DARIAH Winter School in Prague

Open Data Citation for Social Sciences and Humanities

24th to 28 of October 2016

Session 5: Evaluation, Acknowledgement and Credit Circulation

5-Evaluation, Acknowledgement and Credit Circulation

Open Peer Review

Julien Bordier is a sociologist, an independent scholar who works for OpenEdition on an experiment of open peer review and open commentary on a corpus of pre-publication in French.

Introduction

Peer review is what ensures a scientific publication to be a real scientific publication. Traditionally, it is said that everybody has to be masked to make clear and objective review of text. So, a peer reviewer is like Spiderman. And Spiderman says that “With great power comes great responsibility”, because it is a great responsibility for a reviewer to evaluate a text as it will determine if the pre publication will be published or not. The point of *open peer review* is to *open the process* and that means that nobody is anonymous anymore. This practice is developed in the perspective of open access but one can imagine a publication which is not in open access but which also practice open peer review. In the experience we implemented with OpenEdition, everything happened online: the prepublication was displayed on a scientific blog, hosted on Hypotheses. Our hypothesis when we experimented was that we could have a better level, a better quality of scientific communication if everything is open and accessible to everybody. The idea is to make possible conversation between the authors and the reviewers. I will no longer follow the metaphor but you have to know that when Spiderman wants to make a kiss to Marie Jane Watson, he takes off the mask!

Presentation of the experiment

Protocol of the experiment with two different branches:

- Open peer review
- Open commentaries

The report is called: *Open peer review: from an experiment to a model A narrative of an open peer review experiment.* It is available here in open access: https://hal.archives-ouvertes.fr/hal-01302597
You can annotate it with the tool we used for the experiment: hypothes.is

This was implemented on a famous journal called Vertigo, hosted on Revues.org. It is a French-speaking journal based in Montreal about environmental science and ecological issues. We implemented the two different protocols. The first one was open peer review protocol: we published online, on the journal's blog, five pre publications that was submitted to the journal, we found reviewers and asked everybody to know if they agreed

to disclose their identity and to make the process *visible* online - of course they accepted. In a practical way, you have the blog post, which is the pre publication and the comment of the post are composing the evaluation report. So, it is quite identifiable and relevant because it is blog form and people are getting used to it. This was the first and main part of the project and we were very ambitious about it and, afterwards, it worked quite well.
We also implemented another protocol, which is open commentary. At first, this second protocol was not for us so appealing but it revealed itself to be very interesting. In this second protocol, the journal selected some articles to run the experiment after receiving a lot of propositions and especially articles that were not really well written. The idea was to publish online texts that has a real scientific interest but that was not really well formalised and to ask the community, not just the designed reviewers, to comment and to help the author improve the quality of their pre publication. This second protocol appeared to be really relevant and it became a very helpful insight for the authors engaged in this process.
The technical implementation is the editorial environment made by Revues.org (where the journal is hosted) and also by Hypotheses.org (where the journal has its blog). The point was to make both of them communicate and it was important to develop something like that because they represent a legitimate space where scholars recognise and regard.
The traditional review process is a wheel of evaluation: the reviewer makes his remarks about the text, usually in a pdf or a doc file where he inserts comments within the document. Our challenge was to find the right tool to make reviewers annotate the text. We decided to hypothes.is.

Results of the experiment

Open peer review is subject a great discussion and a lot of people think that it will not work while others think that it is the future of scientific publication. I personally don't know but what I saw as a sociologist, playing as an assistant editor implementing the experiment, is that there was a lot of enthusiasm around it. I think that it is the first important point because it means that the scientific community is quite ready to open the evaluation system and especially for publication. Times are changing, like the Bob Dylan's song. Anonymity used to be important in the 70's for example, women scholars asked for it in the review process in order to be fairly evaluated.
Another interesting result is that when you read literacy about open peer review, you have a kind of myth saying that open peer review is quicker than traditional review, but from my perspective and what we experimented, it is not true at all. One can think that as it is on a computer it goes quick and easy, but it is not the case, it is the same as in traditional peer review. Open peer review does need human mediation, it is not robotic or cybernetic peer review. It is still difficult to get reviewers, to get people to follow the deadlines and so on. And it is even more important for the open commentary protocol because the point was just to help the author, so scholars does not have this habit, they usually do it during a seminar but not for publication. It was quite difficult to explain the process and its issue. It was also difficult to get people from outside of the scientific community to comment the text. It is interesting because it means that we still have a lot of work to do about how to get a better integration between scientific community and the rest of society.

Overview of the results

- Enthusiasm
- Need of human work
- The experiment worked because of the integration to a relevant environment (between revues.org and hypotheses.org - seen as a legitimate publishing environment in the French speaking area). For example, Mickael Bon is trying to make an open peer review journal platform, in STM, called Self Journal of Science, but he has a lot of difficulties to get scholars publishing on his platform because it doesn't have the legitimacy.
- In open peer review, reviewers and authors are able to talk to each other and it worked quite well. I do not have empirical results to present now because our corpus was small: it was just five texts open peer reviewed and five texts opened to commentary.
- As a sociologist, I asked questions to all the protagonists of the experience: did you search the person you were doing the review? Of course, everybody did it! And of course, even in classical [closed] review, everybody try to do it and to know who is reviewing and who is the author. And of course, it is really to easy know and that is why classical review is just a myth that now should be opened.
- People did get information about each other and knowing what each one is working on, you can feel how the remarks are legitimate and true. I think that even for the journal or the editor, it is very important because you can have a more fluid conversation in the exchange of information.
- At last, in the community, it shows who is actually working and who is actually doing good work about reviewing. To get a comparison, the text that was openly peer reviewed was also traditionally reviewed, it was a double blind review. The result is funny because in the traditional review, the reviewer just added a line on each section; whereas in the open review, reviewers make efforts to be clear and understandable.
- Besides, it is not only the journal that can judge the level of the evaluation, it is the whole community.

Limits and further potentialities of the experiment

- In order to get users to annotate the text, we asked them to use hypothes.is and it is clearly not ergonomic because you have to ask the reviewer and the authors to create an account on hypothes.is and to install it on their web browser. I had phone calls from older scholars who had some difficulties to use it. -It is not a judgment- but when you have a discussion with someone who do not make difference between Google and its mailbox, when you have to get him using an annotation tool which is not implemented on the platform he uses, it is really difficult. At the end it worked. This is reminder for the idea that you do need this specific human work around open peer review - you have to get people really doing.
- Hypothes.is will be implemented on OpenEdition Books! I think it is a good start, but I think that hypothes.is has some problems, for example on Vertigo Journal, on one

of the text, when you open hypothes.is the annotation refers to another webpage in German which has nothing to do with the displayed text. This is a big problem.
- Potentialities about credit circulation: it makes visible the work of everybody involved in the review of pre publication. It is a very important part of a scholar but it is an *invisible task*. Our commitment in open peer review with this experiment is to make possible to credit everybody who worked on the text. In the way we did it, reviewers are credited by citation, so once the text is published in the journal, you can find the name of the evaluators. So if a scholar wants to refer to it, there is an URL that you can use for example in your curriculum vitae. For the moment, it is just *bricolage* or *craftwork*, it is not yet real digital publishing; an important issue would be to make the name of the reviewer or commentators indexable. It would enable when you use a tool of indexed authors and find what an author wrote, what he reviewed, what he commented.

Conclusion

With this experiment, we tried open commentary and open peer review on pre publication, before the document is really published. We just tested it this way but every form are hybridable: you can open texts to commentaries post publication, you can also close or disclose some parts of an evaluation, it only depends on the agreement you have with reviewers and authors. I think that editors should keep hands on this to find the right policies, the most relevant for their field of research and their publication policy.

Practical Session

- Install hypothes.is on your web browser
- Annotate parts of the report for the European Commission published on HAL: https://hal.archives-ouvertes.fr/hal-01302597
- Give feedback: ergonomic, how you feel with the tool, any particular problem?

Contact

Julien Bordier, sociology PhD, independent scholar, editorial adviser, works on public-space issues. He conducted the open peer review experiment for OpenEdition / Centre pour l'édition électronique ouverte.
Twitter: @julbordier
Email: julien.bordier@openedition.org

Simplifying License Selection

Ondřej Košarko, Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University, Czech Republic

Neglected task

- Research(ers) rely on availability of data & software, but license clarifies what you can do with it
- Authors hold exclusive rights: unless they “give up” the exclusivity, the data are not “usable”; That means assigning a license

Attributing a license to your data is a necessary task as research itself relies on some data. For example, we use data produced or collected by colleagues and we run experiments on them, we evaluate what other researchers did, we use the tools they produced in order to move the field further. Still, without a license that would clarify what you can actually do with the data or software, you cannot work. Of course you can do whatever you want with what you found on the Internet, but the problem is that you shouldn't do it this way because you don't know how the data was collected, if there are private information, if data are reliable, etc. The only person who should know it is the author of the data and it should be the only person that has the right to allow derivative works with the data and tools and publish them. In order to make it available for a wider public, he has to give up on some of his exclusive rights by providing a license that gives either individual persons or in better case the general public the right to operate or to modify the data. As seen during the previous session about open revisions or open reviews, it is necessary for the good advancement of the research to have access to the underlying data. I called this a neglected task because not all of the authors have attributed licences to their work.

Resignation

Some of those reasons might be:

- Not distributing the work
    - Law is scary/boring/complicated
    - License texts are long and hard to read

Without making a license statement, it is like hiding data on your hard drive thinking that you don't need to care about it.

- Not attaching a license
    - I did not assign a license so everyone is free to do as they wish, right?
    - No one reads them anyway

If you want people to do whatever they want with your data, the way you obtained the data allows you give them this right, it better to tell it. There are several licences that say it in a few words.

- Using a license that is familiar with
    - Good attempt, but GPL with data?

You might hit the right one or not, but some of the well known licenses might not be the best solution, for example licenses that are good for softwares are not adapted for data. If it

relates to softwares, executables and libraries, if you attribute it to data, what does it actually mean as there is no executable, no library. It equals to not attributing any license because the user cannot get any special right.

Data and software

- Overall different beasts
    - Regulated by different directives
- Different caveats
    - Software: libraries versus executables
    - Databases
- Different licenses [for data and for software]

There are such differences that it is in fact regulated by European directives. On data, you are often guided by the regulations for databases or collections of work. For software, you might find distinction between library and executable.

Making it easier

One thing you can do each time you want to release something is to ask the person in charge what you can precisely do and what you cannot. You can also build these questions into a process, a tool.

- Get together with people who know this matter
- Prepare a tool/process to help choose
    - Filtering based on requirements
- Capturing differences of data and software
- Adhering to limitations, according to the way you obtained the data
- Promoting open access, as free as possible

When you turn this question into a process, you do one more thing, you can push people in certain direction: advocating open data and open access licensing if the strings attached allow it.

A few tools

- http://licentia.inria.fr/
- http://wizard.elda.org/index.php
- https://tldrlegal.com/
- https://ufal.github.io/public-license-selector/

GPL license: it is nice because you can modify but you have to disclose source code and instal instructions. There is also a quick summary of the provided full text and this can be a collective process with for example the change sets, you can go through the changes of the summary if there are any. There is also a section for comment where people can talk about some not obvious interpretations of the license. Terms of services are even worst than licensing: they are longer and the language is really complicated. When working on the term of services, we should take care of the last changes. If you are using it, you should choose the latest version. It is supposed to be for users, because it explains the framework in which you are allowed to use it, but it is written for lawyers.

Tl;dr legal

Tl;dr stands for *Too Long Didn't Read*, it is used on the Internet either to state that you didn't read it or when you ask for a summary. This perfectly grasps the licenses issues, terms of use, community driven summaries (textual, can/cannot/must), tracked changes, verified content [by a legal expert]

Public License

It grants certain rights not to one particular user, but to the general public (everybody). You can configure the tool to allow modification or not (non derivative is not open data compliant).
According to the permissiveness, you can find open data compliant licenses. From the software development perspective, there is something that might happen is that you might find a combination of licenses that are incompatible: one licence can require you to distribute all the work and another one not to make derivative, it is conflicting. So there is also a licence compatibility issue and it depends on an interpretation that can be different from the official one provided by the authors of the licenses.

- A lot to choose from:
    - 78 OSI approved licenses for software
    - Creative commons, Open data commons, etc.
- Open access
    - Not all public licenses meet requirements of Open data/open access [Open data compliant or endorsed by open data]
    - http://opendefinition.org/licenses/

TOOL: Public license selector

- Tool
- Github
- Pawel Kamocki (legal expert specialised in IP and personal data protection), Pavel Straňák and Michal Sedlák
- Distinction between data & software
- User friendly, asking questions and providing with comprehensive explanations
- Promote open licenses
- Licenses not compatible with the answers are removed from view
    - Question about licenses already used
    - Table of compatible licenses
- Licenses ordered based on openness

How to attach a license?

- Some licenses tell you to insert it prominently in all relevant locations. But, is the download page enough? At the beginning of each file? What about textual only files, how to make the distinction between the text of the license and the text of the file?

- A solution is to mention your license in your metadata [but it doesn't apply to the metadata themselves-descriptive metadata are not copyrighted]
- Download license in RDF (Turtle syntax) to attach it to your data

Resources

- Public license selector: https://github.com/ufal/public-license-selector
- Article about the public license selector: http://www.lrec-conf.org/proceedings/lrec2016/summaries/880.html
- Open Source Initiative: https://opensource.org/licenses/alphabetical
- Creative Commons: https://creativecommons.org/share-your-work/
- Open Data Commons: http://opendatacommons.org/licenses/odbl/

Contact

Ondřej Košarko, UFAL

Ondřej Košarko is a programmer working at the Institute of Formal and Applied Linguistics (UFAL), Prague, Czech Republic. He is one of the developers behind LINDAT/CLARIN repository. The repository is based on DSpace and has been modified to meet the needs of CLARIN centers. This modified version is now deployed in several member institutions. He is also responsible for parts of shortref.org, a tool to ease persistent data citation, and various other bits and pieces like this guide for choosing the adequate licence.
Institutional websites: http://lindat.cz & http://ufal.mff.cuni.cz/

Email: kosarko@ufal.mff.cuni.cz

Evaluation of the SSH and the evolution towards open science

Ioana Galleron, European cooperation in Science and Technology, COST action

General outline

- Evaluation systems in Europe and the place of the SSH
- Challenges of SSH evaluation
- A new player in the field: European Network for Research Evaluation in the SSH (ENRESSH) (COST Action 15137)
- Gathering data about the SSH: main problems
- Some observations about the impact of the open science trend

Evaluation systems in Europe and the place of SSH

They can be classified into four criteria (Geuna et al., 2001).

- Evaluation performer: national, regional or institutional player
- Evaluation purpose: funding allocation or strategy formulation
- Criteria for evaluation: there are many but four main groups: quality, quantity, impact and utility.
- Methods: bibliometrics (impact factor), scientometrics (also takes into account the research environment, size of teams, etc.), peer-review, peer-review supplemented with bibliometrics/scientometrics (*informed peer-review*)
- An interesting dimension in classification terms should be who is evaluated (individuals, teams, institutions?)

There is a lot of discussion around classifications of systems because for example some people think that a funding system is not an evaluation system, it is just a performance based system.

To summarize, evaluation systems can be placed somewhere between two “typical” models:

- evaluation conducted ex-post (after the research), performance-based, indicators, peer review.
- evaluation size based, where allocation of funds or strategy will be formulated looking at the teaching volume, students, the staff, etc.

The general trend is obviously towards the first type performance-based resource allocation system because of the new public management policies and everybody looks for accountability and value for money.

What about SSH evaluation?

When you look at how it is specifically done in SSH, the image is more blurred. We conducted a survey about practices before the beginning of the action to understand the system applied to the SSH. We had 43 participants involved in research evaluation from 25 European countries answering about:

- the level of the evaluation protocol (national, regional, institutional)
- disciplinary differentiation

- who is evaluating
- object of evaluation
- purpose (funding/strategy)
- methods
- timeline
- transparency
- costs

Even if the interviewees had a good knowledge of evaluation, they often had no clue about how to answer our survey; and this is already informative about the situation of the SSH evaluation.
=> Good degree of agreement amongst respondents from the same country about who is evaluated, the methods applied and the link between evaluation and funding.
=> But there is a lot of disagreement or even misunderstanding about **terminology**, ex. evaluation/assessment, "excellence", etc.
=> Evaluation is mainly done nationally and linked to a performance-based funding system.
=> National publication database about SSH production in 13 countries
=> Respondents signal the existence of an SSH specific evaluation in 14 countries (but with a low degree of agreement about the existence of specific methods of evaluation for the SSH).

Methods

- Peer review
    - Most countries use peer review to evaluate SSH
    - Only 9 use informed peer review
    - In most countries, peers apply criteria but there is no agreement about criteria among participants from the same country
- Bibliometrics/scientometrics
    - It seems to be the main method in six countries
    - There is no agreement on the kind of data used to evaluate the SSH

Transparency

Respondents from 14 countries consider their system to be transparent regarding the criteria applied for SSH evaluation whereas 11 consider it opaque. There is a lot of disagreement about this dimension.

Challenges of SSH research evaluation

- Scholars don't want it because of academic traditions. For example, looking at book reviews one observes the tradition of courtesy, of collegiality, you don't say bad things about your colleagues; ideology, some scholars will refuse any form of evaluation as being representative of a relation they reject between individuals and the State; fear, based on real concerns about what is coming out with the evaluation.
- Managers (and decision makers) don't like it because there is a lot of disagreement amongst scholars and no policy maker is happy with an exercise that creates tensions in a field which he/she needs to deal with

- Research Evaluation don't know how to do it, because of the shortcomings of bibliometrics, the problems of peer review and the diversity of the SSH.

Shortcomings of bibliometrics

Ill adapted to the SSH

It has been repeatedly proved that bibliometrics are ill adapted to the SSH:

- specificities of Lotka's distribution; it says that out of a population of scholars, you will have a smaller proportion publishing 3, 4 or 5 articles than the proportion of scholars publishing just one article. We have some data from Italy where it appears that this needs field adaptation when applied to the SSH.
- Also, the Bradford's law is not working because "no core literature in a field can be identified" (Nederhof et al., 1989). Journals in the field are classified in three tiers: the first is the one who will publish the most important papers in a discipline; the other papers, less important, less visible will be published in a second tiers; the third tiers will be with everything else. This law is useful when you are trying to spot a core group of journals where the most innovative ideas from a discipline are published, but as it happens, nobody ever managed to find a core group of journals even in English Language and Literature for example.
- SSH publications are poorly covered in major international databases (WOS, Scopus). There is an under coverage of SSH scholar published books and the coverage of journals is biased with regards to the "language, country, publisher size and age" (Hicks, 2011); you have much more chances of being in WOS if you are an ancient journal than if you are a young one.

War on JIF

Tensions about the Journal Impact Factor concerns all the sciences because:

- It concentrates on journal articles to the detriment of a much more diversified research output landscape
- It is a very approximate proxy of quality: it is not because it is published in a prestigious journal that it will be excellent science. Students in MIT fabricated a scientific paper with a copy and paste method and they had it accepted in a prestigious journal. It is not a guarantee of quality.
- We also know that it is conductive to multiple controversial behaviours (parroting (researchers repeating the same kind of research, sending very little differentiated papers to different journals based on the same research and the same funding, just presented otherwise, to boost their impact and citation factors), psittacism (a sort of "I quote myself"), parochialism "I quote my dear friend X and they will do the same in turn", etc.).
- It under evaluates and under represents the outreach of a publication. This has been demonstrated since we have altmetrics, this is the publication of open access that allowed us to see that there are other parameters for impact: views, downloads, shares, comments on social medias.

Bibliometrics applied to the SSH

Czech example: Malek et al. "System of evaluation of research institutions in the Czech Republic" (2014). It is a performance based research evaluation system based on points that

determine money allocated to a university out of the points they can collect over a given period of time. 61% if you are publishing something in a journal in WOS, books give only 9% of the points. The rationale behind it is that a paper indexed in the Web of Science has a minimum of 10 points but if you look at the formula it is multiplied by a factor. So the idea is that one paper published can easily bring 300 points while if you publish a paper in SSH journals listed in ERIH, it can bring only 30 points, so there is difference 1 to 10 between the two. So it is incitative to publish papers in certain journals than books, even if they are more adapted to the research conducted.

Problems of peer review

- Blind or not, prior to publication peer review may be anti-innovative and can lead to gatekeeping; we have numerous cases of blocking innovation through peer review
- In small countries or disciplines, the pool of evaluators may not be sufficient and this happens easily in Europe because we have a lot of *demographically* small countries.
- Better to it internationally, but criteria and expectations are not the same. A very distinguished Czech professor evaluating a piece of research in French will have the same criteria as an English one? Does the title of professor mean the same thing in all European countries? Even the perimeter of the SSH is not the same when looking at different European countries.
- It is time consuming and the cost may exceed benefits when you put into place huge campaigns of evaluation for an entire system or a discipline.

SHS specific biases

- There is a lack of transparency about the methods and criteria, the selection of reviewers, the treatment of conflicts of interest
- There is a low degree of organisation and quality control over peer reviews
- And there are acute intra and interdisciplinary conflicts about quality.

Survey

We organized a survey in 2014-2015 about peer review in (prestigious) publishing houses, within a project subsidized by ANVUR agency for research evaluation in Italy.
Amongst the questions asked:

- Has the PH a scientific committee assessing the book proposals?
- Has the PH a blind review system?
- Does the PH provide referees with an assessment sheet or guidelines for the evaluation of book proposals?
- Does the PH reject negatively-reviewed book proposals or asks for revision that take into account the reviewers' reports?

We selected publishing houses with specialized series in philosophy, history, literature, languages and linguistics.

- More than 250 publishing houses contacted (100 in Slavic area, 96 in UK and USA, 61 in Italy) => 54 answers
- Up to 9 reminders, high level of opt-out for numerous questions.
- Italian PH: 25% declare not having a scientific committee; more than 33% do not practice blind peer-review; when a peer-review is in place, 35% affirm not using an

assessment sheet as a guidance for peer –reviewers; only 2 PH communicated their assessment sheet.

Peer review in SSH journals and publishing houses:

- There are huge discrepancies with regards to review practices (length, argumentation, style) from one discipline to another. This is a very difficult information to access but we manage to build out a small corpus with the reviews we had in hand from participants to the project. Reviews are sometimes done in one word ("bof" as a peer review evaluation for proceedings to an international conference, see Anne Baillot's paper, Peer review says "bof"). On the other hand, we had 10 pages of observations over an article of 15 pages. So, differences are really huge.
- National incentives and authority involvement are needed to gather a more accurate picture
- When such national incentive exists, it leads to interesting initiatives, for instance the Flemish initiative in Belgium: they decided that the situation is not acceptable, that they are willing to accept books for evaluation but only books published in publishing houses having thorough peer review procedures; whatever procedures, but they have to be thorough, easy to monitor and demonstrate that a certain peer review took place at a certain moment. And they award a label to publishing houses that put into place such mechanisms.

Example of the peer review in the French assessment exercise:

- analysis of 104 reports of evaluation of SSH research units (all research units in two regions, Bretagne and Rhône-Alpes)
- evaluated period: 2004-2008
- conducted using corpus linguistics methodology and tools (Atlas.Ti and AntConc)

Question => How do official criteria for quality are translated into this reports?

Official criteria (AERES)

The words associated with good research are:

- New (original, breaking through, generates new patents, methods, norms, etc.)
- Partenarial (multidisciplinarity is encouraged, as well as extra-academic cooperation).
- Impactful (in the academic community: citation indexes, number of thesis, etc.)
- Useful (to the economy; to the society)
- Recognised (by peers: publications, selection as speaker, leadership, membership; by others: expertise, rewards)

Also, good research in SSH is published in certain journals (« périmètre de scientificité »). It is not really clear if it is an added criterion, a specification, or the only criterion of quality. But what is [scientific] quality?
In practice, reports shows that research appreciated as being good is not individual, the group research must have a thematic coherence. So, in France, individuals work under three constraints from:

- The institution, because they are strongly invited to be a member of an established research group;
- The Research unit, where people are incited to be in conformity with the group
- The discipline, since it is necessary to conform both in choice of research group and research production to the expectations of the CNU (Conseil National des Univesrsités, national council of universities)

3rd Challenge: Diversity of SSH

SSH is a general umbrella for a very contrasted landscape with regards to the publishing habits and underlying representation of quality. Traditional classification in STEM and SSH disciplines are not verified when we are looking at the publication habits.

- Mutz et al. "Types of research output profiles: A multilevel latent class analysis of the Austrian Science Fund's final project report data", 2013. Latent class grouping of publications: it is not unexpected but you have new insight, for example economy is closer to computer science and to mathematics, whereas everybody puts economics into the SSH. So what are we talking about when we designate 'the SSH'? => We are putting in the same basket very different kinds of fruits.
- Another project at ETH Zurich, funded by CRUS (Rector's conference) in Switzerland in 2012. They conducted repertory grid interviews with 21 scholars from 3 disciplines: German literature studies, English literature studies and art history. They observed many of the discrepancies, differences in view mentioned above.
- Ochsner et al. "Four types of research in the humanities: Setting the stage for research quality criteria in the humanities", 2013: 4 types in the SSH in terms of perception. In terms of perception of quality, 7 scholars and we have 4 representations of what quality is.

COST Action 15137: ENRESSH

European Network for Research Evaluation in Social Sciences and Humanities

- Started April 2016 > End March 2020
- NOT a research project: coordination of existing research
- 33 European countries involved and observers from South Africa, Mexico, Moldova

Objectives

- improve the understanding of how SSH fields generate knowledge; because basis of an evaluation is to know what scholars mean, what are the pathways towards producing something.
- to observe what kind of scientific and societal interactions characterize SSH;
- to observe patterns of dissemination and quality representations in the SSH.

Work groups

- WG1: Conceptual frameworks for SSH research evaluation
- WG2: Societal impact and relevance of the SSH research
- WG3: Databases and uses of data for understanding, monitoring and evaluating SSH research
- WG4: Dissemination
- A transversal special interest group for early stage researchers

Gathering data about SSH research

Some people say that SSH scholars publish books. Ok but is it true? Because the evidence so far gathered shows that the SSH scholars start to publish more journal articles than books.

Therefore, we look at:

- Full bibliographical coverage: we want to have visibility of publications not indexed in WoS or Scopus
- Not for citation counts: monitoring and understanding the system
- Our focus is on metadata rather than on datasets and publications.
- Successful development in countries where full coverage is part of a funding related evaluation system: eg. Norway and Belgium (Flanders)

Countries where data are gathered about SSH research:

- Belgium (VABB-SHW)
- Scandinavia: Norway (CRISTin), Sweden, Denmark, Finland (KOTA)
- Czech Republic
- France (RIBAC: only for CNRS researchers, not counting the universities (2/3))
- Italy (CINECA)
- Lithuania
- Latvia
- Poland
- Portugal
- Spain
- Switzerland
- UK (RIN)

Important differences of coverage, methods, categories.

The leader in the field is Norway with CRISTin: Current Research Information System In Norway. They have given the possibility to the scholars of declaring every kind of output, not only WOS. CRISTin: principles behind the use of institutional data in a national information system:

- Completeness: All scholarly publications should be included
- Transparency: Every institution can see and check all other institutions' data. The national database is also online and open to society at large.
- Participation: The indicator is developed and maintained in collaboration between the institutions and the authorities
- Multiple use of the data: CV's, applications, evaluations, annual reports, internal administration, bibliography for Open Archives, links to full text, etc. An important point is that data are imported (from ISI) and they can add data (about other publications). Scholars themselves can check if the number of outputs, not only publications, is correct and add whatever is not there.

Another interesting initiative is VABB-SHW from Belgium.
Engels et al., "Changing publication patterns in the Social Sciences and Humanities, 2000–2009", 2012

- Creation of an "authoritative panel" to select publications (other than indexed in WoS) to be covered by the database
- Five categories of outputs: different coverage and philosophy (a) articles in journals; (b) books as author; (c) books as editor; (d) articles or chapters in books; (e) proceedings papers that are not part of special issues of journals or edited books
- Four conditions: (a) be publicly accessible, not necessarily in open access (b) be unambiguously identifiable by ISBN or ISSN number; (c) make a contribution to the development of new insights or to applications resulting from these insights; (d) have been subjected—prior to publication—to a demonstrable peer review process by scholars who are experts in the (sub)field to which the publication belongs. Peer review should be done by an editorial board, a permanent reading committee, external referees or else by a combination of these.

Challenges

There are countries where full coverage starts to exist, but issues are still faced..

- Gaining sufficient political support and funding for achieving systematic data collection in all European countries
- Interoperate RIS, in spite of differences in scope, degree of exhaustivity, typology

Beyond publications

- Criterium of societal impact brought to the fore the question of how to document engagement with society.
- What place for research data?
- What about submitted/ funded research projects as an indicator of activity and excellence? They are proof about scientific activity, even if rejected, time spent in the preparation should be taken into account.
- Flanders: published papers > 4p.
- Lithuania: book = 40000 characters * field coefficient (SSH=8); if you publish 10 pages less, then it is not a book!

Incomplete reporting and auto-censorship related to different factors:

- Technical barriers: HAL, RIBAC vs. Research gate, Academia
- But incomplete coverage in RG, Academia, Google scholars, etc. (almost same coverage biases as WoS)
- Increase of researcher's workload (double, triple declaration), not interoperable
- Lack of institutional incentives
- On-line CVs: exclusion of "not prestigious enough" outputs

Various typologies

- euroCRIS (CERIF specification):
- comprehensive, but incomplete, for example prosopography, footnotes, glossary; excavation report (a very specific document in archeology) as "report"?

- debatable: PhD Thesis/ doctoral thesis authored book (what is an 'unauthored book'?)/ monograph (in France it would be a book written by one author but in the UK can be a book written by up to three authors) Book = ISBN or not? If you are an African scholar and you want to publish in Ghana, there is no ISBN.

Beyond typologies:

- Genre analysis reveals huge discrepancies between products from the same category: i. e. bilingual abstract and keywords are NOT an universal feature
- Quotation and bibliographical habits are not the same in the various SSH disciplines ("art of the footnote")
- Absent metadata: eg. institutional subsidies and their uses. This is also a potential indication of quality.

SSH evaluation and the open access
Positive

- Stimulates the production of metadata (with the above-mentioned problems of standardisation/ mapping)
- Modifies research practices (more cooperation: intra, inter and international cooperation; more articles than books) > changes the symbolic weight of outputs, and even types of outputs to be taken into account in the evaluation. Open publication result in more cooperation, it is linked.
- Changes some evaluations habits (post-publication evaluation)
- Imposes new metrics (altmetrics rather than JIF) but open access still has a limited influence because perceptions remain biased towards hard copies of books (see interviews conducted in IMPRESSH project, France, 2013)

Less desirable effects

- Large offer of "predatory open access journals"
    - Stimulates fake productivity
    - Lowers quality checking ("we publish within a week")
- Costs of open access
    - puncture already limited budgets;
    - pay capability vs. quality?
- A model to be found for books, proceedings and chapter of books.

Contacts

Ioana Galleron is a Senior lecturer in French language and literature. Her research interests are the French theater of the 17th and 18th century, as well as the evaluation of research in the SSH. She is involved in several projects of electronic edition of plays (see http://www.licorn-research.fr/Boissy.html), and in a research group of the consortium CAHIER, dedicated to computer-assisted literary analysis. Since April 2016, she is the Chair of the COST Action CA15137.

Institutional websites: www.enressh.eu & www.evalhum.eu
Twitter @Enressh & @IoanaGalleron
Email: ioana.galleron@evalhum.eu