you in the various assessments based on publications. In this section I will outline a number of situations where your publication record may play an important role for you and your organization.

Publication lists for CV's and web pages: exhibiting excellence

The most important tool for exhibiting your scholarly impact as an individual researcher is of course the publication list that is a part of your *curriculum vitae* (CV). Many researchers keep their publication list as word-processing documents, in local EndNote databases or on static or dynamic CV web pages. To keep a local list of publication records in a file and on a local website are both labour-intensive and a bit old-fashioned in these network-based times. If your organization runs a publication database – often called a *publication repository* – your chances are good that both the publication listing for your CV and for your personal web page can be generated dynamically from this

publication database. Your tasks are to keep the database updated with your publication records, and preferably also upload the full text of the publications when possible. Contact the local support for your publication database – usually situated at the university library – for information on how to enter records and get listings from the database.

Research evaluations and publication based funding schemes

Research evaluations seem to be a prevailing trend among universities since the turn of the millennium. Every larger university seems to do one evaluation every third or fourth year, and usually publication statistics – *bibliometrics* – play an important role in these assessments. When publication statistics is gathered from commercial bibliographic databases it is very important that the publications can be attributed to you and your organization.

Bibliometrics is also playing a role of increasing importance in performance-based university and department funding all over the world. Among the Nordic countries Norway was first out with a model which allocates funding based on publication counts, Sweden was second with a citation-based model and now Denmark and Finland are introducing publication-based models of the same type as in Norway. When governmental bodies and research funders are starting to use bibliometrics for funding allocation, these kinds of measures are often mimicked at the organizational level by the local university managements. This means that your publication record may well play a role in the funding allocation to your department or research group. The same rules as for publication based research evaluations

apply here – it is important to have the attribution of publications in good order.

University ranking lists

Since the turn of the millennium, worldwide university rankings have become increasingly important and they are growing in number for each year. International students use the rankings when they choose among universities, universities use them to evaluate potential cooperation partners and they are used as a foundation for benchmarking and marketing. Politicians, decision makers and the industry also use the rankings to evaluate higher education institutions for policymaking and allocation of funding. Assessment of university research output in the terms of publications often constitutes an important part of the indicator set used to calculate the rankings.

The three most prestigious international rankings are the Shanghai Jiao Tong "The Academic Ranking of World Universities" (ARWU), the Times Higher Education "The World University Rankings" (THE/WUR) and QS "World University Rankings".

The ARWU-list is published yearly by the Institute of Higher Education at Shanghai Jiao Tong University. It was first produced in the year 2003 as part of a plan to create "world-class universities" in China. The methodology is relatively open, well documented and non-subjective. Universities are judged by unusual achievements, e.g. Nobel prizes and Fields Medals over a very long period. Large and old universities are favoured. Biomedical and physical sciences are given more weight than engineering, social sciences and humanities. The ARWU-list is fairly good at ranking the 50-100 most prestigious universities

in the world. Outside this scope, it is of limited value. This is acknowledged by the ARWU and therefore universities below rank 100 are grouped together in chunks of 50 and 100.

The most important measures in the THE/WUR ranking are the citation measures and the international reputation surveys. Together they account for two thirds of the total score. The QS list has been published in 2004-2009 by the journal Times Higher Education in cooperation with Quacquarelli Symonds Ltd (QS). Starting with 2010, QS is solely responsible for the ranking. Fifty per cent of the score is based on surveys, the rest on quantitative data.

However, and as was touched upon in passing above, there are a number of inherent shortcomings with university rankings – a few of the more notable ones being:

- All measured aspects of a university's activities and duties – education and research – are squeezed into one single measure, while another aspect – societal impact – is even neglected.
- Ranking is a way to make a champions' league and *magnify small differences* in the underlying indicator values. A small indicator discrepancy of 0.01 might be the only difference between two different rank positions in the list.
- When the final composite score is calculated, ranking providers *assign weights to each indicator* in the overall score. This means that the ranking provider's subjective judgement determines which indicators are more important.

Even though university rankings are crude and one-dimensional, they represent an easy-to-digest form of information to the broad public and they are very influential. This is what the European University Association has to say about the rankings in a recent report: “Despite their many shortcomings, biases and flaws ‘rankings enjoy a high level of acceptance among stakeholders and the wider public because of their simplicity and consumer-type information’. Thus, university rankings are not going to disappear; indeed, the number of rankings is expected to increase although they will become more specialised.”¹

Bibliometric studies and indicators: what counts where?

Bibliometrics is simply statistics done on publications, most commonly scholarly publications. A more precise definition might be: “Bibliometrics is the application of statistical methods to publications and is commonly used to assess scientific research through quantitative studies on research publications, primarily articles in peer-reviewed journals.”²

The reason for bibliometrics gaining in popularity and importance is the present urge of measurability in research assessment and funding allocation. Review by peers is still the gold standard in research assessment, but has the drawbacks that it usually not presents hard numbers and may also suffer from personal bias in judgements. Furthermore, it is hard work to do a peer-review of research, so statistics on publications and citations are often used as a shortcut for research assessment.

¹ Ruhvargers, 2011

² Karolinska Institutet, 2011

However, bibliometric indicators should rarely be used alone. If interpreted without caution they might be quite misleading. There are a number of reasons why good research may end up with poor bibliometric indicator values even though the research is of good quality. If the research lab is in a start-up phase, if the research field is very narrow, or the researchers publish their findings in forms and channels not covered by the bibliometric data sources the bibliometric indicators may show poor values, even if the research is of excellent quality.

The best usage of bibliometrics is to *supplement* peer judgement and supply extra statistical information to the reviewing experts, which preferably are knowledgeable of the organization and the research field that is being assessed. If the bibliometric indicators support the expert opinions, the experts can feel a bit more assured in their judgement. If the indicators contradict their opinions, it may be a signal for consideration and rethinking, or at least to try to explain the discrepancy between peer review and bibliometrics. After these initial words on bibliometrics in general I will go into the details on how bibliometric studies are performed.

Sources for bibliometrics

There are only a few data sources that capture enough publication data to be used as viable sources for a bibliometric study. The most important sources for bibliometric data are:

- Thomson Reuters citation indices (approximately the same content as the Thomson Reuters Web of Science)
- Elsevier Scopus

- Google Scholar
- Your organization's own publication database

The most basic forms of bibliometrics, as counting publications and citations, can be done in the online versions of the commercial databases Web of Science (WoS), Scopus and Google Scholar. The publication database of your own organization can only be used for publication counting, since it is not possible to build a citation matching and counting in a one-organization publication database. To do proper citation matching, a large portion of the world's scientific publication production is needed in the same database. The commercial vendors Thomson Reuters and Elsevier are adding around 1.5-1.8 million publication records per year to their systems, which is believed to be around two thirds of a roughly estimated yearly world-wide production of 2.5 million scholarly publications.

When it comes to more advanced bibliometrics, doing comparisons of citation counts to world-wide averages, even the commercial online services won't do the job, since they don't have any world averages to compare citation counts with. To be able to get world citation averages, you have to licence the data for the whole publication indices and build your own analysing system, usually covering about 20-30 million publication records. This is a procedure that involves large costs, both in licenses from the commercial vendors and in costs for personnel building and maintaining the database system, as well as computer hardware.

When you are about to decide which publications to include in a bibliometric analysis of an organization, you need some sort

of identifier that links publications to the organization. Your own organization's publication database usually has the advantage of internal unique ID's for your organization's organizational units and your organization's staff, so publication records may be selected based on those ID's. On the other hand, your publication database does not have any citation counts, so if you want to do citation-based bibliometrics, you need to get data from one of the commercial vendors.

In the commercial databases there are no unique identifiers for organizations or researchers, so the selection of publication records has to be based on error-prone text string matching of author and organization names. This less desirable method of record selection is the reason for the importance of keeping author and organization names unique and consistent over time. Trying to locate publication records for a department or research group using this text-based method is near to impossible, due to the large variation in naming of the organizational units, and frequent name-changes, mergers and splits of departments. In the section about publishing and promoting your research, further down, I will show the best way to state your and your organization's name for your entry in the author list.

Counting fractions of publications

When doing bibliometric studies on co-authored publications, publication and citation counts are often shared between the contributing parties. This is called *fractionalization* and can be based on either author names or addresses. The easiest and most common method when doing analyses of organizations is to do an address-based fractionalization. This is, for instance,

what the Swedish Research Council does when it analyses the output of Swedish research.

The address fractionalization means that if the researchers in your organization have one of four affiliation addresses in a publication, your organization will get attributed one fourth of the publication, regardless of the number of researchers that are affiliated with each of the addresses and regardless of which amount of work each researcher has put into the publication. The share of addresses is also often used as a weight when doing calculation of citation averages, so that publications where your organization addresses have a larger share will weigh heavier in the average calculation.

The methodology opposite to fractionalization is called *full* or *whole counting*, where each contributing organization or researcher gets full credit for the publication and all its citations. This method can on one hand be considered more “fair” to the researchers and the involved organizations, but has the disadvantage of the sum of the parts being larger than the whole. For instance, when doing full counting, the sum of publications from Swedish organizations will be almost twice the total Swedish publication production. What is counted here is not the number of publications, but rather the number of *authorships* or “affiliationships”.

It may also be noted that the new practice of assigning publications to all involved staff in a large research project creates severe adverse effects on bibliometric studies. For instance, some large research cooperatives in particle physics put over 2300 authors and 200-300 affiliations on each publication. If you fractionalize publication and citation counts,

almost nothing of this kind of publication will be visible in the assessment of an organization involved. If you on the other hand do full counting, such a publication can make a large difference in bibliometric indicators for each of the mentioned researchers and their respective organizations, even if the researcher may not even be aware that he or she has a part in the publication in question.

Research fields and average citations

Citation-based bibliometric indicators are based on the assumption that a reference (an outbound citation) from a scientific work to a previously published work represents an indication of scientific impact of the cited publication. It is also assumed that the number of (inbound) citations to a publication can act as a proxy to assess the impact of the scientific work of the author or the group that has produced the cited publication. This assumption does not always hold true at the *micro* level, i.e. for a single article, researcher or research group. There may be negative citations, claiming the cited author to be wrong or that the results are disputable and there are also a number of other reasons to cite a publication that can be considered less valid in relation to the assumption stated above. On the other hand, we also know that if we use bibliometrical methods on a large number of publications, like a thousand or more, we usually find a good correlation between citation-based indicators and a peer review of the work of the studied group.³ This means that the major part of the citations is to be considered as valid in relation to the bibliometric impact assumption. Thus, we can conclude that there is a good reason to believe that high scores in citation-based bibliometric

³ Moed, 2005

indicators are to be seen as a sign of high-impact research when working at the *macro* level.

Different research fields have different publication and citation cultures. In some fields, as for instance mathematics, the publication frequency is low and reference lists are short. In other fields, as for instance biotechnology, publication frequency is high and reference lists are long. This means that the *citation density* in the field of biotechnology will be much higher than the citation density in mathematics and that raw citation counts to publications from the two fields should not be compared without any precautions.

In the commercial databases Thomson Reuters Science Citation Index and Elsevier Scopus, the publications are classified into research subject fields. Thomson Reuters uses 250 field categories to classify each journal issue in 1-6 fields, and the classification of the publications is inherited from the classification of the journal issue they were published. When doing more advanced bibliometrics the classification of the journal issues are used to sort the publications into different research fields and compare the assessed publications only to publications within the same research field, due to differences in publication and citations frequencies between the fields. See Figure 1 for a picture of the differences in average citation rates between research fields.

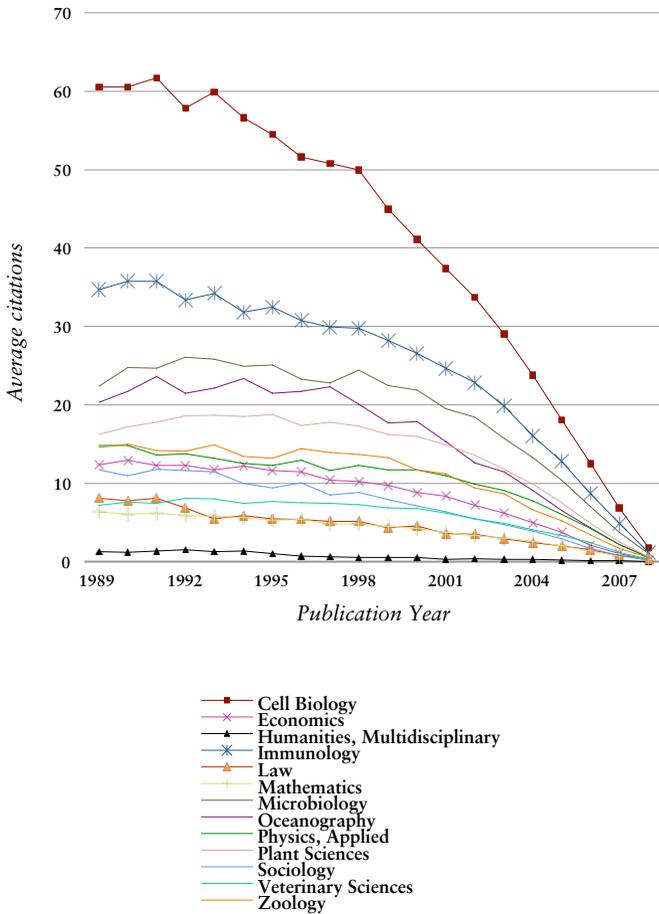


Figure 1. Average citation rate for publications in a number of research fields. Measurements were done in the Swedish Research Council's bibliometric system 2009.⁴

⁴ Kronman, Gunnarsson and Karlsson, 2010. An open citation window was used and self-citations were included. A self-citation is when a

The state-of-the-art bibliometric method to handle differences in citation densities between research fields is the *field normalized citation rate* (c_f).⁵ When calculating the field normalized citation rate, citation counts for publications are compared with the world average citation rate for publications of the same *type* and the same *publication year* and within the same *research field*. Dividing each publication's citation count with the world average citation count for publications of the same type, the same year, within the same research field, results in a normalized value. Using this normalization procedure, the world average within each combination of field, year and publication type will per definition be 1, and a field normalized citation rate value above 1 will indicate that a publication has been cited more than the average in the field.

When publications are measured using the field normalized citation rate it is thus the *journal* in which you publish that decides which field your publication will be compared to. If you publish in a journal classified in a low-cited field as mathematics, humanities or social sciences the citations your publication receives will end up having more weight in the field normalized citation rate indicator than if you publish in a journal that is classified in a field with a higher citation density.

researcher refers to her/his own previous publications in the reference list of an article. Certain data included herein are derived from the Web of Science ® prepared by THOMSON REUTERS ®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA: © Copyright THOMSON REUTERS ® 2010. All rights reserved.

⁵ The field normalized citation rate was introduced under the name *The Crown Indicator* by the bibliometric centre CWTS at the Leiden University in the middle of the 90's and refined and documented by the Swedish Research Council and Karolinska Institutet in the 00's.

There are examples of researchers who are active in multidisciplinary areas that link together computer programming and arts who will publish in journals classified in the arts field. The result of having articles about programming classified in the arts field will often be high field normalized citation indicators due to the relatively high citation rates among computer scientist compared to the low citation rates in the humanities.

Managing your assets: publish for maximum visibility and impact

Now that we know a bit more about bibliometrics and the ways publication records are being used to assess research volume and impact, it is time to take a look at the ways in which you can improve your bibliometric indicators and rankings. First, I will address the importance of choosing the right channel and the right publication type, and then I will address ways to make your publication more visible and influential in bibliometric studies.

Where to publish

The key to research impact, both for you and for your organization, is to make high-quality research and to reach the right audience with your research findings. Choosing the right channel – journal or publisher – for your publication can leverage its impact. Publishing in an international peer-reviewed journal with high impact, covered by the large indexing services, will usually render higher scores in bibliometric studies than publishing in another channel.

The channels with the most prominent outreach and impact in bibliometric studies are international journals covered by the indexing service Thomson Reuters Web of Science (WoS). Thomson Reuters indexes about 11 500 journals and add 1.6 - 1.7 million publication records to their database each year. The Thomson Reuters' indices are usually the main data source for bibliometric studies and therefore it is of vital importance to publish in a journal that is covered by them. If you have a choice when deciding which journal to publish in, consult the Thomson Reuters Master Journal List⁶ to see if you can find an appropriate journal that is indexed.

If you are publishing in a journal, the Thomson *Journal Impact Factor* (JIF) will give you an indication of the average number of citations to articles in the journal. The JIF for a journal is calculated by dividing the number of citations to a journal by the number of articles published in it.⁷ The JIF can be seen as a crude measure of how widely spread and how influential a journal is, and is therefore an indication of how much your article may be read and cited when published in the journal. Journal Impact Factors should not be compared between research fields, due to the differences in publication and citation rates between fields mentioned above, but within a field, the JIF can give you an indication of the most influential journals.⁸

⁶ Thomson Reuters, 2012

⁷ In practice, the Journal Impact Factor is not a clean quota, since some articles are considered "non-citable" and are removed from the denominator.

⁸ The Journal Impact Factor can be found in the Thomson Reuters system Journal Citation Reports.

In what form to publish

The type of publication you choose for disseminating your findings is of great importance for how the research will be assessed in bibliometric studies. Journal articles will usually give better scores than other types of publications such as conference proceedings, monographs and reports, due to the better coverage of journal articles in the bibliometric data sources. Below, I will outline the most common means of publications and what to take into account in respect to each channel.

Beginning with *journal articles*, and as mentioned above, Thomson Reuters primarily indexes about 11 500 international journals in WoS and Elsevier indexes 18 500 journals in Scopus. The reason for focusing on journals is that the journals are the most influential channels in the most fields, but also because journal articles tend to be easier to capture for indexing than other material due to stable titles with re-occurring issues and regular publishing patterns.

When doing bibliometric studies and counting citations, there is a significant difference between the average number of citations to a regular *original* article and a *review* article.⁹ Reviews receive on average 2.5 times the number of citations compared to an original article. This is of course due to the review being easier to digest and covering a broader view of the research field. Reviews get more readers and thus on average more citations. Another finding regarding citation counts is that articles that deal with methodology also tend to gather many citations, since everyone that utilizes the method afterwards will

⁹ This type of scientific review (“overview”) article should not be confused with “Book reviews”, common within the humanities and the social science, that in average reach little attention.

have to refer to the article where it was first presented. So writing reviews and methodology articles could both be considered as acceptable methods to boost citation counts for your research.

Another frequent form of publication within the academia is *conference proceedings*. In the databases and indices used for bibliometric studies the publication types “Article” and “Conference Proceeding” are being used and counted in quite different ways. Original research articles published in regular international journals are usually captured and indexed by the databases WoS and Elsevier Scopus. Conference publications, on the other hand, are a bit more problematic to gather and therefore conference proceedings are not covered by the databases to the same extent as regular articles.

If you do research in an area where conference proceedings are the primary vehicle for dissemination information, consider “repackaging” and republishing your material as an article, preferably in a journal indexed by WoS or Scopus. An article in a prestigious journal with a high journal impact factor will also usually make a better impression in the publication list of your CV.

In many research fields, *monographs* and *reports* are the primary vehicles for sharing research findings. When doing bibliometric studies based on the commercial data sources from Thomson Reuters and Elsevier these types of documents will not be counted, since they are not included in the indexes from these vendors. Bibliometric studies can be extended to include monographs and reports by using local data, such as the publication database of your organization, but currently there

are no methods to count citations to publications that are not covered by the commercial data suppliers.¹⁰ If you are doing research in a field where monographs and reports are of vital importance, the same advice as for conference proceedings apply; try to repackage and republish your findings as an article in a well-renowned journal covered by WoS or Scopus.

Choosing language

Journals with articles written in English is the core of WoS and Scopus, which means that articles in English will always be more influential in bibliometric studies. WoS and Scopus cover some journals in non-English languages but citation counts are usually low on articles in these journals, since the audience for these articles usually is smaller than for an English article.

If you primarily write in a non-English language for a domestic audience, the same repackaging and republishing recommendations as for conference proceedings and monographs apply. For instance, consider if your findings can be targeted at an international audience and republished as an article in an international journal. If you do research in a field where dissemination of results primarily is done via monographs in a national language, incentives for repackaging the result as an English article is of course twofold.

Using cooperation to increase visibility

Cooperation in research is important in many aspects, one of them being the aspect of the “marketing” contact area for the

¹⁰ Google Scholar supplies citation counts for other publication types than journal articles, but there is presently no method to gather these citation counts for batch computations and field normalization.

resulting publications. If more researchers are involved in the research and the publication process, the article will be exposed to a broader audience. Studies have shown that there is a correlation between the number of authors and the number of citations to an article, even if so called self-citations are excluded.¹¹

Figure 2. shows that the average number of citations to publications involving two researchers (7.8) is almost twice as much as the citation rate for single-author publications (4.2). The field normalised citation rate, adjusted for differences between research fields, also shows an increase in average citation rate (+20%) when going from one author to two.

A disclaimer may be in place here; not all cooperation is beneficial per se. As seen from the graphs above, the correlation between the number of authors and the citations start to decrease above six authors. If fractional counting is used when counting publications and citations, the correlation between the number of authors and indicator values will decrease. In addition, bringing in other researchers just to enhance the exposure of the finished publication may not be justified during the phases of actual research and writing.

¹¹ Aksnes, 2006

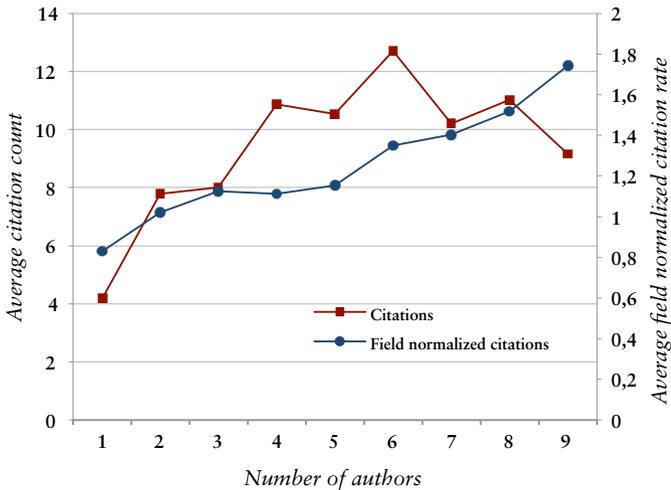


Figure 2. Correlation between the number of authors, the average number of citations and the average field normalized citation rate for publications from KTH Royal Institute of Technology.¹²

Another fair “trick” in the cooperation domain is to cooperate internationally. In bibliometric studies, publications produced as a result of international cooperation are usually seen to give higher citation counts, especially if you cooperate with researchers in countries and regions with high relative citation

¹² Citations are measured in Web of Science July 2011 on publications from KTH year 2005 and field normalized citations are calculated on KTH publications from 2005-2009 in the Karolinska Institutet bibliometric system. Both measures are done with open citation window and self-citations are included. Certain data included herein are derived from the Web of Science® prepared by THOMSON REUTERS®, Inc. (Thomson®), Philadelphia, Pennsylvania, USA: © Copyright THOMSON REUTERS® 2010. All rights reserved.

counts as the United States, United Kingdom or Switzerland. As mentioned before, sometimes international cooperation goes to the extreme as for the particle physicists working at the large hadron collider in CERN, where it is quite common to have around 2300 authors to each article. It is still unclear how bibliometrics should handle this type of publications. If authorship is fractionalized these articles' impact will be reduced to almost nil, if whole counting of authorships is done, they risk to skew the results due to the large and somewhat unfair impact for each researcher and organization involved.

When a researcher puts a reference to your work in her or his reference list, you get a citation and increased impact in bibliometric studies, but you also get increased visibility, since more researchers get aware of your work by studying the reference list of the referring work. This “advertising” effect can lead to more secondary citations from other publications. You can actually do the best to advertise your own work by referring to your own previous publications whenever this is appropriate. This is called a *self-citation* and in many bibliometric studies self-citations are removed, since they are not seen to represent impact in the rest of the scientific community. However, studies have shown that publications with more self-citations still get higher citation counts, even if the self-citations are removed,¹³ presumably by the advertising effect.

Making your publications traceable

A common problem when doing analyses of publications for researchers or research groups is the lack of unique author identifiers in the commercial bibliometric indices. Due to the

¹³ Aksnes, 2006

lack of unique identifiers for authors and affiliations, bibliometric analyses typically involve error-prone searches based on text string matching. To ensure that your publications are credited to you and to your organization it is therefore crucial that names and addresses are stated appropriately.

The names of the authors to the publications are being entered into the database indices in the way they appear in the journal, often just a family name followed by an initial of the given name. If you have a common name like John Smith or Maria Rodriguez, your name may end up like Smith, J and Rodriguez, M in the indices and there might be a lot of other researchers sharing these names. Therefore, the importance of having a unique and consistent author name should not be underestimated.

If you have a common name that you know you might share with other researchers, especially if they are within the same organization and/or field, consider creating a unique author “artist name” by adding an initial from for instance your middle name, for instance Anders Johan Andersson would become Andersson, A J. If you decide to make up a name like this, try to make the decision as early as possible in your research career and be sure to be consistent about its usage, otherwise you might end up having your publication records split up over several “authors” with slightly different names. This is a common problem, especially for researchers with double family names, which might end up with or without a hyphen between the family names or one of the family names interpreted as a given name. For instance, Jessica Wide Cederkvist might end up as author Wide Cederkvist, J; Wide-Cederkvist, J; or even Cederkvist, J W.

There are several initiatives trying to solve the problem with the lack of identifiers for authors, both among the commercial vendors of databases and vendor-independent “global” solutions. Thomson Reuters have their own initiative ResearcherID.com,¹⁴ where researchers can register and do housekeeping of their publication records in the WoS database. This is recommendable to do, especially if you know that your publication records in Web of Science are going to be used for an assessment of your research. Elsevier Scopus also have their own service for author identification, named SciVerse Author Identifier¹⁵ and Google Scholar is building a Google Scholar Citations service¹⁶ with the same purpose. There is also a vendor-neutral global initiative named ORCID – Open Researcher and Creator ID – that was launched in October 2012.

If you change your family name during your research career, it is especially important to make use of the vendors' system for author name unification to keep your publication records together. This is because there are yet no automatic methods other than a unique identifier to detect two different family names as belonging to the same researcher. There are examples of female researchers that keep their maiden family name as a researcher “artist” name after getting married, to keep their publication record together.

As mentioned above, the selection of data material used in bibliometric studies that utilize the commercial data sources is

¹⁴ ResearcherID, 2012

¹⁵ Elsevier, 2012

¹⁶ Google, 2012

usually based on error-prone text string searches. This means that if you want a publication to be credited to your organization, you need to write your organizational affiliation in a way that is easy to understand by an international audience and can be matched using computer-based methods.

Database vendors and other organizations collecting information about scientific publications usually expect author affiliations to be written according to a pattern going from larger organizational units to smaller, followed by city and country information:

Organization, Faculty, Department, Unit, City, Country

If you choose to write your affiliation using a form that starts with the name of your research lab or centre, it may happen that your main organization won't be identified and attributed, since its name will be buried further down in the address and maybe not detected by the system doing the publication selection. If you are affiliated with an organisation with a non-English name, also check that you are using the proper English name of your organization, rather than trying to guess.

If you do research in a very large collaborating team, make sure that the main author of the publication at least gets information about the proper English name of your organization and the country information to put in the address list:

Your organization, City, Country

It is interesting to notice the changing role of the address here. In a publication economy, the function of the address is

changed from that of a postal address to an organizational affiliation. There are still researchers who believe that it is important to put the street name and the zip code in the address. To what use? Do you expect people to write letters to you, so you need the mailman to find his way? In the publication economy, the main purpose of the address is the identification of the right organization to credit the publication to.

Making your publications accessible

The world of scholarly publishing is right now going through a transition where the old paper-based reader-pays subscription model is replaced with a new more internet-savvy producer-pays model. This means that journals are beginning to cover the costs for peer review and publishing with a fee from the publishing researcher or her/his organization or funding agency, or by being a part of a publishing-funding learned organization. When the cost of publishing is moved from the reader to the producer, articles can be published on the Internet free for all to read without any barriers as subscriptions or tolls and that is why this new publishing model has been named *Open Access*.

Another way to make the content of the publications freely available to the public is to do *self-archiving* of articles that have been published in a subscription-based journal. The publishers usually gives authors the right to publish the reviewed and accepted last manuscript before publication in an institutional repository, sometimes after an *embargo period* of six to twelve months or even longer after publication. This is called *post-print* self-archiving. The conditions for self-archiving

and the length of the embargo periods for various publishers can be checked at the online service SHERPA/RoMEO.¹⁷¹⁸

There are a number of reasons why you should try to get your publications freely available on the Internet:

- It improves the speed and efficiency of research, and also enables interdisciplinary research.
- Your publication will be more visible in the international search engines and may be found and read by a broader audience.
- Studies show that articles published for free access on the Internet gain more citations.¹⁹
- You have to publish your findings as Open Access if you have funding from a body that mandates it.
- Your organization may have a policy for scientific publishing that mandates you to publish your results in Open Access journals or in the organization repository.

While talking about Open Access publishing, a final word of warning may be in place. In the turmoil of the transition of scholarly publishing, a new breed of non-serious, so-called “predatory” publishers with poor or non-existent peer review are entering the scene. If you get invited to publish in a journal that will charge you for accepting your manuscript, check for signals of non-seriousness as spamming e-mails, an amateurish

¹⁷ University of Nottingham, 2012

¹⁸ The conditions presently seem to be in a constant flux, so it is safest to do a final check at the website of the publisher or the contract you signed before publishing.

¹⁹ Eysenbach, 2006, Hitchcock, 2012

looking website, a non-existent postal address, poor contact details, etc.²⁰

Another way to make your work more influential is to publish the underlying research data for public re-use. There are studies showing a correlation between public research data and the number of citations to the publication(s) based on the data.^{21 22} However, it is still unclear whether there is a *causal* connection between the publication of data and the increased number of citations or if the correlation is caused by some other related parameter as the funding or the number of researchers involved in the study. On the other hand, no one has so far shown a negative correlation between published data and the number of citations.

Using social media for increased visibility

In today's digital age, the old saying “publish or perish” can be augmented with a more modern counterpart “get visible or vanish”.²³ Besides making publications and data public, developing a comprehensive online presence can leverage the impact of your research. Using online services as Twitter, Facebook, Google Plus, LinkedIn, Mendeley or Zotero can make your research visible to a larger audience and create a debate around your work.^{24 25} Establishing a blog focusing on your research is even better, especially in combination with the

²⁰ Beall, 2012

²¹ Dorch, 2012

²² Sears, 2011

²³ Science Online, 2012

²⁴ Mendeley, 2012

²⁵ Zotero, 2012

other social media tools.

Ensuring findability and preservation

If you want to reach out with your research results and gain impact, it is important that your publications are searchable in the global search engines on the Internet and also preserved for future reference. This is where the publication database of your organization - the *institutional repository* - can play an important role. Publishing in a *subject repository* can also increase the findability and preservation of your work.

In the publication economy of today, most research organizations run a publication database where information about the publications of its researchers is stored. The primary content of the publication database is not publications as such, but *metadata* records with information about the publications and it is used to market and keep track of the output of the researchers in the organization. These publication records are often used as a basis for bibliometric studies.

Publication databases are often used for the following purposes:

- to generate publication lists on web pages for departments, research groups and individual researchers
- to generate publication lists for CV's and project applications
- to visualize and market research results from the organization
- as a source for bibliometric analyses

- to make research output more visible to search engines as Google and Google Scholar

In many cases, records can be imported to the publication database from commercial databases as WoS and Scopus. This is often done by the staff at the university library. However, if you have produced publications as monographs, reports and conference proceedings papers that not are indexed in the commercial databases, you usually have to register them manually yourself or get someone to do it for you.

The publication databases are often extended into institutional *repositories*, which can cater for the *full text* of publications, usually as PDF files, besides the metadata records needed for marketing and bibliometrics. This is where your organization's repository comes in handy for disseminating your publication in full text, doing self-archiving, as previously mentioned in the section about Open Access. If you find that you have the right to do self-archiving of your manuscript, do this to increase its visibility and impact.

Subject based repositories give you an opportunity to increase the effectiveness of your reputation building by giving your research and early visibility and allowing your researcher community to cooperate more efficiently. The fast dissemination to your peers is crucial when it comes to impact and citations. There are studies that show that manuscripts published in subject repositories as *arXiv* "... yields a citation advantage of a factor five."²⁶

²⁶ Gentil-Beccot and Mele, 2012

If you care about the number of citations to your work, you should try to refer from the pre-print to the finally published article, since several versions of an article in different locations can lead to a phenomenon that is known as *citation fragmentation*. Fragmentation occurs when each variant of your publication captures only a portion of the citations the unified publication would get. The effect of citation fragmentation can often be spotted in Google Scholar, where you may find several incarnations the same publication with different citation counts. In Google Scholar, registering for an account and bringing the variants together to one single record can alleviate this phenomenon. This is not possible to do for the other bibliometric database vendors, since they will only index the journal variant of the publication, and then the citations referring to the pre-prints and any other variants of the publication in repositories will get lost.

Discussion: why an emerging publication economy?

Why have publication records and bibliometrics started to be of such importance that we now even are inclined to call it a publication economy?²⁷ The underlying reasons can probably be spelled globalization and tightening competition for resources and knowledge around the world, together with a historic development of society. Have we perhaps not only transcended the farming era and the industrial era, but also the newly celebrated information era and are now heading into a new *era of knowledge*?²⁸ One in which education and research is the *industry* and the higher education and research institutions are the *factories* that produce this knowledge?

²⁷ Larsson, 2009

²⁸ See: Castells, 2000

Research once used to be reserved for an academic elite, consisting of a few wealthy aristocrats that could support themselves while getting educated and producing science. At that time, research did not put any large expenses on the society and the scholars could therefore have a large degree of freedom in their research. Today, on the other hand, a large portion of the population goes to university and higher education and research is a major financial undertaking for the society. If you are putting a lot of resources into the production of something, don't you then want to be in control over what you get in return for your invested money? At least this is what industrial managers have been doing for over a century now, running business intelligence systems with statistics on their production.

But what is the output of a knowledge production? Knowledge is a much more esoteric and multi-faceted product than, for instance, cars, refrigerators, computer programs or civil services. What should be measured if we want to assess the results of a knowledge production? In the urging need for something to measure, governments and university managements turn to what *can* be measured, rather than what *should* be measured, since no one seems to know the answer to the latter question. Publications and citations are some of the few measurable results of a knowledge production, and that is why they so frequently are being used to assess the return on investment in research. Governments and university managements seem to be acting a bit like in the old joke about the man who lost his car keys one night and started looking for them, not where he lost them, but beneath the lamppost, because there it was light so he could see.

Using bibliometrics to assess research is both right and wrong at the same time. On one hand, there is a legitimate reason to try to find measures on return of investment in research. On the other hand, many more factors should be taken into account when doing the measuring.

To further complicate the picture for the assessment of science, the output of the scholarly society is not a static product, which can be measured without intervening with the production process. In other words, if we start to measure research in certain ways and allocate funding according to the results, researchers will adapt to this and the measurements will start to be an incitement, driving research in directions towards the measurable. To quote a recent critical article: “Metrics of quantity once were the means to assess the performance of researchers, but now they have become an end in their own right.”²⁹

Here, I would like to issue a call for help from scholars in various disciplines. If you know that governments and university managements want to measure and put numbers on the results of your research; which measures should be used to make the right assessments and drive the research in the right direction? You are the ones that should know, and if you don't help, your research will only be measured with inferior, one-dimensional tools as bibliometrics.

Looking from this economic perspective, we can see how the societal role of the scholar is changing over time. What used to be an economically independent scholar with freedom to do research driven by curiosity is now a worker in a production

²⁹ Fischer, Ritchie and Hanspach, 2012

machine for the knowledge society. In the short run it is off course good that the society doesn't spend money on research that don't give any apparent benefits in return. But how do we know in the end what benefits are to be gained from which research? If researchers only focus on delivering short-term accountable results and managing their publication assets, what will happen with the long-term basic research that may deliver results that are important in 20-30 years?

There is the school of old academics that claims that researchers should be given funding and then left alone to do their research in peace. It is a bit like the sayings of a famous entrepreneur nearly a century ago: How do you run a successful company? You hire competent and talented people and leave them alone to do their job as they think best fit.

But how do we know which researchers are talented and should be recruited and get this safe long-term financing? And how many of them should be financed? And are we not re-building the old academic ivory tower with an elevated elite of untouchables if we do so?

I would like to end this discussion on the publication economy with a quote from a recent critical letter in the journal *Trends in Ecology and Evolution*:³⁰

The modern mantra of quantity is taking a heavy toll on two prerequisites for generating wisdom: creativity and reflection.

³⁰ Fischer et al., 2012

References

- Aksnes, Dag. "A macro study of self-citation." *Scientometrics*, Vol. 56, no. 2, 2006, pp. 235–246. DOI: 10.1023/A:1021919228368.
- Beall, Jeffrey. "Criteria for Determining Predatory Open-Access Publishers.", 2012. <http://scholarlyoa.com/2012/08/04/criteria-for-determining-predatory-open-access-publishers/> [Retrieved 10 August 2012].
- Castells, Manuel. *The Information Age : Economy, Society and Culture. The rise of the network society*. [2], Malden Mass., Blackwell, 2000.
- Dorch, Bertil. "On the Citation Advantage of Linking to Data.", 2012. <http://hprints.org/hprints-00714715/> [Retrieved 28 July 2012].
- Elsevier. "Author Identifier - SciVerse.", 2012. <http://www.info.sciverse.com/scopus/scopus-in-detail/tools/authoridentifier> [Retrieved 5 August 2012].
- Eysenbach, Gunther. "Citation advantage of open access articles." *PLoS Biology*, Vol. 4, no. 5, 2006, pp. e157. DOI: 10.1371/journal.pbio.0040157.
- Fischer, Joern, Ritchie, Euan G. and Hanspach, Jan. "Academia's obsession with quantity." *Trends in Ecology & Evolution*, Vol. 27, no. 9, 2012, pp. 473–474. DOI: 10.1016/j.tree.2012.05.010.
- Gentil-Beccot, Anne and Mele, Salvatore. "Citing and Reading Behaviours in High-Energy Physics. How a Community Stopped Worrying about Journals and Learned to Love

- Repositories.", 2012. <http://arxiv.org/abs/0906.5418>
[Retrieved 5 August 2012]
- Google. "Google Scholar Citations.", 2012.
<http://scholar.google.com/intl/en/scholar/citations.html>
[Retrieved 5 August 2012].
- Hitchcock, Steve. "The Effect of Open Access and Downloads
(‘Hits’) on Citation Impact: a Bibliography of Studies.",
2012. <http://opcit.eprints.org/oacitation-biblio.html>
[Retrieved 25 June 2012].
- Karolinska Institutet. "The Bibliometrics Group’s publishing
recommendations for research scientists at Karolinska
Institutet.", 2011.
http://kib.ki.se/sites/kib.ki.se/files/FS_rekommendationer_till_publicerande_forskare_en.pdf [Retrieved 7 June 2012].
- Kronman, Ulf, Gunnarsson, Magnus and Karlsson, Staffan.
"The Bibliometric Database at the Swedish Research
Council: Contents, Methods and Indicators.", 2010.
<http://vr.se/bibliometrics> [Retrieved 27 July 2012].
- Larsson, Staffan. "An emerging economy of publications and
citations." *Nordisk Pedagogik*, Vol. 29, 2009, pp. 34–52.
- Mendeley. "Mendeley: Free Reference Manager and PDF
Organizer.", 2012. <http://www.mendeley.com/> [Retrieved 8
August 2012].
- Moed, Henk. *Citation Analysis in Research Evaluation*. New
York, Springer, 2005.
- ResearcherID. "ResearcherID.com.", 2012.
<http://www.researcherid.com/> [Retrieved 5 August 2012].
- Ruhvargers, Andrejs. "Global University Rankings and Their
Impact". European University Association, 2011.

http://www.eua.be/pubs/Global_University_Rankings_and_Their_Impact.pdf [Retrieved 26 July 2012].

Science Online. "Science Online 2012: Day Two Blitz - Get Visible or Vanish.", 2012.

http://www.youtube.com/watch?v=fs4mZrxl7gg&feature=youtuube_gdata_player [Retrieved 30 July 2012].

Sears, J.R.L. "Data Sharing Effect on Article Citation Rate in Paleoceanography." 2011 Fall Meeting, AGU, San Francisco, Calif., 2011. <http://www.agu.org/cgi-bin/wais?tt=IN53B-1628> [Retrieved 20 June 2012].

Thomson Reuters. "Master Journal List - Thomson Reuters.", 2012. <http://ip-science.thomsonreuters.com/mjl/> [Retrieved 5 August 2012].

University of Nottingham. "SHERPA/RoMEO - Publisher copyright policies & self-archiving.", 2012. <http://www.sherpa.ac.uk/romeo/> [Retrieved 7 August 2012].

Zotero. "Zotero.", 2012. <http://www.zotero.org/> [Retrieved 8 August 2012].

Ulf Kronman is a physicist who has pursued doctoral studies in musical acoustics at the KTH Royal Institute of Technology. He has worked as a bibliometric analyst at the Karolinska Institutet, The KTH Royal Institute of Technology and the Swedish Research Council. He is currently the coordinator of the programme OpenAccess.se at the National Library of Sweden.