Revista de Sistemas de Informação da FSMA
n. 11 (2013) pp. 2-10

SISTEMAS DE INFORMAÇÃO
http://www.fsma.edu.br/si/sistemas.html

TRILHA PRINCIPAL

New technologies and paradigms in education and research: some reflections

Amit Bhaya, *Ph. D. in Electrical Engineering, Universidade Federal do Rio de Janeiro*,
Ricardo Linden, *D. Sc. in Electrical Engineering, FSMA*

New technologies and paradigms in education and research: some reflections

***Abstract*—This article calls attention to the fact that higher education and research worldwide is being transformed by new tools, new methods of evaluation and collaboration and so on. Although some notice has been taken of this in Brazil, by and large, there has not been much debate or discussion on the impact of these novelties on university education and research. We briefly discuss some of these new ideas and technologies and speculate on the changes that they might or should engender in the Brazilian system of higher education.**

***Index Terms*—On-line education, Internet, evaluation, bibliometrics, collaborative research.**

I. Introduction

The last decade has witnessed the Internet making itself indispensable in all areas of human activity, ranging from commerce to the most esoteric mathematical research. In this article, we will concern ourselves with the impact of this pervasiveness, ubiquity, anytime availability and provision of new tools or 'apps' on university education, more specifically focusing on science and technology, which we are more qualified to comment on.

We will discuss three areas of activity in higher education: teaching, research and evaluation. First, we will briefly describe the new technologies that have become available in the last few years and then speculate on how this might or should impact the corresponding area in Brazil.

A. Teaching and learning

The buzzword today is MOOC, which refers to Massive Open Online Courses. Here is the Wikipedia definition.

A massive open online course (MOOC) is an online course aimed at large-scale interactive participation and open access via the web. In addition to traditional course materials such as videos, readings, and problem sets, MOOCs provide interactive user forums that help build a community for the students, professors, and TAs. MOOCs are a recent development in distance education. Features associated with early MOOCs, such as open licensing of content, open structure and learning goals, and connectivism may not be present in all MOOC projects, in particular with the 'openness' of many MOOCs being called into question.

These courses are proliferating and, interestingly, are currently being offered by some of the most prestigious universities based in the United States. It remains to be seen whether this will set standards for MOOCs that will be followed by others, or whether followers will dilute the quality and content of these courses in order to try to achieve a larger following. Initial reports, both from the organizers themselves and from our students who have taken these courses, are positive, both in terms of the learning experience and in terms of sheer numbers of students.

The first and so far the most famous example of this type of course was the artificial intelligence class taught by Stanford University professors Peter Norvig and Sebastian Thrun (https://www.ai-class.com/). The course attracted 160,000 students from all over the world, 20,000 of whom finished the course, receiving a statement of accomplishment, but no Stanford credits. This course received massive coverage in both print and Internet media, due to the importance of the hosting university, its lecturers and the number of students enrolled.

As pointed out in [1], the tools used in this course were rudimentary: a course home page which centralized all course information and posted each class in video format, hosted on YouTube. Students were evaluated using a multiple choice test and code that was instantly corrected directly on the browser. Students were encouraged to provide each other with feedback and help each other using a discussion forum provided in the class site.

The course organizers operated on the assumptions that the community tends to filter out "bad" collaborators and that professors could help this effort, by pointing out errors in the discussion, even though a truly massive course with thousands of students would generate too much traffic for a single professor to handle.

The framework created for the course spawned a company (Udacity), which calls itself "the future of higher education". In spite of this bold declaration, the questions that such initiatives raise for conventional university education, in Brazil and elsewhere, are manifold. Some of the more immediate ones are:

- What are the basic assumptions for MOOCs to work and when can they be validated?

Autor correspondente: Amit Bhaya, *amit@nacad.ufrj.br*

- Do MOOCs represent competition for conventional university courses?
- How should professors offering conventional courses react to MOOC offerings similar to theirs?
- Does the existence of MOOCs signal a paradigm change in education?
- How can we really evaluate the quality of those that graduate from MOOCs?
- Are students ready for self-teaching and self-motivating paradigms? Is the job market also ready for them?
- How will MOOCs affect the current job pool in academia?

An important point about MOOCs is the change that they will impose on the economics of education. In the 1960's Baumol and Bowen pointed out that some industries suffer from a "cost" disease: namely, that wages always go up but productivity remains the same. Their original study was done on the performing arts, observing the fact that the number of artists needed to perform a piece of classical music always remains the same, but their wages are significantly higher than in the past. An analogous situation holds for traditional education, which is based on lectures for which the number of students per class remains constant. Hence, since the average wage received by teachers has increased, it follows that productivity (as measured in results per dollar spent) has decreased with time.

MOOCs change this picture by allowing for a much larger number of students per class. Thus, following an iTunes type of business model, even if each student pays a small tuition fee for the classes, the possible revenue could be much larger than that obtained from regular classes, since MOOCs can be expected to enroll many orders of magnitude times more students than regular classes.

Notice that enrollment fees, even if they are significantly low, can decrease the number of students and limit them to those who are really interested in the course. Currently available data indicates that most of those enrolled do not finish the course (some MOOCs have a 90% dropout rate) and that it may be the case that many prospective students will be discouraged from enrolling if they have to pay even a small fee. Nevertheless, even if only a small percentage of prospective students actually enrolls and pays fees, the large number of them might make MOOCs economically attractive.

This economic diagnosis seems to be the main reason why many entrepreneurs and capitalists are flocking to online education, seeing it as a target of opportunity, in which income scales, but costs do not and, in fact, infrastructural costs are extremely low, compared with conventional education. This is specially true in Brazil where, even though the law imposes a lot of requirements on online education (see, for instance, the decrees 5622/2005 and 6303/2007), many institutions still see massive online education only as a tool that decreases costs and increases revenues. This perception is one of the reasons why there is considerable prejudice against online education, even though it is valuable tool that can provide high quality education in many areas of knowledge.

Theoretically, online education is not an easy source of revenue. Course materials need to be improved and transformed in order to work online, creating multimedia effects and a way to increase the relationship among students and the instructors. Also, since the material produced will be seen by very large numbers of people, instructors could be expected to spend more time preparing and teaching online courses than their traditional counterparts.

On the other hand, the extremely popular Khan academy model (https://www.khanacademy.org/) seems to be a counterexample to the previous paragraph. Its minimalist approach relies on a no-frills video of a professor lecturing, using a blackboard that might or might not have some technological enhancements. It is effective in many situations, explaining its popularity, and thus might easily become a business model, overlapping many of the requirements of the law and dispensing with deeper considerations of educational value and pedagogy.

Nevertheless, in spite of the hype that surrounds it, the Khan academy model should not be seen as a panacea for education at large. Talbert (http://chronicle.com/blognetwork/castingoutnines/2012/07/03/the-trouble-with-khan-academy/) points out that the Khan academy model is excellent for teaching mechanical processes, such as solving equations or performing algebraic decompositions, but not to elicit dialog and reflection and therefore should not be taken as a substitute for professors and institutions of higher education .

Another aspect of MOOCs that will generate debate is their impact on conventional course offerings within the academic domain. In short, if a top university offers a MOOC on a certain topic, how should the less well known or smaller universities deal with it? Supporters of MOOCs say that course offerings from flagship universities should be subscribed to by the other universities, through a licensing fee, for example, and that the latter can then use these as supplements to their in-class courses, or require students to take these lectures as preparation for in-class sessions, which will then become more interactive and lead to more discussion, problem solving or whatever else the in-class instructor sees fit to do. They also add that anytime availability of recorded lectures allows students to repeat material they have difficulty with, so that when they take their doubts and difficulties to an internet forum or to the in-class session, they are much better prepared to ask specific questions as well as to appreciate the answers.

Detractors of MOOCs point out that if many smaller colleges start using MOOCs, then they would be inclined to make cuts in their faculty, reducing the job market. Furthermore, it is alleged that indiscriminate use of MOOCs would tend to diminish the diversity of approaches to creating and disseminating knowledge that currently exists in the various sites of higher education. This concern seems to be more immediate and threatening in the Humanities, than in the hard sciences, which tend to have a much more

standardized set of concepts and methodologies. In the alarmist limit, one could even argue that the emergence and subsequent dominance of a MOOC-elite would cause a spiral of death in academia, in which the elite prospers, and peripheral endeavors, no matter how worthy, perish.

On the other hand, smaller colleges could benefit from MOOCs by performing the transition from "sage on a stage" to "guide on the side" model **Lewis** meaning that they could include MOOCs as part of their curricula and use them to impart the basic knowledge. Hence, their professors could be free from having to impart the knowledge and could proceed directly to questions and discussions, assuming their students had already accessed the material provided by MOOCs.

This paradigm change seems to be fundamental in order to keep up with times. In-class lectures were fundamental when knowledge was not easily available, professors were the guardians of knowledge and students had access to few other sources. Nowadays, the Internet provides an unending source of information (and misinformation) and could be used as the foundation for the classes. In other words, the pedagogic ideal seems to be in transition to the "flipped classroom" – a model in which teachers preassign whatever lecture-type material is needed, as homework, and reserve the classroom time for peer and interactive learning.

Given that many Internet sources are inaccurate, biased or worse, MOOCs from reliable institutions can become a staple in a participative model, helping the students prepare, much as a book from a well known professor should have done in the past, when assigned readings were used. The idea is that, in this new model, professors at smaller colleges who use MOOCs, would not be liable to lose their jobs. On the contrary, the tasks of MOOC-adopting professors would change from being the sole knowledge imparters to hybrid roles as being sources of knowledge and guides, helping students find the knowledge by themselves, leading discussions, proposing assignments and performing tasks designed to help the students mature and become better professionals.

Another important observation is that education is a product with a consumer base made up of society at large, the future employers and the students themselves. All these consumers will sooner or later be affected by the advent of MOOCs, either in a positive or negative manner. For instance, the effect of MOOCs on society may be positive by diminishing costs, increasing the spread of knowledge, providing quality education in remote locations and by the increasing realization by the students of the different cultural and intellectual realities, which may turn them into better professionals and citizens. On the other hand, society may be negatively affected by the possible loss of jobs, the diminished cultural variety in education, by the loss of regional adaptation in the courses offered and even by the disincentives to the existence of independent thinking faculty members in smaller colleges. Whether the positive or negative effects predominate will clearly be context dependent. In the major centres, concerned with rising costs, the positive effects will probably be regarded as dominant. Small colleges will probably reason that the negative effects are too pernicious for MOOCs to be considered a positive force.

Summing up, there is little doubt that MOOCs represent a major potential change in the way learning and teaching will be carried out in the next few decades. Nevertheless, there is a real danger that MOOCs can be turned into "cash cows". We feel that there is an urgent need to study and debate the issue of maintaining quality in the face of strong incentives to accentuate the positive economic aspects and downplay the educational objectives. In addition, there is a great need to carry out experiments that will provide data that allows more informed positions and decisions to be taken.

B. The emergence of collaborative research

Although individual patterns of doing research may not have changed much over time, the availability of tools that make the discovery of related work much easier has accelerated the dynamics of research in ways that might lead to a greater need and incentives to do collaborative research without the need for collaborators to be in actual physical proximity.

There has been a considerable development of social network tools that promote web-based organization and sharing of knowledge between researchers. Indeed, the current generation of researchers and most certainly coming generations are more likely to interact with their peers through one of the many fora available on the Internet, utilizing the large and ever-increasing variety of tools available to do this.

As is well known, the world wide web and the internet revolution was the fruit of collaborative efforts between physicists at CERN, almost three decades ago. In recent times, even amongst mathematicians, who are well known for solitary work habits, some important collaborative efforts have emerged. Terence Tao, a Fields Medallist, has been instrumental in launching the Polymath Project, which aims to promote "massively collaborative online mathematical projects". It has already been successful in examining a serious attempt to settle the P = NP conjecture in the sense that it was instrumental in coordinating a constructive collaborative effort to examine the proof and ultimately to find a fatal flaw.

The importance of collaboration in science seems to be recognized by technocrats worldwide, and new initiatives to foster it are being proposed. For instance, in June 2013, 50 countries launched the Global Research Council, whose goal is to promote cooperation among funding agencies and increase scientific cooperation between countries.

Given such initiatives as well as the high level of connectivity provided by the Internet, one would imagine that international or even intranational cooperation would be booming. However, at the same time, universities within a country and worldwide usually operate in highly competitive mode, because of pressures relating to funding,

patents and intellectual property rights. Paradoxically, some of the same agencies that promote and even fund international cooperations also put out extremely competitive funding calls that promote competition between potential collaborators. In fact, in the case of Brazil, there is very little cooperation between the top universities, which see each other as competitors, and a significant amount of cooperation with foreign universities, which are generally the favored partners. This tendency is very likely to be strengthened by the new Brazilian Science without Borders program, which aims to send unprecedented numbers of undergraduate and graduate students to be trained abroad. Foreign universities, even those which have never been involved in collaborative efforts in Brazil, are avidly vying for their share of Brazilian-funded students and researchers. Although such an initiative may be seen as a science diplomacy effort, creating social networks which might give rise to scientific cooperation, it is still too early to evaluate this. In addition, the impact, on Brazilian science and technology, of the exodus of a large number of the best students, as well as the funding associated to them, needs to be evaluated very carefully.

The issue of incentives for collaboration has not been addressed in a consistent way and currently it is difficult to discriminate between true collaborations in which both partners bring resources, human and financial, with a genuine two way flow, and collaborations in which the situation is more unilateral. There are many reasons for this: in the past, in a Brazilian context, asymmetrical collaborations with 'advanced' countries were considered acceptable; evaluation of collaborations is still more quantitative (based on the number of publications, theses, etc.), rather than qualitative, masking asymmetry. Quantitative evaluation is pervasive in academia and most of the time, its popularity derives from the fact that it is easy to measure and hard to contest. True cooperation is an intangible asset that is difficult to measure numerically and funding agencies will need to spend time and money learning how to assess it in order to foster national interest.

For this pattern to be broken, funding agencies need to come up with reasonable metrics to evaluate collaboration as well as incentives that tend to mitigate the cooperation-competition dilemma. This is a difficult task, but a necessary and urgent one, in the face of an expensive program like Science Without Borders.

II. Evolution of publication paradigms

Computer scientists and mathematicians have been at the center of battles with some of the major publishers of scientific journals in recent times. Landmark events were the en masse resignation of the entire editorial board of the *Journal of Algorithms* in 2003, protesting the successive price increases in journal subscriptions, although, by this time, almost all technical typesetting, at least for computer science and math journals was done by the authors themselves, at little or no cost to the publishers. Significantly, this revolution in typesetting was caused by the invention of the now famous and indispensable scientific typesetting language TeX, which was the brainchild of Prof Donald Knuth, who was the Editor in Chief of the Journal of Algorithms at the time. To be fair, ever since the late nineties, various declarations of independence and movements towards an open model of publishing had been occurring `http://legacy.earlham.edu/ peters/fos/lists.htm#declarations` and this trend has continued, with the recent boycott of Elsevier journals by mathematicians world wide, which has led the journals to start thinking about changes in their pricing and accessibility policies.

In parallel, there has been a strong movement towards open access models of various types, as well as the strengthening of large archival sites that are now in the process of creating new tools to make their resources more accessible, searchable and customizable. A pioneering initiative was the arXiv website, set up by theoretical physicist Paul Ginsparg. The electronic version of the preprint, called an e-Print, is typically submitted to a journal and simultaneously posted to the arXiv website (http://arxiv.org/). This is now the standard method of scholarly communication used by a very large fraction of the physics, astronomy and mathematics communities. The idea of simultaneous posting on a publicly accessible server and submission to an archival journal is known as the Green Open Access model. Significantly, in recent times, some reputed closed or toll access publishers are now encouraging or even requiring prospective authors to use the Green Open Access model.

We quote from Tony Hey's description: *arXiv is the primary daily information source for hundreds of thousands of researchers in physics and related fields. Its users include 53 physics Nobel laureates, 31 Fields medalists and 55 MacArthur fellows, as well as people in countries with limited access to scientific materials. The famously reclusive Russian mathematician Grigori Perelman posted the proof for the 100-year-old Poincaré Conjecture solely in arXiv.*

In addition, open journals using other models are also growing in popularity. There are specialized directories for them, such as the Directory of Open Access Journals (which lists over 9,400 journals). Jubb [2] points out that the "Gold open access" publication mode in which publishers receive publication charges and access is granted to anyone who might be interested in the research might be an interesting alternative to the Closed Access mode. Notice that such a solution will still limit publication by authors who cannot afford to pay publishers in order to have their articles published in Open Access mode.

Apart from the monetary issue, there is also an evaluation issue. Most researchers need to conform to a certain standard in order to receive grants, promotions or simply to be well regarded by their peers. One problem with this is that most evaluation systems tend to overqualify the existing closed access journals with a long history and disqualify more recent open access ones.

One example of this situation is the Brazilian Qualis system proposed by CAPES. Due to the large number of journals evaluated, they are simply graded by the curve on

impact factors. Even though this might seem fair, it is a huge disincentive for new journals and tends to inflate the importance of journals that have few articles with a large amount of citations (giving them a high average citation count). This is not an easy situation – the Computer Science publication directory contains thousands of entries and evaluating all of them in detail would require a huge and expensive effort. However, if one really intends to foster new journals that publish high quality open access articles, any evaluation system, including Qualis, will have to look beyond clustering by impact factors.

The current publication model also creates a differentiation between the "haves" and "have-nots" in the scientific community. Being competitive in todays fast moving research scenario necessarily implies having access to worldwide publications. Nevertheless, due to high prices for subscriptions, this creates a situation where researchers from smaller and poorer institutions, to their detriment, do not have access to the best journals available.

In Brazil this situation is even clearer. In a pioneering and extraordinarily important initiative, Federal institutions were provided with access to the "Periodicos" portal, allowing them to download articles from thousands of journals, including the best repositories available. Nevertheless, if a private institution wants to provide access to its researchers, a basic package including some of the best repositories would cost hundreds of thousands of dollars. Of course, this has nothing to do with the Federal Government's management of the portal"Periodicos": it has to defray the extra and exorbitant cost demanded by the repository owners for additional accesses. Once again, this points to the need to seriously examine the open-access publication paradigm.

Jubb [2] suggests that this situation might be remedied by using the "Gold open access" publication mode, in which publishers receive their money from publication charges and access is granted to anyone who might be interested in the research.

For the last decade at least, in addition to doing the research and bearing its costs, authors are also responsible for typesetting their articles and providing graphs and figures in compatible formats. Reviewers usually work voluntarily and editors sometimes receive a meager sum for secretarial support (if they receive any money at all). Hence, the cost structure is more skewed toward the researcher. Yet, per page rates can come up to US$1,000, once again making it difficult for small institutions to cope with the cost (and also militating against readability).

Free open access for all is an ideal situtation, where all the problems described above would be taken care of, but it may not be a sustainable situation. Even totally free online journals (such as the one you are reading right now) incur several costs that must be paid by someone. For instance, this journal has a minimalist structure, consisting of an editor in chief, a web person and an editorial board, all of which must have their wages paid accordingly. Besides, there are the costs of Internet hosting and the reviewers wages, who must also be paid. In Brazil, as in most other places in the world, these costs are ignored, given that most reviewers come from public institutions, but their expert work takes time and must be paid (in this case, by the government).

Totally free open access also accounts for a situation where the research is paid for with government grants. It is reasonable to argue that these grants were paid for by the public and hence the public must have free access to them. Access fees make research inaccessible to part of the public. Obviously, Gold open access also solves this problem, but it does create a situation where government funded research must pay fees (usually also coming from government pockets) to enrich a private society that was not a major contributor either to the research or to the paper preparation process.

This is not to say that publishers do not perform a valuable service; however, most publishers earn a lot of money and their participation in the cost structure that leads to the final publication is very small. Obviously, publishers need to earn enough to pay for editorial staff, archival infrastructure and other aspects of the publishing business, but should their profit margins be so high? This is too important an issue to be solved by relying solely on free market or other ideological dogmas.

The Internet has more to contribute than solely reducing costs and fomenting access, even though these are issues of the utmost importance. Palazzo [3] suggests a model that he calls "cooperative edition". In this model, each author would receive a score that comes from his publication record and from the quality of his articles as assessed by public discussion. Hence, each published article would include a forum where the scientific community could discuss the topics proposed and also rate the article. Authors with low scores would require reviews before being published, but well evaluated authors would be granted the right to publish directly. This model is interesting since it foments cooperation among peers, not only among reviewers and writers, but also readers, who usually have a distributed expertise that could complement the author's background.

Although on the one hand, the Federal Government's initiative (Portal CAPES) must be applauded, being unique in the world at the level of a whole country providing access to all Federal institutions, it is a very expensive option, even more so if we compute the extra costs that must be borne by private research institutions that wish to improve the quality of their research and education by paying for access to it. Furthermore, by providing access to bibliometric tools (Web of Knowledge and Scopus) and citation analysis, it has also introduced some curious distortions in the rating of journals, such as the Qualis system promoted by CAPES. It is not our intention to analyze this system here critically, but it does seem to run counter to a world-wide initiative, sponsored by the other major player in the Brazilian science and technology scene (CNPq), called the Global Research Council (GRC) which states that its main objective is to 'agree on an action plan for implementing Open Access to Publications as the main paradigm of scientific communication in the following

years'.

Qualis is intended to stimulate publication in high ranking journals, where rank is based on ISI's Web of Knowledge impact factors. This means that Qualis ranking gives higher grades to older journals with higher impact factors, which tend to be closed-acess. On the other hand, the GRC inititiative, also subscribed to by the Brazilian government, intends to foster open access publication, which usually means newer journals with lower impact factors.

An inititative that would probably go a long way towards the goals stated by the GRC would be the arXiv and episciences model. To understand better why this is true and why Qualis and GRC represent conflicting directions, consider the following data from Tony Hey's blog (http://tonyhey.net/2012/12/19/a-journey-to-open-access/).

The arXiv repository is now over 20 years old and has a submission rate of over 7,000 e-Prints per month and full text versions of over half a million research papers are available free both to researchers and to the general public. More than 200,000 articles are downloaded from arXiv each week by about 400,000 users. Most, but not all, of the e-Prints are eventually published in a journal and this amounts to a sort of post-publication quality stamp. The apparent drawback of multiple, slightly different versions of a paper turns out not to be a serious drawback in practice. Citation counts for high energy physicists usually count either the e-Print version or the published version. A detailed study of the arXiv system by Anne Gentil-Beccot, Salvatore Mele and Travis C. Brooks is published as Citing and Reading Behaviours in High-Energy Physics. How a Community Stopped Worrying about Journals and Learned to Love Repositories. The paper is, of course, available as arXiv:0906.5418.

In the arXiv model, articles become fully available to readers and to those that ultimately pay most of the costs incurred by science (the taxpayers). Episciences, in their own words, is an initiative that provides a technical platform for peer-reviewing; its purpose is to promote the emergence of epijournals, namely open access electronic journals taking their contents from preprints deposited in open archives such as arXiv or HAL, that have not been published elsewhere. The idea is that those epijournals would promote peer reviewing and then publish the articles after adding the value of scientific evaluation. This type of journal is also being referred to as an overlay journal or as an arXiv overlay e-journal.

The arXiv+Episciences+overlay model is open to criticism, since it competes directly with the existing economic model of closed-access publishing. More specifically, why would libraries be willing to pay for access to a journal, if e-prints of the journal articles are freely available at some arXiv? Indeed, this has led to drastic reductions in susbscription revenues for this class of journals. In addition, epijournals assume that articles were not published elsewhere, which may cause a decline in the number of submissions to closed-access journals. The latter may begin to receive less high quality manuscripts and, in order to keep on publishing, they may need lower their standards, causing a spiral of decline of currently highly regarded closed-access journals.

The advent of free and open publications is, in our opinion, a very positive development. Nevertheless, the issue of funding remains. Someone must pay for open access and the cost of an editorial and librarian staff is not low. Will the arXiv+episciences model be able to cope with the costs?

On the other hand, this model seems to be the object of Paul Ginsparg's conclusions in "As We May Read":

"On the one-decade time scale, it is likely that more research communities will join some form of global unified archive system without the current partitioning and access restrictions familiar from the paper medium, for the simple reason that it is the best way to communicate knowledge and hence to create new knowledge. Ironically, it is also possible that the technology of the 21st century will allow the traditional players from a century ago, namely the professional societies and institutional libraries, to return to their dominant role in support of the research enterprise."

arXiv and its counterparts can become a force for the free distribution of knowledge, taking leverage from the paid subscription model and allowing for scientists to spread the knowledge for free (or almost so). Nevertheless, in order for this to become a reality, funding agencies must promote such an approach and the current academic evaluation model needs to evolve, as we point out in the next section.

A. Evaluation and bibliometrics

In the area of evaluation and bibliometrics, prior to the explosion of the Internet, the Institute for Scientific Information had a virtual monopoly on so called bibliometrics. Here is the Wikipedia description:

Bibliometrics is a set of methods to quantitatively analyze scientific and technological literature. The term was coined by Alan Pritchard in a paper published in 1969, titled Statistical Bibliography or Bibliometrics?. He defined the term as "the application of mathematics and statistical methods to books and other media of communication". Citation analysis and content analysis are commonly used bibliometric methods. While bibliometric methods are most often used in the field of library and information science, bibliometrics have wide applications in other areas. In fact, many research fields use bibliometric methods to explore the impact of their field, the impact of a set of researchers, or the impact of a particular paper. Bibliometrics are now used in quantitative research assessment exercises of academic output which is starting to threaten practice based research.

Once again, the Internet and, more specifically, search engine science has transformed bibliometrics and now there are several metrics to measure performance, either of an individual (Hirsch's h-index and its several variants), of a journal (impact factors of different types) and so on.

These indices suffer from one basic problem, deriving from the fact that an observer alters the behavior of

the observed system. Whenever a metric is established, researchers tend to adapt their publication and research tactics in order to receive better evaluation. This is not necessarily bad – if reliable metrics are developed and used to identify and fund good research, they can be an important tool for feedback.

There are other issues associated with most metrics. For instance, Brembs et al [4] have shown that the journal rank, as measured by its impact factor, is predictive of the incidence of fraud and misconduct and that there is also evidence that research unreliability is higher in higher ranking journals, even for non-retracted publications. There are incentives for publishing in those papers that might cause the researcher's ethics to become flimsy, leading to less than reliable results. This paper also points out several other problems associated with the impact factor, raising the question as to whether it should be a fundamental tool for assessing researchers and choosing grant recipients.

Other metric used have similar issues and of course, the problem is not with the metric, but rather with the way it is used. As stated above, researchers tend to transform any metric into a score for their competition for limited resources (tenure, academic vacancies or grants) and objective metrics are the single solution proposed by our systems for two main reasons: the need to be perceived as fair and the lack of resources.

People are reluctant to accept subjective evaluation, specially those in which they are classified as inferior in some aspect, and tend to disqualify the evaluation, if not the evaluators themselves, and this often leads to lawsuits and to confusion in selection processes. On the other hand, if we have numbers obtained from measurable events, such as publishing 3 papers in B1 journals giving you X points versus publishing 1 paper in A2 journals giving you Y points, we can evaluate the researchers results using a simple sorting function. Observe that this process does not evaluate the research itself, but only the journal where it was published. There may be no other option – but this does not mean that such methods are completely fair. Also, since there are thousands of researchers in any country publishing thousands of papers, it would not be practical to analyze them individually and assess their importance and scientific value. Hence, we need to simplify matters and the impact factor metric is one way of doing this.

Alternatives to the impact factor that rely on social networks and the Internet have been proposed [5]. For instance, asking readers who have not published with the authors to assess numerically and textually the papers they downloaded, using this information to gather the importance of each article. Notice that the lack of assessments would also carry some information on the importance of the article. This "crowd metric" is also subject to fraud - researchers can team up and offer positive reinforcement for each others articles, for instance, requiring further nontrivial graph analysis. Besides, spurious evaluation could be granted a lot of weight when the number of analyses is small, which is bound to happen when the journal is less well known. Nevertheless, it is a solution that could be considered and improved in order to bring subjective aspects into play and allow specialists to provide their feedback, which has become increasingly viable in this Internet age.

B. Brazilian initiatives in open access and online courses

The Khan academy which targets mostly high school to first year college education is already a reality in Brazil, through the efforts of the Federal Government and funding from the Lemann Foundation. Its videos are being translated and made available without charge at http://www.fundacaolemann.org.br/khanportugues/.

As mentioned before, the Khan Academy is a low cost initiative that creates classes that do not require a lot of production and can be created at a low cost. It can serve as material to be used as a prelude to the discussions and in-depth knowledge to be provided by in-class lecturers and professors. Nevertheless, despite the hype (such as http://oglobo.globo.com/tecnologia/khan-academy-aponta-para-futuro-da-educacao-online-2824034), even Khan himself does not believe that his model is a panacea for education everywhere.

Another important step was taken when Brazil joined the Global Research Council (GRC), a virtual organization with members of the Governing Board from the US, Germany, Brazil, Saudi Arabia, Japan, China, Europe, Canada, Russia and India. GRC is proposing an action plan for implementing Open Access to Publications as the main paradigm of scientific communication in the coming years. If this becomes successful, it will represent a strong push towards the open access and worldwide collaboration models discussed in this paper.

On the other hand, there are inertial forces in Brazil that tend to counteract these efforts. The academic evaluation models utilized in Brazil rely strongly on a points system, where publications are divided into strata and each article published in them gives the author or his department a certain number of points. Each academic is evaluated according to his points, becoming eligible for grants for instance, and similarly for the academic programmes at each institution offering university-level education. CAPES and CNPq, the Brazilian institutions that evaluate and fund science and technology, have not so far shown intentions of revising this model, or promoting public discussions in order to find a better alternative for the future. This points-based model is a strong force, engendering competition instead of collaboration, and one that will keep open access from prevailing, as we will discuss further in the next section.

III. Discussion

Can we discern a trend in all these developments? We offer some thoughts on this, aware that all predictions are dangerous and also that this is just a first attempt to initiate a larger debate on some of these issues.

It is clear that there is a strong emerging trend towards openness, whether it be in terms of our teaching and our classrooms, our research and our publications and even of our evaluation by peers. It no longer seems an exaggeration to say that, if one does not participate in this open revolution, then one will run a very real risk of ceasing to "exist", or, at the very least, ceasing to be relevant.

This trend also comes from the government itself, recognizing that results coming from public funds cannot be made private and that they have a fiduciary obligation towards spreading the benefits achieved by publicly funded research.

There is no question that the feedback provided by openness is beneficial to all the participants in the process. For instance, in the specific case of teaching, which varies greatly in quality, even at prestigious institutions, just the fact of knowing that one's lectures will be recorded and put out on the Internet or even on an Intranet, will, in our view, give most teachers sufficient incentive to deliver their lectures to the best of their abilities. In addition, it will induce more professional conduct, both from students and from teachers. Other benefits accrue over time: students will be able to use their time more effectively, reviewing material that they had difficulty with when trying to absorb it in real time. A good set of lectures, recorded and made available for posterity, will free good lecturers to prepare new sets of lectures in other areas, while others can take over the older material, aided by a good set of on-line lectures.

The issue of intellectual property is quite sensitive and institutions need to define their policies on this matter quite clearly, in order to avoid future conflict. Besides, professors and lecturers need to increase their awareness to those matters, as well as creating a culture of citation to the original lecturers. Nevertheless, the rights should be defined with the concept that cooperation is not only acceptable, but also desirable - hence, the copyright adopter should not prevent editing and altering existing classes either for making them more adequate to a local reality or for simply updating or reflecting the lecturer's views.

There is also the issue that that prospective students may choose possible advisors by looking at their research performance as measured by the many metrics that are now openly available. Once again, transparent behavior by researchers is a good policy: by adhering to such initiatives as registering on ResearcherId, ORCID or on one of the larger research social networks on which their scientific work can be made accessible. However, an informal evaluation of the Lattes database of researchers curricula indicates that few researchers have actually signed up for a ResearcherId number (official Thompson Reuters statistics indicate that there are less than 50.000 ResearcherIDs in Brazil, while there are 1.6 million curricula in the Lattes database). In the case of Brazil, where many researchers have long names with many possibilities of being parsed into a last name, first name format, this is a very important step in disambiguating a citation record.

Although we have no clear answers at this time, we feel that the following list of topics merits urgent attention.

- Impact of MOOCs on our university courses. Should we update our courses and launch similar initiatives? Should we ignore MOOCs from outside Brazil, at the risk of losing the interest of our students and our relevance worldwide?
- Should we make the transition from no classes on the internet to more or even all classes on the Internet? If so, how should this be achieved, and how could we make government agencies cooperate with this objective?
- Should we be giving incentives to new efforts to transform Brazilian journals into international journals hosted or managed by the large publishing houses that currently promote the closed-access, bundled model of journal publication? Or should we be thinking about the creation or merging of existing efforts like Scielo with large, open-access archives like arXiv?
- Should we change the way we evaluate professors in Brazil? How can we foster collaboration and the transformation of Brazilian academia into a real web of science?
- Does the current structure of Brazilian evaluation system promote innovation? Do the incentives that professors and lecturers receive point them toward collaborating, innovating and improving our social and scientific landscape?
- How do initiatives towards openness (such as the GRC) conflict with existing evaluation systems? How do we solve those conflicts?

These are important questions to which we have no straight and simple answers. But it is clear they should be kept in mind and discussed urgently, if Brazil wishes to become a real scientific powerhouse.

On the issue of publication, the tendency towards openness is even stronger. Sites such as arXiv and ePubSciences point towards a direction where science will effectively become public property, allowing for greater knowledge sharing and faster generation of better results. Still, there must be someone willing to pay for the necessary infrastructure. However, given the current amount of money spent on subscriptions to closed-access publications, this could be considered to be a minor issue, for we could probably maintain open-access archives and review systems, for a fraction of the amount that libraries spend each year on closed-access publishers. The point underlying the previous paragraph merits re-emphasis: the system must be restructured to validate open-access sources. If funding agencies continue to evaluate researchers based on metrics that favor traditional closed-access publications, the tendency towards open-access can be considerably slowed down.

In the final analysis, we should not lose sight of the fact that we ourselves are the makers of the system. Universities can lead the way by defining merit and promotion in a way that takes these new open-access approaches

into consideration. Once this trend grows, it will define a natural path for funding agencies and governments. If we wait for the revolution to come from the technocrats above, we will most certainly never progress towards truly open science.

References

[1] C. O. Rodriguez, "Moocs and the ai-stanford like courses: two successful and distinct course formats for massive open online courses", *The European Journal of Open, Distance and E-Learning*, 2012. [Online]. Available: http://www.eurodl.org/?article=516#AI-Stanford (visited on 06/10/2013).

[2] M. Jubb, "Go for gold in the scientific publishing revolution", *Research Fortnight*, p. 22, 2012. [Online]. Available: http://www.ornaverum.org/reference/pdf/local-copy/019.pdf (visited on 06/10/2013).

[3] J. P. M. Oliveira, *Como vencer na universidade*. 2012. [Online]. Available: http://www.amazon.com/Vencer-Universidade-Portuguese-Edition-ebook/dp/B006ZBX2T4 (visited on 06/10/2013).

[4] B. Brembs, K. Button, and M. Munafo, "Deep impact: unintended consequences of journal rank", *ArXiv.org*, 2013. [Online]. Available: http://arxiv.org/abs/1301.3748v3 (visited on 05/10/2013).

[5] J. Priem, H. A. Piwowar, and B. M. Hemminger, "Altmetrics in the wild: using social media to explore scholarly impact", *ArXiv.org*, 2012. [Online]. Available: http://arxiv.org/html/1203.4745v1? (visited on 03/20/2012).