

SHORT TERM SCIENTIFIC MISSION (STSM) – SCIENTIFIC REPORT

DRAFT 21 FEBRUARY 2018

Action number: CA15137
STSM title: Studying the VIRTAs pilot data: Analysing the (co-) occurrence of journals in the national and regional databases (2014-2015)
STSM start and end date: 26/01/2018 to 12/02/2018
Grantee name: Joshua Eykens

PURPOSE OF THE STSM

The main purpose of this working stay was to explore the data that has been collected for the VIRTAs-ENRESSH pilot. By doing so, I wanted to illustrate the potential for bibliometric explorations and analyses of the Social Sciences and Humanities in Europe. This pilot (VIRTAs-ENRESSH-POC, further abbreviated as VEPOC) is one of the results of the COST Action ENRESSH (CA1537 “European Network for Research Evaluation in the Social Sciences and Humanities). It is an outcome of a collaborative project developed by working group 3 (databases and the uses of data for understanding SSH research). The main aim of which is to investigate the possibilities for a European bibliometric database for SSH research.

As an extension of the Finnish VIRTAs publication information service, the European version would provide a potential solution for a European decentralized system aimed at integration and visibility of data about and for the SSH and other fields of science. In addition, it would provide a complete overview on European research publications - including all types of scholarly publications and potentially other research outcomes in the future as well.

The database that was compiled for the pilot consists of a compilation of bibliographic metadata of 52,948 scholarly publications, from 6 research institutions, and for the period of 2014 – 2015:

- The University of Antwerp (Flanders, Belgium)
- The University of Oslo (Norway)
- The University of Carlos III (Madrid, Spain)
- The University of Helsinki (Finland)
- The University of Jyväskylä (Finland)
- Tampere Technical University (Finland)

The analysis performed during my stay at CSC focused on peer-reviewed articles and the scientific journals they appeared in. The analysis of scientific periodicals (or journals) is being considered as one of the most appropriate ways to study the dynamics of scientific communication. More than books and monographs, journals continuously add knowledge to – and alter our understanding of certain areas of expertise. Questions regarding the ways in which this distribution of newly developed knowledge is happening (or happened) are increasingly becoming a matter of interest for scholars interested in the history, the sociology, and science of science (e.g. scientometricians).

More specifically than, I wanted to study the (co-)occurrence of journals in the different datasets that were compiled in the VIRTAs database. This undertaking will result in a report. In the end, expectations are that the analysis of the data will result in a publication on the use of bibliometric indicators in the context of VIRTAs.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSM

My two-week working visit at CSC can be divided into three parts: (1) preparation of the data, (2) analysis, (3) and reporting of findings.

- (1) The preparation of the dataset was done during the first 3 days of the stay (29th of January – 31st of January). First, I tackled the problem of missing or ‘problematic’ codes (day 1). Second, I enriched the dataset with additional data sources that contain meta-data and information about the scientific journals (day 2). After the merging of these different datasets, a manual check was done, and the final dataset was cleaned once more (day 3).
- (2) During the second part, which took place during days 4 to 7 (1st of February – 2nd of February and 5th of February – 6th of February), I performed different statistical analyses to make an inventory of the different variables (day 4 and 5) and describe the characteristics of dataset (day 6 and 7).
- (3) Part three mainly consisted of writing up the results of the analyses and tackling additional problems with missing or ‘problematic’ codes which were encountered during the actual analysis. The writing of the report was done during days 8 to 10 (7th of February – 9th of February).

The working procedure and results were frequently discussed with the staff members working on the VIRTAs pilot:

Mon 29 th CSC	Tue 30 th TSV	Wed 31 st CSC	Thu 1 st CSC	Fri 2 nd CSC
9:00 Start at CSC 13:00 – 14:00 Get-together-meeting with VIRTAs people	11:00 – 13:00 Working at TSV with Janne Pölönen 13:00 – 16:00 Meeting with CSC at TSV			9:30 - 10:00 Introduction of work to Research Information Management group at CSC

Mon 5 th CSC	Tue 6 th CSC	Wed 7 st CSC	Thu 8 th CSC/TSV	Fri 9 th CSC
12:00 – 13:00 Meeting with VIRTA people			15:30 - 17:00 Wrap up meeting at TSV with Janne Pölonen and CSC	

DESCRIPTION OF THE MAIN RESULTS OBTAINED

The original aim of this project was to map to what extent scientific periodicals co-occur in different institutional bibliographic databases. That is, we wanted to be able to understand if (and how) the different institutions involved made use of the same publication fora to communicate their research. A key finding in this respect was that publications in social sciences and humanities appear to be very scattered over different journals and disciplines. Almost 60% of the journals contained only one publication of authors affiliated to one of the institutions.

Additionally, articles classified as SSH research did not only appear in journals dedicated to SSH research. 10% of the journals in the list are associated to either natural sciences, engineering and technology studies, medical and health sciences, or agricultural sciences. Journals mainly related to social science research and humanities make up 59.3% and 30.8% of the total list. The fact that humanities journals make up less than a third of the journal list relates to the fact that scholars working in these fields are more inclined to book publishing than their colleagues working in fields related to social sciences. Another possible explanation for this difference, could be the number of (FTE's) researchers affiliated to the research groups that work in a certain discipline.

When we took a closer look at the different languages preferred by journal editorial teams, we could see that English is the most important one (almost 70% of the journal list). We are looking at a European dataset, so it might be no surprise that journals allowing articles written in different languages (category: multiple languages) follows next. Scandinavian, Dutch, and Spanish language journals are also quite prominent. This is, ofcourse, one of the artefacts of the dataset studied.

This difference in language seems to be discipline specific. A closer look at the 'hard' sciences shows us how English language journals dominate in these fields – that is for the case of our data, of course. For social sciences and humanities, the picture is quite different. In the case of social sciences, this share of 90% plus drops to a total that is slightly higher than 70. When we break up the generic category 'social sciences' into different disciplines, again a quite variable picture emerges. Journals related to psychology, for instance, are in 90% of the cases English language journals. Whereas for law, English language journals account for less than half of the total share (40.3%). For journals related to fields in the humanities, the share of English language journals drops even further: here they represent slightly more than 50%

A comparison of the different institutions also delivered some interesting insights. The language of the journals preferred for publication is an obvious difference. The disciplinary productivity (measured) as the total output of research articles than, also differs between the institutions. Whereas we can see that the social sciences are more than humanities in all institutions (at least when it comes to article publishing), we

can observe a remarkably high share of research in social health sciences conducted by authors affiliated to Helsinki University. The University of Antwerp seems to host a very strong tradition in language and literature studies.

FUTURE COLLABORATIONS

Future research guided by the data gathered by implementing the European publication service would allow for more detailed bibliometric research. Besides that, there will also be new possibilities for empirical studies into the sociology and history of science, and more particularly European SSH. The 'new' mode of knowledge production (mode 2), for example, that has manifested itself throughout Europe over the past two decades perhaps plays an important role in what we have observed so far.

It has been noted by Nowotny et al. (2003) that one of driving forces behind this transformation of knowledge production (i.e. a more specialized and problem-oriented way of doing research) was the turn towards a top-down coordination of research priorities. In Europe this is taking place at different levels. The European Union, nation states, and university level research councils all increasingly emphasize thematic research programmes.

Consequently, two characteristics should become observable. First, 'knowledge is generated within a context of application', and, second, knowledge is increasingly becoming problem-oriented (Nowotny et al., 2003, 186). These new characteristics stand in contrast with the old process, wherein problems arose from within the theoretical/experimental debates in 'pure' scientific fields. The second characteristic (more problem-oriented research) increasingly turns 'traditional' disciplinary boundaries into cross-roads. This transformation of knowledge production can also be observed in the journal landscape.

First, in tandem with the expansion of academia, there is an ongoing drive towards specialization, which also institutionalizes itself through journals dedicated to SSH research. One of the manifestations of this process can be observed through the emergence of new journals dedicated to different specialities, or specific societal problems. Whereas the more traditional and 'canonical' scientific periodicals were dedicated to disciplinary streams of knowledge production, new 'problem oriented' journals are on the rise. Because of this growth of available publication fora on the one hand, and the differentiation of these periodicals into (new) (sub)disciplines and specialities on the other, a de-concentration of articles is an obvious result.

Second, national language journals appear to remain of importance for almost all fields in social sciences and humanities. Although it is clear (also from our dataset) that English is indeed the new lingua franca, a good deal of research by scientists affiliated to the different institutions is being published in national language outlets. The authors affiliated to Helsinki University, for example, published almost 15% of their articles in Finnish language periodicals. The latter are, on their behalf, less used by Dutch or Spanish speaking authors working at Flemish or Spanish research institutes.

To get a deeper understanding of these relatively new processes that are changing the journal publishing landscape in Europe, longitudinal studies and cross-national comparisons are necessary. On the one hand, the VIRTAs project would facilitate this kind of research. When more countries or institutions decide to join in

on the project, the amount of available data will automatically increase. On the other hand, comparisons between publication patterns in Flanders and Finland also form an interesting opportunity in this light.