Virt&I-Comm.3.2012.24

RESEARCH EVALUATION AND ACADEMIC QUALIFICATION: AN APPLICATION TO CHEMISTRY

Antonio Laganà, Department of Chemistry, University of Perugia, Italy

Abstract

In this paper the general problem of handling data necessary in discipline specific evaluations for academic qualification is examined in the context of knowledge management systems and related active and passive filtering techniques. The different active and passive filtering parameters on which the evaluation criteria adopted in the on-going Italian full professor qualification procedures are based is examined by referring to a chemistry sector. A merging of active and passive filtering procedures into a active-passive one relying on collaborative interventions of candidates and peer reviewers is suggested.

KNOWLEDGE MANAGEMENT AND FILTERING TECHNIQUES

By definition, research is a difficult activity to categorize and evaluate. There are not standard rules for carrying out research and for working out its evaluation because, by serendipity, one might obtain high quality unexpected products even without adopting a proper scientific method. On the other hand one may obtain no products at all even from high level research. However, the more indispensable research and the resulting innovation are becoming for our society and its economic development [1], the more rigorous the implementation of objective procedures for ranking and evaluating scientific achievements and qualifications become.

The handling of research data is a typical Knowledge Management (KM) activity in which the necessary information is collected into a data base and a conceptual platform is utilized in order to automate a ranking process. In other words, KM builds up its procedures on a set of practices able to represent, identify, create and elaborate knowledge. Within this, the management of scientific knowledge is, in particular, a specific challenge because of the large variety of science subjects and specializations.

A first crucial aspect of this application of KM is the volume of data handled that may amount to the order of TBs and PBs even for simple cases because the sets of scientific data are themselves increasingly numerous (and almost always geographically distributed). Accordingly, the application of modern KM may require the appropriate automatic management of large caches, transfer of data over wide area networks, their coordination in databases, related elaborations for comparison/evaluation and other advanced techniques as well [2].

A second crucial aspect of KM is the standardization of data and of their storage procedures. In this respect, it is of fundamental importance to arrange, whenever possible, a common framework able to gather several kinds of information coming from different sources in a proper way in order to progress with a rigorous evaluation of data quality. At the same time, owing to the fact in that various tools (e.g. meta schedulers) can be used, a data standardization appears to be necessary so as to allow the mentioned tools communicate each other by sharing information and functions.

The last crucial aspect of KM for carrying out evaluations is the structuring (or mapping) of collected data into a space more suited for ranking using appropriate functions like the filtering ones. Filtering techniques are sufficiently simple automatic tools singling out

ISSN: 2279-8773

information and patterns embedded in collected data [3]. Passive Filtering (PF) techniques elaborate predictions out of people user behaviours counting on the fact that if one liked (or disliked) a service in the past he/she will continue to do so in the future. In practice this means that in PF users of a service leave (not necessarily consciously) a sign of their preference in similar ways (natural behaviour). In practice, PF software:
- looks for the number and opinion of people sharing the same rating patterns
- makes use of the frequency of the opinions of like-minded people to set a ranking.
PF methods are usually based on a large sample of data and are, accordingly, very stable and variants of the method are mainly based on the different technologies used for monitoring the just mentioned natural behaviour.

As an alternative, one can use Active Filtering (AF). AF is based on the fact that peers explicitly intend to share their opinions and offer them (possibly in an agreed shared version) for the evaluation. There are some advantages in using AF. One of them is the fact that the explicitly formulated opinion of the peers could be of interest for the evaluated subjects (whose behaviour could be modified accordingly) and useful for the development of better metrics. Yet, they show the disadvantage of relying on a number of collected opinions definitely smaller than those of PF techniques and, accordingly, less reproducible and stable. For example, the opinion expressed by the peers could be biased (this is a frequent complaint) and conflicting if no appropriate metrics are adopted and publicized before starting the evaluation.

EVALUATION PROCEDURES AND PARAMETERS

Typically, research work develops along the following procedural scheme:

a) undertake preliminary systematic studies of the problem
b) plan related theoretical and/or experimental and/or computational activities
c) carry out planned activities
d) rationalize collected results and if necessary iterate on steps b, c and d
e) assemble the research product

The most measurable and verifiable step of such procedure is the last one that in general requires also writing reports and papers (after all, even an unsuccessful research can be an important source of knowledge worth being reported). It is, therefore, not surprising that in research evaluation much emphasis has been given to objective bibliometric parameters [4] (bibliometric parameters are a **set of mathematical and statistical indicators measuring the quantity and quality of publications)** for which several AF and PF procedures (and mixed ones) have become of popular utilization**.**

There are three types of bibliometric indicators:

1. quantity indicators, which measure the productivity of a particular researcher;

2. quality indicators, which measure the quality (or "performance") of a researcher's output;

3. structural indicators, which measure connections between publications, authors, and areas of research.

The followings are the more popular bibliometric parametrs

NUMBER OF PUBLICATIONS

The number of scientific publications produced by the analyzed unit during the analyzed time span. While the acceptance for publication of each single paper is on an AF process based on the peer evaluation of the publications [4] (at least for the internationally quoted journals) the counting of the number of published paper is to be considered a PF process expressing the acceptation of an author by a scientific community. As a matter of fact, different parameters are referred to different sources (sometimes even selected by degree of specialization, by field, by level of ranking of the journal, by the bibliographic database, etc.) and are often presented separately per document type. Verified sets of data are taken from publication sources like ISI [5], CWTS [6], top journals, CEST [7] world share of publications, CEST degree of specialization, CEST relative activity index.

NUMBER OF CITATIONS

The (total, average per publication, CWTS field normalized, logarithm based, limited to top ranking, to relative audience, averaged, journal normalized) number of citations to articles published by an analyzed unit during the analyzed time span. This is also a PF method that sums the citation values of all articles published by the analyzed unit during the analyzed time span. A particular type of citation parameter is the H-index that is the number of publications (H), attributed to the analyzed unit during the analyzed time span, that have at least H citations and some of its version for age correction.

COLLABORATIVE PUBLICATIONS

Another PF set of indicators is the quantification of to what extent an analyzed unit has cooperated with other units in the production of articles (International collaboration – share of publications with co-authors from organizations in at least two different countries; National collaboration – share of publications with co-authors from at least two organizations within the same country; Department collaboration – share of publications with co-authors from at least two departments within the same organization; Unit collaboration - share of publications with co-authors from two or more research units). For this parameter the share of articles co-authored between the units is accounted for by dividing the number of articles published by the analyzed unit during the analyzed time span by the number of papers co-authored together with a selected other unit. For internal collaboration the number of paper is divided by the number of co-authors. Often a particular emphasis is given (if the author list is not strictly alphabetical) to the first author or to the corresponding one.

IMPACT FACTOR

The ISI impact factor is, indeed, another PF filtering indicator whose value corresponds to the average number of citations a publication in a specific journal has received during the two years following the year of publication. The ISI impact factor for a specific journal and one specific year is calculated by counting the number of citations to articles in that journal in the two preceding years (with respect to the publication one) divided by the number of publications defined by Thomson ISI as “citeable” in that journal in the two preceding years.
A more advanced journal impact factor than the ISI one takes both journal subject areas and document types into consideration. This makes comparison possible between journals in different subject areas. Other improvements are the extension of the period of measurement to more years (for instance 5) and the inclusion among “citable publications” of documents of type “letter” (this allows to have the same publication types in both the numerator and the denominator).

ISSN: 2279-8773

COMBINING AF AND PF PARAMETERS

As is apparent from the above list, there is a great deal of variety among the methods used for defining, collecting and using PF bibliometric parameters for research ranking. For example, a parameter like the H index, being a field specific cumulative index, does not encourage work in niche fields and/or changes of field. Moreover it is not linear with the growth of scientific production and, in some cases, is even further scaled by the age of the researcher. Other choices are related to the inclusion or exclusion of self citations, the fractionalization or any other form of weighting publications between the contributing authors which discourage collaborations. Moreover, these PF bibliometric parameters can be further criticized because the databases or quoted references are insufficient (if not incorrect) and this may induce a bias in the metrics (fostering, for example, short term rather than long term research or quantity over quality) and, after all, influence the behaviour of the researchers and distort the effect of the filtering action.

This can be mitigated by corrective actions undertaken with the participation of people to be evaluated. The validity of most of the indicators improves, in fact, if the authors themselves validate or supply information about their publications before the indicator values are calculated. This practice is already partially implemented (as we shall see later for the specific case of the Italian call for Professors qualifications) by combining active and passive filtering techniques and inviting the candidates to help with AH one. This is indeed the case in which researchers are invited to properly tailor their curriculum. A proper representation of the scientific career, of the recognitions within the scientific community of pertinence, the ability of transferring scientific achievements to society, the fundamentality of the research carried out, etc.. And this is, probably, the best countermeasure to the bias of some bibliometric parameters and leaves no room to the temptation of having no rules (and therefore to arbitrariness) that is sometimes "justified" with the pretext of preserving the autonomy and the dignity of the evaluation committee..

As already mentioned, the usual way of complement PF indicators is to combine them with the AF ones in which a self assembled curriculum of a researcher is checked and evaluated by a committee of peers selected among recognised "experts" of the scientific community. Among the curriculum items particularly considered in that respect are invitations to deliver lectures and talks, chairing conference sessions and workshops, acting as reviewer, member or director of the editorial board of a journal or magazine, supervising PhD students, coordinating a research group, a funded research project following a competitive call, chairing a learned scientific society, organizing a conference, acting as technology transfer agent, bearing patents, owning spin-offs, etc. which are difficult to automatize.

THE 2012 ITALIAN CALL FOR PROFESSORSHIP QUALIFICATION

Following the experience gained in the COST CMST domain for evaluating European cooperative projects and in peer reviewing for other international agencies and calls (like the NSF, EC and ECTNA ones) the author has contributed to setting evaluation criteria for the recent Italian public Qualification selections for Professors. The latter has involved full time the analysis of the activities of more than 400 researchers. The key guidelines of the Qualification criteria dictated by the Ministry are (we report below a free partial translation of criteria published on the Ministerial MIUR site for full professor qualification for sector 03/B1):

"PREAMBLE

The overall evaluation of the candidate has to be grounded on the analysis of the value of his/her scientific production. The matching of the bibliometric specific parameters is neither a necessary or

a sufficient condition for the achievement of the qualification. As a general rule, the qualification can be obtained only by candidates fulfilling both conditions (scientific merit evaluation and bibliometric parameters). Notwithstanding, the committee can deviate from such rule. This means that the committee could either not pass candidates overcoming the prescribed bibliometric parameters (yet not qualifying for a positive scientific merit) or pass candidates not overcoming the prescribed bibliometric parameters (yet qualifying for an extremely positive scientific merit).

GENERAL CRITERIA FOR EVALUATING SCIENTIFIC ARTICLES AND OTHER TITLES:

the evaluation is aimed at assessing the full scientific maturity of the candidate as testified both by the importance of the scientific subjects investigated and by the high quality and originality of the achieved results placing the candidate in a recognised position in the international research context. The evaluations is also aimed at testifying:

• ability to direct a research group characterized by international level collaborations
• experience in supervising doctoral students
• capacity to attract competitive research funds as coordinator at international level and as promoter of technological transfer

ENCLOSED SCIENTIFIC PUBLICATIONS:

Criteria:

a) coherence with the subjects of the disciplinary and interdisciplinary competition
b) personal contribution to the collaborative publications
c) quality of the scientific production seen as part of the international research on the ground of originality, methodological rigor and innovative character using merit classifications (excellent, good, acceptable, limited) as from annex D of D.M. 76/2012
d) editorial placement of science products of national or international level using standard peer review procedures

Parameters:

a) number and type of presented publications and their distribution along time with particular reference to the five consecutive solar years preceding the date of presentation of the application (keeping into account the periods of maternity leave or other periods of leave considered by the law and differing from study periods)
b) impact of the publications within the competition sector (keeping in mind of the academic age and, where necessary, of the specific characteristics of the scientific sector or part of it)

Possibility of deviating from above mentioned criteria/parameters including the use of possible criteria/parameters different or integrative (art. 3, comma 3, and art. 6, paragraph 5, DM 76/2012):

• The committee believes important to integrate the above mentioned criteria/parameters adding elements of evaluation specifically related to the overall scientific production of the candidate in the last five years using particularly the Impact Factor of the journals.

ISSN: 2279-8773

• Another element to be taken into account in addition to the other factors affecting parameter a) is the possibility of delaying submission of one or more scientific paper because of the deposit of a patent.
• As to parameter a) the committee, aware of the long times intervening between the manuscript acceptance and its effective publication, will consider also scientific papers not yet available in print yet bearing a DOI.

TITLES OTHER THAN PUBLICATIONS:

Parameters:

a) impact of the overall scientific production in the 10 years preceding the publication of the call up to the date of application evaluated according the indicators (median) of which at the article 6 and annexes A (bibliometric) and B (non bibliometric) of D.M. 76/2012
b) scientific responsibilities of international and national accepted for funding in competitive calls based on peer review
c) direction of journals, magazines, editorial series, encyclopaedias and prestigious reviews
d) participation to editorial committees of journals, magazines, editorial series, encyclopaedias and prestigious reviews
e) assignment of official fellowship and teaching duties at high qualification foreign and international universities and research institutions
f) direction of bodies and research institutions of high qualification
g) participation to prestigious academies of the area
h) achievements of prizes and awards for the scientific activities
i) results obtained for technological transfer in the proper scientific areas through technological transfer in terms of participation to spin off, development, utilization and commercialization of patents
l) possession of titles allowing a better definition of the scientific profile of the candidate as set by the committee (art. 3, paragraph 3, and art. 4, paragraph 4, letter l), DM 76/2012):
m)
• be "corresponding author" of articles published in the last 10 years.

• Documented experience of the candidate in managing complex organization either in academy or in research.

• A series of invitation to deliver talks at international conferences at Universities of research institutions of high international level.

• Supervision of PhD students or participation to evaluation committees or external referee for PhD candidates at Universities of international levels.

• Ability to attract PhD students or Post Doctors even through international competitions (e.g. Marie Curie fellowship);

• Scientific citations of excellence, like the "highlights", journals covers and any other element that can testify the scientific impact of the products of the candidate

• be referee of national or international agencies.

ISSN: 2279-8773

Possibility of deviating from above mentioned criteria/parameters including the use of possible criteria/parameters different or integrative (art. 3, paragraph 3, and art. 6, paragraph 5, DM 76/2012)

• The committee will use the indicators (medians) of Art. 6 of DM 76/2012 as from the basic criterion of the qualification according to what is suggested in Circular n. 754 of January 11, 2013.

In particular the Committee shall consider:

1) satisfactory to the end of the qualification the overcoming of 3 medians out of three after the obtainment of additional 4 points deriving from the overall evaluation of the publications and titles other than publications as specified below;

2) satisfactory to the end of the qualification the overcoming of 2 medians out of three after the obtainment of additional 8 points deriving from the overall evaluation of the publications and titles other than publications as specified below;

As mentioned by Circular n. 754 of January 11, 2013, the Committee shall consider also candidates overcoming only one or none of the three medians, as specified in the next paragraph.

TYPE OF PUBLICATIONS AND PONDERING OF CRITERIA AND PARAMETERS (art 3, comma 3, DM 76/2012):

Scientific publications:

The committee will evaluate first of all the scientific publications enclosed by the candidates for a maximum number of 20 using the above mentioned criteria and parameters. As to the type of publications will be considered only articles and reviews bearing bibliometric index (such index can in fact be considered as a means of pondering in agreement with what is given in annex D of DM 76/2012). It will be considered as suitable for being proposed for qualification only the candidates obtaining a positive evaluation by the committee on the ground of the already mentioned criteria shortly listed herewith: a) coherence with subjects of the competition sector or with pertinent interdisciplinary subjects, b) individual contribution if collaborative publications, c) quality of the scientific production, d) editorial level of the scientific products.

The committee, in addition to the enclosed publications will take into consideration those of the last 5 years and will assign 5 points to those with the highest Impact Factor (IF) as specified below:

1. An average of the IFs larger than 6 corresponds to "excellent" (10 points);

2. An average of the IFs larger than 3.5 and lower than 6 corresponds to "good" (5 points);

3. An average of the IFs larger than 1.5 and lower than 1.5 corresponds to "acceptable" (2 points);

4. An average of the IFs smaller than 1.5 corresponds “limited” (0 points).

The source of the Impact factor will be the Journal of Citation Reports and related values will be those of the year 2011, independently from the year of publication.

TITLES (DIFFERING FROM PUBLICATIONS)

ISSN: 2279-8773

For the parameter of letters going from b) to l8) the committee will assign the following points:

to the parameter of letter l1) a maximum of 10 points will be assigned according to the following scheme:

be "corresponding author" in at least 20 publications corresponds to "excellent" (10 points);

be "corresponding author" in at least 10 publications corresponds to "good" (5 points);

be "corresponding author" in at least 5 publications corresponds to "acceptable" (2 points);

be "corresponding author" in less than 5 publications corresponds to "limited" (0 points).

to the parameters of the letters b), c), e), f), l2) and l3) will be assigned a maximum of 4 points;

to the parameters of the letters d), h), i), l4)-l7) will be assigned a maximum of 2 points;

to the parameters of the letters g) and l8) will be assigned a maximum of 1 point.

Following a positive evaluation of the enclosed scientific publications, the Committee shall consider the case of candidates surmounting 3 medians of parameter a) satisfactory for the qualification if additional 4 points are obtained from the overall evaluation of the publication and titles other than publications.

Following a positive evaluation of the enclosed scientific publications, the Committee shall consider the case of candidates surmounting 2 medians of parameter a) satisfactory for the qualification if additional 8 points are obtained from the overall evaluation of the publication and titles other than publications.

Following a positive evaluation of the enclosed scientific publications, the Committee shall consider also the candidates surmounting only one or none of the medians as foreseen in the Circular n. 754 of January 11, 2013. In these cases a candidate will be considered suitable for being proposed for qualification only if in the overall evaluation of the publications and titles other than publications gathers a number of points at least even to the sum of the per cent deviations from the not overcome and in any case not lower than 15 points."

SOME CONCLUDING COMMENTS

To the end of carrying out a proper analysis of the effectiveness of the procedure one should wait for the outcomes of the qualification process. At present, in fact, not all the committees have completed their work (the present deadline is the end of September) and, therefore, results are not available. However, a critical reading of the structure and founding elements of the adopted procedure can be already made.

A first element of consideration is the source of data. For sure taking as a data source ISI or Scopus is an element of guarantee. However, the candidates should be allowed to see the outcome of the search made and to integrate and further validate it. In this respect the use of a more comprehensive data source (such as scholar.google.it) on which the candidate clicks to select the appropriate ones for the specific application. This is important for the first median of the Italian competition. In addition even the possibility of correcting misquotations in the papers quoting the candidate work

would be desirable for the appropriate definition of both the contemporary H index and the average number of per paper citations.

Another important indication is the adoption of a more structured approach to the curriculum information. It would be useful, for example, to insert the data for the curriculum in a verifiable data base format. For example information for the IF of selected best papers and the first or corresponding author could be input/corrected in such way. Even other data like that on (coordination of or participation to) international and national funded projects, direction of journals, magazines, editorial series, encyclopaedias and prestigious reviews, participation to editorial committees of journals, magazines, editorial series, encyclopaedias and prestigious reviews, assignment of official fellowship and teaching duties at high qualification foreign and international universities and research institutions, direction of bodies and research institutions of high qualification, participation to prestigious academies of the area, achievements of prizes and awards for the scientific activities, results obtained for technological transfer in the proper scientific areas through technological transfer in terms of participation to spin off, development, utilization and commercialization of patents, possession of titles allowing a better definition of the scientific profile of the candidate.

To focus directly with the foundations of the undergoing process of "qualification for professorship" (presently being carried out by the Italian scientific community, the massive and intensive nature of such evaluation has singled out some peculiar aspects of the techniques adopted for quantifying the evaluation of the outcomes of research activities to the end of singling out the elements of a qualification to professorship. Within the limits of the selected medians (whose validity is not dealt with in the paper although a serious discussion of this matter is in order) the process shows positive aspects mainly for its intention of introducing objective parameters and for the possibility of compensating (within a certain per cent) PF bibliometric parameters with AF curricular ones. This suggests further work once the results of the qualification process will be made available along an increase in number and a more appropriate choice of the medians in order to better serve the scope of the selection.

REFERENCES

[1] see for example the emphasis given to this aspect in the next European Framework project in http://ec.europa.eu/research/horizon2020/index_en.cfm

[2] A. Laganà, S. Crocchianti, N. Faginas Lago, A. Riganelli, C. Manuali, S. Schanze, From Computer Assisted to Grid Empowered Teaching and Learning Activities in Higher Chemistry Education in Innovative Methods in Teaching and Learning Chemistry in Higher Education, I. Eilks and B. Byers Eds, RSC Publishing (2009) p. 153-190 ; ISBN 978-1-84755-958-6

[3] Mani A, Nagarajan A, Understanding Quality of Service for Web Services. Improving the performance of your Web services (2002); http://www.ibm.com/developerworks/webservices/library/ws-quality.html, cited 02 Dec 2011.

[4] http://kib.ki.se/sites/kib.ki.se/files/Bibliometric_indicators_definitions_1.0.pdf
[5] ISI Thomson Scientific, formerly known as Thomson ISI
[6] CWTS Center for Science and Technology Studies, Leiden University
[7] CEST Centre d'études de la science et de la technologie, Switzerland

ISSN: 2279-8773