

# **Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project**

A report produced for DFG, ESRC, AHRC, NWO, ANR and ESF

8 March 2010

# **'Towards a Bibliometric Database for the Social Sciences and Humanities'**

## **Project Board Members and Affiliations**

Prof. Ben Martin (SPRU, University of Sussex), Chair of Project Board

Dr. Puay Tang (SPRU, University of Sussex), Project Secretariat (Manager)

Molly Morgan (SPRU, University of Sussex), Project Secretariat

Prof. Wolfgang Glanzel (University of Leuven)

Prof. Stefan Hornbostel (iFQ)

Prof. Gerhard Lauer (University of Göttingen)

Prof. Gerard Lenclud (College de France)

Prof. Luisa Lima (ISCTE, Member of the ESF  
Standing Committee on the Social Sciences)

Prof. Charles Oppenheim (University of Loughborough)

Prof. Peter van den Besselaar (Rathenau Institute)

Prof. Milena Zic-Fuchs (University of Zagreb, Chair of the  
ESF Standing Committee on the Humanities)

# Table of Contents

EXECUTIVE SUMMARY .....	I
INTRODUCTION .....	1
PART A .....	5
RECENT BIBLIOMETRIC DEVELOPMENTS IN THE SOCIAL SCIENCES AND HUMANITIES .....	5
THE ROLE OF BIBLIOMETRIC INDICATORS IN RESEARCH ASSESSMENT EXERCISES .....	7
STRENGTHS AND WEAKNESSES OF SSH BIBLIOGRAPHIC DATABASES AND LISTS ....	8
PART B .....	12
CREATING THE SSH BIBLIOMETRIC DATABASE: KEY OPERATIONAL AND STRATEGIC CONSIDERATIONS .....	12
UNDERLYING CONSIDERATIONS.....	12
OPERATIONAL ISSUES.....	13
STRATEGIC OPTIONS FOR DEVELOPMENT .....	17
PART C .....	20
POTENTIAL APPROACHES FOR CONSIDERATION.....	20
SUGGESTION 1.....	21
SUGGESTION 2.....	22
SUGGESTION 3.....	24
SUGGESTION 4.....	25
SUGGESTION 5.....	26
SUGGESTION 6.....	27
PART D .....	33
RECOMMENDATIONS FOR THE DEVELOPMENT OF A COMPREHENSIVE SSH BIBLIOMETRIC DATABASE.....	33
RECOMMENDATION 1: DEFINE THE CRITERIA FOR INCLUSION OF SSH RESEARCH OUTPUTS AND ESTABLISH A STANDARDISED DATABASE STRUCTURE FOR NATIONAL BIBLIOMETRIC DATABASES .....	35

<b>RECOMMENDATION 2. EXPLORE THE OPTION OF INVOLVING A COMMERCIAL SUPPLIER IN THE CONSTRUCTION OF A SINGLE INTERNATIONAL SSH BIBLIOMETRIC DATABASE.....</b>	<b>41</b>
<b>RECOMMENDATION 3. CONDUCT A PILOT STUDY ON ONE OR MORE SELECTED SSH DISCIPLINES .....</b>	<b>43</b>
<b>RECOMMENDATION 4. LONGER-TERM EXPANSION AND ENHANCEMENT OF THE SSH BIBLIOMETRIC DATABASE .....</b>	<b>44</b>
<b>SELECTED REFERENCES .....</b>	<b>48</b>
<b>ANNEX 1.....</b>	<b>49</b>
<b>ANNEX 2.....</b>	<b>50</b>
<b>ANNEX 3 – ROADMAPS.....</b>	<b>51</b>

## Executive Summary

The aim of the project was to explore the possibility of developing a database for capturing the full range of research outputs from the Social Sciences and Humanities (SSH). SSH research outputs include not just those articles published in international journals, but also articles in national journals, academic book chapters and books, books aimed at a more popular audience, monographs, reports in the 'grey literature', and non-published outputs from fields such as the performing arts.

Such a database is intended to serve a number of related purposes. One is to address the growing pressure from policy-makers and research funders to demonstrate 'accountability' and to ensure 'value for money'. A second is to develop performance measures for assessing research quality and impact. For basic research in the natural sciences, there are fairly well established indicators of research output and impact (based on publications and citations), but these work poorly for the social sciences and barely at all for arts and humanities.

Third, research funders and others may want to use the database to provide an overview of SSH research outputs in Europe. Fourth, funders and policy makers may use it or as a source of information of to identify areas of strong research capacity and those that are perhaps in need of capacity-building or support. Lastly, research councils<sup>1</sup> may seek to use the database as a tool for mapping emerging areas of (often interdisciplinary) research.

The question addressed in this project is, 'What is the potential for developing some form of research output database that could be used for assessing research performance in SSH?'

An important aspect to the background context of the study relates to recent developments with respect to various databases and bibliographic lists on which a SSH bibliometric database might build. These include the growth of 'Open Access' publications, significantly improved coverage of journal literature by the Web of Science and

---

<sup>1</sup> In what follows, we use the generic term 'research council' to include all research-funding agencies.

Scopus, the emergence of Google Scholar as a new source of publication and citation data for books as well as journals, the growing availability of digital data on publications from book publishers, the development of various national or disciplinary bibliographic databases, and the establishment of a range of institutional repositories for research outputs.

The report identifies the main problems and issues to be confronted in any attempt to construct an inclusive SSH bibliometric database. It analyses a number of key considerations for the creation of an inclusive SSH bibliometric database, along with certain operational issues. On this basis, it sets out various strategic options.

The report concludes with a number of recommendations chosen on the basis of their practicality and cost-effectiveness. We propose a way forward based on four main recommendations. The first centres on the definition of criteria as to which SSH research outputs should be included in a bibliometric database and the establishment of a standardised database structure for national bibliometric databases. The second explores the option of involving a commercial supplier in the construction of a single international SSH bibliometric database. In both of these, the focus is on published scholarly outputs appearing either in peer-reviewed journals, or in books that have likewise been subject to peer-review before publication. The third recommendation involves conducting a small pilot study focusing on one or more specific SSH disciplines. The fourth recommendation deals with the longer-term expansion and enhancement of the SSH bibliometric database to include other SSH outputs (that is, in addition to peer-reviewed articles and books). A hybrid approach (i.e. a combination of 'top-down' and 'bottom-up') has been recommended for the implementation of each recommendation because this appears to offer the best of both worlds – impetus, guidance and authority from the top, and expertise and experience from the bottom.

The report argues that these four recommendations offer the most promising way forward in exploring and then establishing an international bibliometric database for the social sciences and humanities. As the approach outlined involves various stages, it provides the opportunity to curtail the process at any point if the

problems prove to be intractable or excessively expensive to overcome. It also suggests that the prospects of success will be greater if, for pragmatic reasons, the initiative starts with a relatively small group of research councils and countries, allowing others to join in subsequently as momentum builds and as the necessary resources become available.

# **Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project**

## **Introduction**

The primary aim of this Scoping Project was to investigate the possibility of developing a comprehensive database for capturing and assessing the full range of research outputs from the Social Sciences and Humanities (SSH), preferably including non-published research outputs from fields such as the performing arts.

Such a SSH bibliometric database is intended to serve five main purposes. The first is to permit the construction of indicators needed to demonstrate accountability with regard to the public funds devoted to research. The second is to develop indicators for assessing research excellence. Third, policy makers, research funders and others may wish to use the bibliometric database to provide an overview of SSH research outputs in Europe. Fourth, funders may use it as a means for assessing research capability and for identifying areas in SSH that may require capacity-building. A research output database might permit a 'portfolio' analysis of SSH research, contributing to decisions on resource allocation, particularly in the Humanities where there is often little relevant information. Fifth, research councils may use the information provided by the bibliometric database to map emerging areas of (often interdisciplinary) research with an aim of ensuring that they are adequately resourced.

In recent years, the European Commission has underscored the importance of developing public information systems on higher education institutions, including data on research performance, as a vital part of the emerging research infrastructure of the 'European Research Area'. The availability of a bibliometric database for SSH would thus represent an essential component of this infrastructure.

This report analyses the main obstacles that would have to be overcome in developing such a bibliometric database. It also examines developments with regard to a range of bibliographic and well as bibliometric databases, suggesting how a SSH bibliometric database might build upon these developments. It concludes with a



number of recommendations for how such a project might be taken forward to the next stage. These are based on ideas developed by the Project Board Members, on discussions at two international workshops of invited experts, and on two commissioned studies, one from Prof. Diana Hicks and Jian Wang (Georgia Institute of Technology, U.S.) (see Annex 1) and the other by Henk Moed (CWTS Netherlands) and Felix de Moya (SCImago Research Group-CSIC, Spain) (see Annex 2).

In order to avoid confusion, it should be stressed right at the outset that throughout this report we have adopted a relatively broad definition of 'bibliometrics', one that goes well beyond 'just citations'. The aim here is to establish whether it is feasible to construct a database covering the full range of SSH research outputs to help fulfil the five purposes outlined above. For this we use the term 'bibliometric database'.

Ideally, what is required is a database that brings together, in a consistent and comparable form, data on the main research outputs of SSH (i.e. the number, kind and quality of the outputs) and also provides an indication of the impact of those research outputs not only on fellow academic researchers but also more widely (whether in the form of economic and social impact, or impact in terms of enlightening the general public).

To achieve this, one would need a reasonably inclusive bibliometric database that encompasses different forms of research or scholarly output from SSH – i.e. one that includes published articles in international and national journals, book chapters, monographs and books, and other non-published and non-textual research outputs. Some of these data may be obtained from high-quality bibliographic databases, so these are likely to be an essential building block for the construction of an inclusive SSH bibliometric database. However, many bibliographic databases currently lack the data needed to enable that database to be used for bibliometric purposes, such as the institutional addresses of all the authors in a consistent and comparable form.

Thus, an inclusive database suitable for use in the bibliometric analysis of SSH will differ substantially from existing bibliometric databases like the Web of Science and Scopus, which consist

primarily of scholarly journal articles published in international journals, and mainly written in English. These two databases involve strict data-collection and verification protocols, enabling them to be used for the construction of various bibliometric indicators. For example, citation data are often used to assess the scholarly impact of published research outputs, as required in the research performance assessment exercises that are increasingly being introduced by national governments worldwide. However, neither database is currently suitable for assessing SSH research.

In the recommendations suggested in this report, we have attempted to address the main problems that the construction of a more inclusive SSH bibliometric database poses. These problems include:

1. The scale and variety of research outputs from SSH. Unlike in the Sciences, in the Social Sciences and the Humanities we need to include a much wider range of outputs, such as books and book chapters, more 'popular' books and articles aimed at the general public rather than academic peers (or 'enlightenment literature'<sup>2</sup>, as we term it here), 'grey' literature such as policy reports, as well as (for some fields at least) research outputs with a non-textual content. Thus an inclusive SSH database must allow for variety in the range of indicators that may eventually be constructed from the database.
2. The need to consider national journals and research outputs (in particular, those published in languages other than English). However, this raises the issue of what criterion (or criteria) should be used in determining which research outputs

---

<sup>2</sup> 'Enlightenment literature' is the term used here for publications aimed primarily at the general public rather than academic peers: "Enlightenment literature represents knowledge reaching out to application and is found in periodicals whose goal is knowledge transfer or 'enlightenment' of non-specialists" (Hicks and Wang 2009, p. 4). The authors also refer to a study by Burnhill and Tubby-Hille, which found that in the UK "projects in education [were] reaching practitioners through the *Times Education Supplement*, with researchers in sociology, social administration, and socio-legal studies publishing in such periodicals as *New Society* and *Nursing Times*" (see Burnhill and Tubby-Hille, 2003). In Norway, Kyvik found that one-half of social scientists published contributions to public debates (see Kyvik, 2003, as quoted by Hicks and Wang, p. 4).

should be included and which excluded. As one of the commissioned reports points out, there is a need for “a consistent, evidence-based criterion for journal scholarly quality”, in particular, one “that can be applied impartially and without favouritism across the range of European languages ... [This] will be crucial to building a respected bibliometric infrastructure for SSH” (Hicks and Wang, 2009, p. 12).

3. The highly variable quality of existing SSH bibliographical databases and lists, from which the new bibliometric database is likely to draw substantially. This variability reflects the uncertainty and inconsistency inherent in the quality criteria currently used to select entries for existing bibliographical databases, as well as variations in the editorial standards of the databases themselves.
4. The lack of a standardised database structure for the input data. It is vital to ensure consistency of data fields, a task made all the more difficult here by the fact that, for published outputs, the publication and referencing characteristics of those outputs vary widely across SSH disciplines. Moreover, it is important to keep in mind the non-published research outputs that need to be considered in any eventual database.

The Report is structured as follows. **Part A** provides overviews of (1) recent developments in two of the main bibliometric databases, the Web of Science and Scopus; (2) the role of bibliometric indicators in research assessment exercises; and (3) existing SSH databases and lists, including brief descriptions of their characteristics and problems or limitations with regard to extending them to use for bibliometric purposes. **Part B** sets out the main issues to be resolved in establishing a possible ‘road map’ to the creation of an inclusive bibliometric database, in particular focussing on (1) the key underlying considerations; (2) the operational issues; and (3) strategic options for development. **Part C** begins to bring everything together and presents a range of potential approaches for the construction of the SSH bibliometric database. **Part D** concludes with a number of Recommendations, which have been synthesised from the various approaches discussed in Part C.

## **PART A**

### **Recent bibliometric developments in the Social Sciences and Humanities**

There are a number of interesting developments currently being undertaken by the Web of Science and Scopus to expand their coverage of Social Sciences and Humanities (SSH) journals. The Web of Science (now published by Thomson-Reuters, but previously known by its constituent parts, the Science Citation Index, the Social Sciences Citation Index, and the Arts and Humanities Citation Index) has increased the number of SSH journals it covers from 1,700 in 2002 to 2,400 in 2009. As from 2009, the Web of Science journals includes 1,200 'regional' journals, defined as "journals that typically target a regional rather than international audience by approaching subjects from a local perspective or focusing on particular topics of regional interest" (Moed et al., 2009, p. 29).

With regard to Scopus (which is produced by Elsevier), a key development is the addition in June 2009 of 1,450 journals, which takes its collection to 3,500 SSH journals (i.e. nearly 50% more than the Web of Science). Moreover, Scopus is starting to add bibliographic meta-data on highly cited books (in particular, data on the full title of the book, the publisher, all the authors and their institutional affiliations). This is important because in some SSH fields a very substantial portion of the published research output consists of books and book chapters rather than articles in journals covered in the Web of Science. These non-'citation-indexed' books and chapters are often well cited in articles in journals scanned by the Web of Science (CWTS 2007, p.48). This illustrates the vital importance of including books and book chapters as source records in a comprehensive SSH bibliometric database.

For several decades from the 1960s onwards, the sole source of bibliometric data was the Science/Social Sciences/Arts and Humanities Citation Index (then published by the Institute for Scientific Information), which later became the Web of Science. However, this situation changed in 2004 with the appearance of Scopus, which provides publication and citation data for a

somewhat larger number of journals for 1996 onwards. It would appear that there is currently an element of competition between these two main bibliometric database providers as to which will be seen as providing the more comprehensive SSH database. Furthermore, the effort by Scopus to include books signals a move to be more inclusive with regard to the full range of published research outputs from SSH. The Web of Science's recent inclusion of regional journals reflects a similar desire. It is quite possible that competition between the two established database providers may result in further expansion and inclusion of other SSH research outputs.

The third and newest competitor to the Web of Science and Scopus is Google Scholar (along with Google Book Search<sup>3</sup>). According to Moed et al., Google Scholar, along with Web of Science and Scopus, is a bibliographical database that facilitates "desk-top or poor man's bibliometrics" (Moed et al., 2009, p.19). This database is the only one of the three that currently covers books and book chapters extensively, and it provides simple indicator data, such as numbers of citations. It is relatively easy to find books in Google Scholar because it has full-text indexing.<sup>4</sup> Together, Google Scholar and Google Book Search apparently scan millions of books. However, it remains to be seen whether Google will develop this extensive database into a fully-fledged source for bibliometric analysis. At present, Google fails to provide clear information on what is, and what is not, covered in the database. As a result, their records are not presently usable for systematic, rigorous bibliometric analysis, and there is a concern about the accuracy of their citation links.<sup>5</sup> However, with the continuing developments in software applications

---

<sup>3</sup> Google Book Search scans books from a range of sources, including digital repositories, and enables users to access and read extracts from them.

<sup>4</sup> For example, when B.D. White searched for material on Gabriel Plattes, a 17th century utopian and scientific author; in Google Scholar and JSTOR (also full text), he found 50-60 articles, while in the Web of Science, which is bibliographic rather than full text, he found less than five (see B.D. White (2006), 'Examining the claims of Google Scholar as a serious information source', *New Zealand Library and Information Management Journal*, 50(1), 11-24, as quoted in Hicks and Wang (2009), p. 8).

<sup>5</sup> See Hicks and Wang (2009) and Moed et al. (2009).

and advances in computer science, there is certainly the potential to overcome these concerns.

## **The role of bibliometric indicators in research assessment exercises**

Research assessment exercises are increasingly being undertaken by national agencies and individual research institutions worldwide. There are a number of reasons for conducting these assessment exercises, including: (a) the evaluation of research excellence; (b) the adoption of a funding formula to distribute funds between universities and/or research institutions; (c) ensuring accountability with regard to the use of public monies (and assessing the return on investment of public research funds); (d) as a marketing mechanism in the competition for prospective students; and (e) as a benchmarking tool employed by higher education and research institutions.

Many of these exercises rely, in part, on the Web of Science or Scopus for the creation of citation impact indicators. While there is merit in using citation impact indicators, there are problems with a sole reliance on these databases, notably the fact that these bibliometric databases have not comprehensively included books. As noted above, the publication characteristics of the various disciplines in SSH vary significantly. In history, for instance, books can account for as much 'impact' as an economics article in the *American Economic Review*. Thus, while they are more than likely to appear in institutional and national bibliographical databases or lists, books are still under-represented in the current two main bibliometric databases. The problem is not just confined to existing bibliometric databases, but extends to bibliographical lists as well, many of which are of varying quality. All of this poses problems for the robust development and use of citation impact indicators from existing databases. In the next section, we provide an overview of the strengths and weaknesses of the main databases and lists.

## **Strengths and weaknesses of SSH bibliographic databases and lists**

It is worth stressing here the key difference between a bibliographic and a bibliometric database. The main aim of a *bibliographic* database is to aid in literature retrieval (as opposed to assessing research outputs). However, there are some bibliographic databases that allow for a degree of structured bibliometric analysis. For example, ECONLIT, Sociological Abstracts and Psycinfo record author affiliations and cite references, thereby providing some of the data needed for bibliometric analysis. However, these databases all focus primarily on journal articles.

*Bibliometric* databases, in contrast, are used expressly for measurement applications, in particular for research assessment purposes. These databases include details of the references cited, and contain full institutional and author details. They also permit the creation of rather more sophisticated indicators (for instance, citation totals, the average number of citations per publication, numbers of highly cited publications, and the Hirsch index or 'h-index') that can be used to help assess the impact of a body of research output.

The main shortcomings of current SSH bibliographic databases or lists, as identified by Moed et al. (2009), are:

1. a lack of standardisation of author names and institutional affiliations, including the fact that many bibliographic databases list the corresponding address of the first author only;
2. a lack of cited references in source publications;
3. a failure to list all the authors of a multi-authored source publication;
4. differing quality in terms of data capture;
5. uncertainty with regard to the quality criteria used in selecting which outputs to include in (and which to exclude from) the database, which highlights the necessity of robust selection criteria;
6. errors, for example with regard to journal status, including the inclusion of journals that are no longer published, are

suspended or are published irregularly, and journals whose status is unknown.<sup>6</sup>

A few themes are worth elaborating on with regard to these points. Firstly, points 1-4 highlight the necessity of a standardised database structure, including a standardised set of data fields, if the database is to be used for bibliometric purposes. According to Moed et al. (2009), the family of SSH databases within CSA-Illumina<sup>7</sup> exhibit some degree of standardisation but several bibliographic SSH databases are not part of this group. Furthermore, again we find that books are inadequately represented in these databases.

Secondly, an essential condition for a robust bibliometric database is the scholarliness and accuracy of its contents. Yet according to Hicks and Wang, the Web of Science and Scopus journal lists both exhibit certain problems in this respect, in particular, that there is a significant level of non-scholarly literature contained within them. The authors analysed three other Social Sciences and/or Humanities journal lists – the 'Norwegian reference list' (developed at NIFU-STEP, Norway), the European Reference Index for the Humanities (ERIH), and the Australian ERA Humanities and Creative Arts list (ERA HCA)<sup>8</sup> – to see if the same problem existed in these lists. A brief description of each list is given below before discussing Hicks and Wang's findings.

The 'Norwegian reference list' is the list of journals accepted by and submitted in the Norwegian research evaluation process.<sup>9</sup> The list covers all fields of science, social sciences and the humanities. Scholarly publications are defined as "presenting new insights in a form that allows the research findings to be verified and/or used in new research activity in a language and with a distribution that makes the publication accessible for a relevant audience in a publication channel with peer review" (Hicks and Wang 2009, p. 6).

---

<sup>6</sup> See Moed et al. (2009) and Hicks and Wang (2009) for a comprehensive list of SSH bibliographic lists and their details.

<sup>7</sup> CSA Illumina provides access to more than 100 full-text and bibliographic databases. The databases cover Arts and Humanities, Social Sciences, Natural Sciences and Technology.

<sup>8</sup> For more details, see Hicks and Wang (2009, pp. 18-20).

<sup>9</sup> For more details on, and problems with, the Norwegian reference list, see Hicks and Wang (2009, pp. 19-20).



Hicks and Wang analysed the Social Sciences and Humanities journals on this list.

The ERIH list claims to cover top-quality European Humanities research published in academic journals in English and non-English languages (including national journals). The list is peer-reviewed by 15 European expert panels, who select and aggregate input received from funding agencies, subject associations and specialist research centres across Europe. A main aim of the ERIH is to enhance scholarly outputs in the Humanities.

The Australian ERA HCA (Excellence in Research for Australia Humanities and Creative Arts) contains 19,500 unique peer reviewed journals to form a draft list of ranked journals. Each journal has a single quality rating and is assigned to one or more disciplines, and the list has been reviewed by discipline-specific experts “to strengthen sector confidence in the accuracy of the journal rankings.”<sup>10</sup> Recently, a range of performance indicators has also been created for 136 disciplines, each of which may choose the indicators that are most appropriate for them. For instance, the Social Sciences have selected a mix of quantitative and qualitative indicators, such as citations and peer review, while the Sciences, not surprisingly, have selected quantitative indicators. Such an approach has gained widespread acceptance from the Australian academic community.

Hicks and Wang’s analysis provides some interesting insights into the claim by all five databases/lists to be based solely on “scholarly refereed material”. For instance, the Web of Science was found to contain 16% of non-refereed and 4% of non-scholarly journals (out of 2,600 SSH items), while Scopus contained 32% of non-refereed journals and 12% of non-scholarly journals, and ERIH 43% of non-refereed and 10% of non-scholarly journals (out of the initial 3,900 humanities journals examined).<sup>11</sup>

---

<sup>10</sup> [http://www.arc.gov.au/era/era\\_journal\\_list.htm](http://www.arc.gov.au/era/era_journal_list.htm)

<sup>11</sup> Hicks and Wang (2009) only analysed the initial ERIH list containing the 3900 journals verified in Ulrich’s, and they adopted Ulrich’s definition of what constitutes a ‘refereed’ or ‘non-refereed’ journal. The analysis does not address all the 5200 journals that ERIH now covers, many of which are not published in English.

In summary, the above analysis suggests that no bibliometric database or bibliographic database is 'perfect'. Moreover, despite the widespread criticism of the Web of Science for being too 'Anglo-centric', it remains widely used in many research assessment systems primarily because the articles published in its indexed journals are seen as having reached an internationally recognised standard. As Hicks and Wang (p. 7) note, "Journal editors feel it an honour to meet the criteria for inclusion in WoS [Web of Science]".

## **PART B**

### **Creating the SSH bibliometric database: Key operational and strategic considerations**

In order to establish a possible 'road map' for the creation of an inclusive SSH bibliometric database, we first need to address a number of key issues. In what follows, these have been divided into (1) key underlying considerations; (2) operational issues; and (3) strategic considerations. Examination of these issues will serve to clarify the challenges that the development of a SSH bibliometric database faces, and the range of options that may be adopted to construct such a database are provided in Part C. Careful consideration and selection from among these options will help to facilitate a speedier and more effective implementation, as presented in the Recommendations in Part D.

#### **Underlying considerations**

There are a number of central issues to consider prior to the creation of a SSH bibliometric database. These include the following:

1. The need to raise awareness among research funders, policy-makers and others of the significant length of time required for the development of a SSH bibliometric database, in the same way that bibliometric databases for Science, Technology, Engineering and Medicine (STEM) subjects have evolved over a period of many years.
2. A SSH bibliometric database must allow considerable flexibility in terms of coverage. While initially it may, for pragmatic reasons, focus on scholarly articles and books, over time it will need progressively to bring in more popular books, magazine or newspaper articles and other 'enlightenment literature', 'grey publications' such as policy reports, and (ideally) details of non-published outputs like artwork, exhibitions, excavation reports and photos for assessing SSH impact. Such a database should also allow the creation of different indicators to serve other purposes than those

specified in this Report. Likewise, those indicators should be such that they can be used by individuals, groups, research organisations, etc. for their own assessment or other measurement purposes. Consequently, the imposition of a STEM-like bibliometric database focusing mainly on international journal articles cannot be a solution because of the very different communication modes of SSH researchers.

3. The practicality of attempting to build a SSH bibliometric database from bibliographic lists of institutional and national repositories. This will require that bibliographic lists containing SSH research outputs from publicly funded research and those published in national journals are compiled in a consistent form and are made available to the 'creators' of the new database.

## Operational issues

A number of operational principles need to be considered here. One of the most important of these centres on whether to adopt a top-down approach, or a bottom-up approach, or some combination of the two. Let us consider this issue first (addressed in points 1-3 below), before examining other important operating principles (points 4-7 below).

1. **A top-down approach.** This would involve either creating a database at the European (or some other international) level or strong central coordination of national organisations with the establishment of standardised rules in order to ensure full comparability of nationally provided data. A decision is also needed as to who should initiate this process.
2. **A bottom-up approach.** This would entail the producers of existing national bibliographic databases and lists working together to develop common rules and procedures that would result in their respective databases becoming more comparable and, in due course, capable of being integrated in some form. To achieve this would require that the compilers of such national databases work very closely together to ensure convergence towards common standards. As with the above, a decision has to be made for starting the process.

Input from the SSH scholarly communities with regard to the coverage (i.e. the range of research outputs) of these bibliographic databases is also vital in order to ensure full disciplinary involvement and support.

3. **A hybrid approach.** A third possible approach might involve some combination of the above two approaches. For example, a supra-national European organisation might begin by developing a 'bibliometric manual' that would set out the requirements of a SSH bibliometric database, including appropriate definitions, what data are required and in what form, systematic criteria for determining what types of research output should be included and excluded, and so on. The compilers of national bibliographic databases would then be invited to supply data according to those common conventions. Here, too, a decision on who should start the process should be undertaken.

It is worth noting here a possible analogy with the establishment of the first truly comparable data on R&D funding in the early 1960s. In this case, OECD took the lead and, working with international experts, drew up the first 'Frascati Manual', which set out definitions for what was to be included and what data were required. Over time, national bodies made the necessary improvements to their data-gathering processes, and hence the quality and comparability of the data supplied by member states improved. Gradually, more and more countries have come to produce their R&D data according to the Frascati Manual, driven by the incentive that they can then make comparisons with other countries.

In the case of all three of the above approaches, the following questions will need definite answers before a plan of action can be undertaken:

- Who should decide which approach is optimal? Who should initiate the decision-making process?
- Should a European organisation set the standard for the structure of the database?
- What would be the role of European national research councils in standard setting? Would they work together,

or should they be responsible for identifying the European organisation that is going to oversee the task?

- Should an existing international standard be considered, which allows for expansion to include the other inputs that are required for an inclusive SSH database?<sup>12</sup>
- Who is to be responsible for maintaining the database (as this implies a long-term commitment of resources)? Should there be collective funding from national research councils or should the funds come from the European Commission?

4. If they are to provide an input to the European bibliometric database, bibliographic databases/lists need to be able to demonstrate that they include **high-quality national research outputs** that have been validated at a national (or even international) level by leading academics and bibliographic experts. To achieve this, it will again be essential that the respective database compilers talk extensively to each other in an attempt to ensure full comparability of their respective lists of SSH outputs.
5. An essential step is the establishment of a **basic threshold criterion** (or set of criteria) for determining which SSH research outputs are of sufficient quality or importance to merit inclusion. One starting point is consideration of the various criteria currently adopted by national institutions across the various SSH disciplines. (An alternative would be to adopt a 'liberal selection policy' in which SSH academics include as many research outputs as they see fit.) However, the great variety of criteria (both explicit and implicit) currently in use (or new ones created, for example, as a consequence of adopting a liberal selection policy) means it is likely to prove difficult to reach some consensus among the disparate European research councils, institutions and academics. Equally importantly, the pursuit of such a consensus may well delay the start of the SSH bibliometric

---

<sup>12</sup> Such a database can be used by governments for metrics and research evaluation without sacrificing the other components in the database.

database. In view of this, it may be sensible to proceed in stages, beginning with a relatively short and simple set of criteria for a range of clearly specified scholarly outputs, and then expanding this with further quality criteria once the initial SSH bibliometric database has been constructed. *In the light of these considerations, a basic or minimum threshold criterion could focus on initially on **scholarly articles in peer-reviewed national and international journals**<sup>13</sup>, and on **scholarly books that have been subject to a peer-review process.***

6. In developing a SSH bibliometric database, we will need to carefully **monitor the various impacts or consequences**, both intended and unintended, on the research process. Any attempt to introduce performance indicators of whatever type may have undesirable effects in terms of influencing what research is undertaken and the kind of outputs that are produced. For example, the use of publication counts in Australia as part of the formula used in distributing research funds to universities resulted in a proliferation of publications in lower quality journals. To avoid this, one may well need to distinguish between higher and lower quality research outputs (as the Norwegians have done in their research assessment process) as well as having threshold criteria for determining the minimum quality needed to be included in the database (see point 5 above). Similarly, an initial focus on international scholarly outputs could result in researchers strategically changing their publishing behaviour to the detriment of production of national language and enlightenment outputs. The risk of such an unintended consequence has to be weighed against the benefit of adopting a pragmatic phased approach to the construction of a SSH bibliometric database.
7. Ideally, there would be merit in commissioning one or more studies on the development of SSH quality or impact

---

<sup>13</sup> We are aware that a few high-status journals in the humanities (for example, in philosophy) do not operate a formal peer-review system. However, their editors may instead operate a more informal review system in consultation with colleagues, which may still constitute a form of 'peer review' and therefore entitle them to be included.

indicators. However, in the light of several available and ongoing studies on this issue, it may not be a priority to commission yet another one alongside the implementation of the SSH bibliometric database. Nevertheless, it is **essential that experts on impact assessment techniques and methodologies should be fully consulted during the development of a SSH bibliometric database**. To undertake the development of the database in isolation from studies on these techniques, particularly for the SSH, would be to overlook an important potential contribution to the formulation of appropriate performance and impact indicators for the SSH bibliometric database.

## **Strategic options for development**

Once decisions have been made on the operational issues, there are various strategic options to be considered. These include:

1. Deciding whether the new SSH database should be developed by a European agency or by national governments (through national research councils working with their respective research institutions). Who should make this decision? Alternatively, perhaps a group of research councils, as illustrated by the examples of HERA and NORFACE (ERA-NET projects for the Humanities and for the Social Sciences, respectively) could be considered to spearhead the development.<sup>14</sup>
2. Determining whether, in the light of the commercial competition between existing database providers, one of these might be approached and persuaded to assume overall responsibility for the development of an inclusive SSH database.

---

<sup>14</sup> NORFACE is a partnership of 12 European research councils to increase research and research cooperation policy in Europe. HERA is a project formed of 13 research councils aimed at strengthening Humanities research and its 'profile' in Europe. Both projects are funded by the ERA-NET scheme the objective of which is to increase the cooperation and coordination of research activities carried out in the Member States of the European Union.



3. Establishing whether there is any prospect of working with Google Scholar to create a more rigorous bibliometric database. Google Scholar has the advantage in that it already covers books and book chapters. However, at present, there is little or no information on exactly what is included in the Google database, and the data are not sufficiently systematic and rigorous to be used for serious bibliometric purposes. Furthermore, Google Scholar only covers books that are available in whole or in part on the Internet, which remains a significant limitation.
4. Determining whether to support the further development of digital repositories of research outputs in universities and public research institutes by encouraging them to move towards the adoption of common standards and data formats, so that their data can be used as an input to a new European SSH bibliometric database.
5. Deciding whether to build on existing initiatives, drawing on lessons learned. For example, if it is decided that a European organisation is to develop the SSH bibliometric database, then there may well be important lessons to be learnt from the European DRIVER project (Digital Repository Infrastructure Vision for European Research). One could imagine a follow-on project to DRIVER that would build the SSH bibliometric database on the infrastructure already developed by DRIVER.<sup>15</sup>
6. Considering the long-term viability of an SSH bibliometric database and its resource requirements. If a European organisation is to be asked to develop the SSH bibliometric database, then consideration needs to be given to the possible sources of funding. Should this be a collaborative venture of European research councils? Might it be worth approaching the European Commission, perhaps in conjunction with a group such as ESFRI (the European Strategy Forum for Research Infrastructure), for the funding

---

<sup>15</sup> The DRIVER project aims to establish an infrastructure of European digital repositories for researchers and the general public (see <http://www.driver-repository.eu/> ).

needed to help enable smaller Member States, in particular, to develop digital repositories and bibliographic databases?

## **PART C**

### **Potential approaches for consideration**

This section provides a synthesis of suggestions by Moed et al., Hicks and Jiang, and the two workshops held in Brighton and Berlin, on how the main challenges identified in the study may be addressed. We present these as a background to the specific recommendations put forward in Part D. The recommendations are thus drawn from a consideration of these suggestions and their implications.

To reiterate, these challenges are:

1. the need to include a wide range of Social Sciences and Humanities outputs, such as books and book chapters, in any new SSH bibliometric database;
2. the need to cover national journals (in different languages apart from English) as well as international journals;
3. the variable coverage and quality of existing Social Sciences and Humanities bibliographic databases and lists from which a new SSH bibliometric database will draw extensively;
4. the lack of a standardised format for the input of data into bibliographic databases and lists, including the problem of listing the affiliations of first authors only. To some extent, this reflects the wide variation in publication and citation practices across the Social Sciences and Humanities disciplines, but without this consistency in data fields, it will be impossible to ensure comparability and to begin to integrate data from these different sources;
5. uncertainty over the quality criteria used in the selection of entries in the different bibliographic databases.

The numbering of the suggestions below is not intended to indicate any sense of priority. In addition, they are not to be seen as mutually exclusive. Each suggestion has a number of advantages and disadvantages (see Box 1 at the end of the six suggestions for a summary of those advantages, disadvantages and implications). As noted above, the recommendations presented in Part D will

select and/or combine the most pragmatic aspects of the suggestions after due consideration of their implications.

## **Suggestion 1**

*Create more comprehensive national bibliographic systems through the development of institutional repositories.*

Moed et al. note that a study conducted in 2006 by Van der Graaf and Van Eijndhoven on European institutional repositories found that only about a quarter of European higher education institutions (HEIs) have created digital repositories of their research outputs. Moreover, among these repositories, it would appear that only just over one third of the research outputs for a given year have been included. Taken together (and even assuming some improvement in the intervening three years), these figures imply that only about 10% of recent publication output of European HEIs is included in institutional repositories.<sup>16</sup> There is therefore considerable scope for this coverage to be extended.

### **Actions Required**

1. Assistance with capability-building for those countries that currently lack the necessary institutional repositories.
2. Help in designing and coordinating the introduction of institutional deposit policies to capture the full range and extent of research outputs (in particular, published books and journal articles) in each institutional repository.<sup>17</sup>
3. Collection of standard bibliographic meta-data for the deposited research outputs across institutional repositories.<sup>18</sup>
4. Encouragement of institutional repositories to begin capturing the cited reference lists contained in the published outputs (to supplement existing citations drawn from scanning international journal articles only).<sup>19</sup>

---

<sup>16</sup> Moed et al. (2009), p. 49.

<sup>17</sup> See Moed et al (2009) for more details.

<sup>18</sup> See Moed et al (2009) for more details.

<sup>19</sup> See Moed et al (2009) for more details.

## **Implications**

The main implication here is the vital need for development of the relevant capabilities of institutional repositories, both of which are apparently lacking in the majority of European HEIs. This suggests that such developments may have to be undertaken in tandem with selecting the most practical measure to 'kick-start' the creation of the SSH bibliometric database.

## **Suggestion 2**

*Enhance and build upon existing national documentation systems through the creation and standardisation of institutional research management systems.*

This recommendation is largely aimed at Higher Education Institutions (HEIs) that have developed lists of their research outputs for the purposes of research evaluation, for instance, for submission to a national research assessment exercise. A well designed national documentation research system should allow the flexibility to include not only international journals but other SSH research outputs, such as articles in national journals, books and book chapters.

### **Actions Required**

1. Build upon an existing research information system (e.g. METIS in the Netherlands), in which those submitting the data must specify (a) fields, (b) a list of relevant journals, and (c) some categorisation of journal levels.
  - To ensure that national journals receive the appropriate weight, a separate component within the system for national literature will need to be created for fields in which it is important (i.e. for fields that are not internationally homogeneous in terms of subject matter and approach, but instead focus more on nationally or regionally specific topics). The national journal list, which would need to be validated by national academic experts and academies, would have different criteria for determining the level of journals, and would be assessed separately from the international literature.

Agreed criteria for what constitutes 'peer-review' will also need to be established.<sup>20</sup> There will thus be two interacting but somewhat separate systems.<sup>21</sup> According to Moed et al. (2009), there are existing rules and protocols to build **interfaces** between such separate systems and databases.<sup>22</sup>

2. Expand on an agreed research information system through the **development and application of interfaces** to lists that include books and monographs.
  - For a database of books, this could be built with records that include author affiliation by adopting an international standard, such as the ONIX electronic international standard, which is currently used for representing and communicating book industry product information including author affiliation.<sup>23</sup>
  - Books and monographs, which will (like journals) be assigned to different levels, can be incorporated in a dedicated component from an acceptable and identified list of scholarly publishers agreed and validated by national academic experts and academies.<sup>24</sup>
3. **Or** agree on an existing research information system being used in institutions in the European Union, and then perform tasks (1) and (2) above.
4. **Or build on the DRIVER** initiative (if it is supported for further development and utilisation – see "Strategic options for Development" in Part B), and then perform tasks (1) and (2) above.
5. Link institutional repositories to this research information system.

---

<sup>20</sup> See footnote 13 above.

<sup>21</sup> See Hicks and Wang (2009) for more details.

<sup>22</sup> See Moed et al. (2009), p. 51.

<sup>23</sup> See Hicks and Jiang (2009).

<sup>24</sup> This is already being done by the Norwegian model; see Hicks and Wang (2009) for more details.

## Implications

There are three main implications here. The *first* is the development of a research information system for countries that do not currently have one. The *second* is that a minimum threshold criterion be established for the inclusion of selected outputs. The *third* is the possible adoption of a variant or combination of existing database systems, such as METIS, DRIVER or a commercial system, but this requires further investigation including an examination of the possibility of orchestrating some convergence between these alternatives. This, in turn, implies a capability in, or available resources for, the development and implementation of interfaces to enable 'convergence' to take place. Overall, these implications suggest that Suggestion 2 is likely to face considerable difficulties that need to be overcome in the initial stages.

## Suggestion 3

*Create a new database of SSH research outputs from publishers' archives and institutional repositories (articles and books), and (in due course) add to this appropriate data on enlightenment literature and curated events.*

A possible exemplar here is the new database being developed by the Spanish Research Council from publishers' archives.

### Actions Required

1. Create a new database from scratch that includes all publications and citation data obtained directly from publishers.
2. Identify **enlightenment** books and periodicals, perhaps categorised by readership, and then assign levels for this database.<sup>25</sup>
3. List and assign levels for curated events and other **non-textual outputs** for this database, which will be agreed by national experts.<sup>26</sup>

---

<sup>25</sup> See Hicks and Wang (2009) for more details.

<sup>26</sup> See Hicks and Wang (2009) for more details.

## Implications

The main implication here is the resource-intensity (time and cost) and complexity of creating and maintaining such a database. This suggests that this may not be a suitable pragmatic measure to 'kick-start' the process of creating a SSH bibliometric database.

## Suggestion 4

*Try to take advantage of the competition between the Web of Science and Scopus to strengthen the coverage of SSH research outputs, and of the potential of Google Scholar to become a more rigorous bibliometric database provider.*

As noted above in Part A, the Web of Science and Scopus are already expanding their coverage of SSH journals, including the introduction of books. Clearly, the main advantage of these two databases is their international acceptance as a source for structured bibliometric analysis. However, Part A has also discussed Google Scholar as a potential supplier of a bibliometric database if improvements are made in terms of transparency, systematisation and rigour. The advantage of Google Scholar/Google Book Search is its uniqueness in being the only database currently covering books as sources of citation links.<sup>27</sup>

## Actions Required

1. Decide on who should approach and explore whether a deal might be negotiated with Thomson-Reuters, Elsevier or Google to ensure not just best value for money (as significant public monies will be involved, for example, in providing the bibliographic lists) but also compatibility with the intended purposes of the SSH bibliometric database.<sup>28</sup>

---

<sup>27</sup> Part B, under "Strategic options for consideration", has also offered a suggestion for enticing the company into becoming such a supplier.

<sup>28</sup> Workshop participants and research council representatives were insistent that the purposes of the database be clarified at the outset. NWO representatives commented that what is 'commercially feasible' may be somewhat at odds with what is 'scientifically feasible' with regard to a SSH bibliometric database. Hence it is important that the main purposes of the database should not be compromised in discussions with commercial suppliers, if the latter are approached to help construct the database.



2. Approach Thomson-Reuters (Web of Science) and Elsevier (Scopus) with the idea of expanding their book coverage to include complete bibliographic meta-data on highly cited books, chapters and monographs.
3. Try to find out about Google's future plans regarding the integration of Google Scholar and Google Book Search, then approach Google with the idea of eventually becoming a 'fully-fledged' bibliometric database provider.

### **Implications**

The main implication here is the need for a nominated party who has the extensive knowledge on bibliometrics required to negotiate with the bibliometric publishers.

### **Suggestion 5**

*Integrate the specialised SSH bibliographic lists into one comprehensive bibliographic database.*

As noted in Part A, there are several problems with these specialised lists, not least of which is the absence of a standardised database structure and data fields.

### **Actions Required**

1. Move towards an agreed standardisation of the database structure among the main producers of these bibliographic lists.
2. Examine the existing selection criteria for the sources included in these lists, and how these might be standardised.
3. Introduce books as they are currently very under-represented in the majority of these lists.<sup>29</sup>

### **Implications**

The main implication here is the need for a body of bibliometricians to spearhead the process of standardisation.

---

<sup>29</sup> Moed et al. (2009, p. 47).

## Suggestion 6

*Encourage the further development of the Open Access approach, as this offers a potential means to overcome barriers of accessibility and to enhance the visibility of SSH journals and books published by small European publishers.*

The U.S. National Research Council has adopted this model, while some European university presses are engaged in developing an Open Access SSH library. The advantages of such a system are:

- a. It could build on existing schemes to support small European SSH publishers (such as the OAPEN project, which is funded as part of the European Commission “e-Content Plus” Programme).<sup>30</sup>
- b. It would improve the availability and ‘promotion’ of European SSH outputs.
- c. By providing scholars with access to this database, it would help to overcome the accessibility problems posed by a currently rather fragmented publishing industry.
- d. All the electronic items will be indexed by Google Scholar (as is already being done by Google Scholar of all electronic full texts), thereby further facilitating access to European SSH research outputs.
- e. It would provide a revenue source as users would be permitted to read only single pages of the publication, with full publications then being sold relatively cheaply for downloading and saving or for printing.<sup>31</sup>
- f. It would open up the potential for citation analysis, although ‘Open Access’ databases do face certain difficulties because of the content and structure of

---

<sup>30</sup> This programme aims to develop and implement an Open Access publication model for SSH books. It uses the DRIVER infrastructure. OAPEN consists of a number of European university presses and universities, such as Amsterdam University, Göttingen University, Manchester University and Firenze University (see Moed et al., 2009).

<sup>31</sup> See Hicks and Wang (2009).

individual repositories, conflict of interests with commercial e-publishers, and so on.<sup>32</sup>

### **Actions Required**

1. **Build and maintain** an electronic full-text SSH journal infrastructure, including the establishment of meta-data fields (author, institution, journal name, etc.).<sup>33</sup>
2. This database will include peer-reviewed journals not on-line and not indexed by WoS and Scopus.<sup>34</sup>
3. Build upon the **OAPEN digital library** and include more European book publishers.
4. Integrate (1) with (3) through the development and application of appropriate **interfaces**.
5. Consider an agreed set of metrics, such as number of downloads or links to related electronic documents.

### **Implications**

The main implications here are (1) the potential redundancy of effort and (2) the potential conflict of interest with the current publishers of bibliometric databases.

Box 1 below summarises the suggestions discussed above. It must be emphasised here that each approach requires a considerable amount of time to develop, the extent of which is difficult to specify as it depends, amongst other things, on the state (quality and

---

<sup>32</sup> See Moed et al. (2009, pp. 51-52), who highlight the problem with commercial e-publishers because of the need for an Open Access database to know the download and sales figures of each book. Such data will be hard to obtain. They also argue that it is necessary to obtain library loan figures for books; these, too, are not readily available.

<sup>33</sup> Hicks and Wang do not recommend working with institutional Open Access repositories, mainly because their quality and coverage may be questionable, thus making these lists unsuitable for assessment purposes.

<sup>34</sup> Hicks and Wang, however, recommend a scrupulous *needs-assessment* for this approach because it entails large upfront costs and maintenance in much the same way as any effort to combine existing special SSH bibliographies. In addition, a significant expansion in the coverage of journals and books by the Web of Science and Scopus could render such an Open Access database redundant.

comprehensiveness) of the SSH bibliographic databases and the institutional capabilities for the production of these databases.

**Box 1. Summary of the advantages and disadvantages of each approach**

Suggestions	Advantages	Disadvantages	Implications
<p>1. Create a more comprehensive national documentation system through the development of institutional repositories.</p>	<p>Provides opportunities to countries with inadequate skills to develop comprehensive institutional bibliographic lists and a national documentation system.</p> <p>The data could be harvested for the SSH bibliometric database.</p>	<p>Identify which countries – what selection criteria for identification?</p> <p>Need ‘political will’ to design policies to develop lists and national documentation system.</p> <p>Resource-intensive.</p> <p>Time-intensive as have to start from ‘almost scratch’.</p> <p>Will likely slow down the development of the SSH database.</p>	<p>The vital need for development of bibliometric capabilities and of institutional repositories, both of which are apparently lacking in the majority of European HEIs. This suggests that such a measure may have to be undertaken in tandem with selecting the most practical measure to ‘kick-start’ the creation of the SSH bibliometric database.</p>
<p>2. Enhance and build upon existing national documentation systems through the creation and standardisation of institutional research management systems.</p>	<p>A well designed national documentation system should have the flexibility to include a range of SSH outputs.</p> <p>Enhances the possibility of integrating national institutional repositories.</p>	<p>Aimed at HEIs that already have institutional bibliographic lists.</p> <p>Potential difficulty in arriving at consensus on which institutional research management system to adopt.</p> <p>Excludes HEIs that have not developed a national documentation system AND bibliographic lists.</p>	<p>The <i>first</i> is that a minimum threshold criterion be established. The <i>second</i> is the possible adoption of a variant or combination of existing systems, such as METIS, DRIVER or a commercial system, but this requires further investigation including an examination of the possibility of orchestrating some convergence between them.</p>
<p>3. Create a new database of SSH research outputs from publishers’ archives and institutional repositories (articles and books) and (in due course) add to this data on enlightenment literature and curated events.</p>	<p>New database from publishers’ archives is already being created by Spain – showing that it can be done.</p> <p>Allows for a wide range of SSH outputs.</p>	<p>Resource-intensive.</p> <p>Time-intensive.</p> <p>Need standardisation of institutional lists.</p> <p>Risk of ‘reinventing the wheel’?</p>	<p>The main implication here is the resource-intensity (time and cost) and complexity of creating and maintaining such a database. This suggests that this may not be a pragmatic measure to ‘kick-start’ the process of creating a SSH bibliometric database.</p>

Suggestion	Advantages	Disadvantages	Implications
<p>4. Try to take advantage of the competition between the Web of Science and Scopus, and the potential of Google Scholar to become a rigorous bibliometric database provider.</p>	<p>Web of Science and Scopus have international acceptance in terms of their use for structured bibliometric analysis.</p> <p>Google Scholar is the only database that comprehensively covers books; it is easy to find book references; and it collects simple citation indicators. Currently used for 'desk-top bibliometric analysis'.</p> <p>Promotes competition.</p> <p>Likely to be most cost-effective solution, depending on negotiated terms for production and public usage.</p> <p>Accelerates the creation of SSH bibliometric database.</p> <p>Maintenance of database more assured.</p>	<p>Google Scholar's current business model is non-transparent; its multiple sources are still unknown; its records are not usable for structured analysis; and there is concern about the accuracy of the citation links.</p> <p>Who negotiates the 'deal' to ensure value for money for the users (as public resources will be incurred) and commercial publishers?</p>	<p>The main implication here is the need for a nominated party with extensive knowledge of bibliometrics to negotiate with the bibliometric publishers.</p>
<p>5. Integrate the specialised SSH bibliographic lists into one comprehensive bibliographic database.</p>	<p>Some specialised bibliographic databases, such as the family of CSA-Illumina databases, already have a standardised database structure.</p> <p>These databases cover specific (sub-)disciplines.</p> <p>Many of these databases are accessible through a common interface.</p>	<p>Time-intensive.</p> <p>Resource-intensive – major effort needed to standardise and de-duplicate these databases for bibliometric analysis and for maintenance of database.</p> <p>Who maintains the database?</p> <p>High risk of redundancy, especially if Web of Science and Scopus continue expanding their databases.</p> <p>Unclear selection/quality criteria for inclusion of outputs.</p> <p>High incidence of absence of institutional affiliations of publishing authors.</p> <p>Stark under-representation of books.</p>	<p>The main implication here is the need for a body of bibliometricians to spearhead the process of standardisation.</p>

Suggestion	Advantages	Disadvantages	Implications
<p>6. Encourage the further development of the Open Access approach, as this offers a potential means to overcome barriers of accessibility and to enhance the visibility of SSH journals and books published by small European publishers.</p>	<p>The use of public money to support small European SSH publishers.</p> <p>Facilitates the availability, accessibility and 'promotion' of European SSH outputs.</p> <p>Facilitates easier access and helps to overcome the accessibility problems posed by a fragmented publishing industry.</p> <p>A revenue source, as users are permitted to read only one page of the article, so full articles need to be purchased for downloading.</p>	<p>Resource-intensive for standardisation because of variability in the structure and content of the national institutional bibliographic lists and high maintenance costs of database.</p> <p>Difficult to arrive at standardisation.</p> <p>Time-intensive.</p> <p>Conflict of interest with commercial publishers – need to know the download and sales figures of commercial e-publishers.</p> <p>Need to know loan figures for each book, or each article from libraries, which are seldom available.</p> <p>Who maintains the database?</p> <p>High risk of redundancy, especially if Web of Science and Scopus continue expanding their databases.</p>	<p>The main implications here are the potential redundancy of effort and conflict of interest with the current publishers of bibliometric databases.</p>

## PART D

### **Recommendations for the development of a comprehensive SSH bibliometric database**

To reiterate, the SSH bibliometric database is intended to fulfil a number of functions:

1. to provide accountability with regard to the use of public funds;
2. to assess research quality and to permit the development of performance indicators;
3. to provide a comprehensive overview of SSH research outputs in Europe;
4. to map the directions of SSH research, indicating, for example, which areas are under-researched, or which exhibit an established research capacity and which are lacking this;
5. to identify new emerging areas of interdisciplinary SSH research.

Given the above objectives and the desire of research councils to initiate the construction of a SSH bibliometric database in a timely fashion, we advocate that the development of a comprehensive SSH bibliometric database be carried out on the basis of **four** recommendations. These involve:

1. defining the criteria for inclusion of SSH research outputs and establishing a standardised database structure for national bibliometric databases;
2. exploring the option of involving a commercial supplier in the construction of a single international SSH bibliometric database;
3. conducting a pilot study of one or more specific SSH disciplines;
4. longer-term expansion and enhancement of the SSH bibliometric database.

For each recommendation, a **hybrid approach** is commended based on **a combination of top-down and bottom-up actions**, with the emphasis on extensive bottom-up involvement in the



production and development of the bibliographic databases and lists that will then underpin the SSH bibliometric database. Practicality of implementation and cost-effectiveness are the two main criteria underlying the choice of the recommendations that follow.

This part of the report presents the four main recommendations, for each of which we lay out a series of actions. The recommendations here, as noted in Part C, combine various aspects from the suggestions presented there.

Recommendations 1 and 3 may be undertaken in parallel in order to save time. Such a decision to conduct them in tandem, however, will depend on the views of the research councils or organisations charged with the task of creating a SSH bibliometric database and the resources they are able to make available. Recommendation 2 can only commence after significant progress has first been made with Recommendation 1. Likewise Recommendation 4 is probably best left until Recommendations 1 and 3 have been largely completed so that the insights gained into what other research outputs and indicators need to be considered, particularly for the Humanities, can be fully taken into account.

After due consideration of the substantial difficulties and large upfront investment highlighted by the two commissioned studies, we have decided not to recommend pursuing certain of the options presented in Part C for reasons of practicality and cost-effectiveness. Those not pursued here are (i) the Open Access approach; (ii) the integration of specialised SSH bibliographic lists; and (iii) the creation of a new database of SSH research outputs from publishers' archives and institutional repositories (see Part C above for a discussion of the difficulties inherent in each of these suggestions and Box 1 for a summary of their respective advantages, disadvantages and implications).

We suggest that a hybrid approach (i.e. a combination of 'top-down' and 'bottom-up') be adopted with regard to each recommendation. A hybrid approach is likely to prove most cost-effective as it should ensure that the coordination, comparability and integration of the respective databases are achieved without incurring unnecessarily heavy additional cost to countries that have already invested significant resources in national bibliographic databases. A hybrid

approach also would seem to offer the best of both worlds, with impetus, guidance and 'clout' being provided by the top, and expertise, inputs (providing and validating content) and feedback coming from the bottom.

## **Recommendation 1: Define the criteria for inclusion of SSH research outputs and establish a standardised database structure for national bibliometric databases**

This recommendation focuses on (1) the establishment of the minimum criteria for the inclusion of scholarly peer-reviewed articles and books, and (2) the creation of a standardised structure for the various national bibliometric databases so that they provide comparable data across countries. The achievement of this will be underpinned by five key components:

1. strong coordination and close working between national organisations, in particular research councils and institutional repositories, to provide the necessary impetus for the development of internationally comparable bibliographic databases/lists and their gradual transformation into full bibliometric databases;
2. bottom-up involvement of national institutions and repositories in consultation with bibliometric experts, users and SSH scholars on the provision, validation and development of the eventual bibliometric databases;
3. resources being made available to national institutions and repositories that have inadequate capabilities to develop bibliographic databases/lists into full bibliometric databases.
4. in order to avoid the process of data collection getting out of control and to ensure the harmonisation of collected data among the involved countries, the definition of the standardised structure must be established at the outset as any subsequent changes<sup>35</sup> to it will prove extremely costly;

---

<sup>35</sup> Subsequent structural changes, such as including new meta-data in later phases will incur disproportionate expenditure of labour, time and costs.

5. Important meta data<sup>36</sup> should be included even if they are not used in the initial phase of database utilisation.

## **Actions Required**

### *Top-down*

- In order to ensure consistency of criteria and standardisation of practice and to avoid many of the operational pitfalls discussed in Part C, as well as to provide the necessary 'platform' or infrastructure on which the other recommendations will build, the following **actions are recommended**:
  - That a small group of research councils from a diverse range of (large and small) Member States<sup>37</sup> should take the lead in setting in motion the process of standardisation and the establishment of the minimum criteria for inclusion of research outputs in a SSH bibliometric database. (These are hereafter referred to as 'the lead research councils'.)<sup>38</sup>
  - That the lead research councils should seek to expand the composition of this group incrementally, for instance, by including the Norwegian Research Council (given that its SSH bibliometric database was favourably reviewed by Hicks and Jiang) and others, in order to 'keep up the momentum' of development so as

---

<sup>36</sup> Such meta-data included in contemporary 'bibliometric' databases are, for instance, references, all authors/editors/contributors, full affiliation/address information, author-affiliation assignment, and acknowledgments including funding information. In the light of a major initiative forthcoming from CrossRef to obtain a Universal Researcher ID, which will take over the Thomson-Reuters and Scopus researcher IDs, the metadata should also include Researcher ID.

<sup>37</sup> This could follow the examples of HERA and NORFACE [NWO] (See point one under Section on "Strategic Options" in Part C for an explanation of NORFACE and HERA).

<sup>38</sup> Although a larger lead group of research councils would allow for broader coverage in terms of participation and diversity, the Board Members have in the majority agreed that, for practical reasons, a smaller group of research councils is needed to kick off the process, while clearly recognising that participation from other research councils will also be necessary, as explained in the subsequent bullet point. However, as the process gets underway and gains momentum, we fully expect the group of 'lead research councils' to expand so as to reflect the full diversity of European research traditions, cultures and languages.

to eventually include all EU Member States, and so as to avoid the danger of producing a distorted picture of the diversity of research traditions and cultures as exist in different European member states.

- That the lead research councils appoint **a standard-setting body**, which will include a combination of expert bibliometricians and library or documentation experts (from Europe and elsewhere, in particular, those familiar with the specificities of SSH research), and SSH researchers to ensure that, in setting the standards, there is due consideration of different disciplinary communication modes. Some of those library/documentation experts might be selected from Higher Education Institutions (HEIs) that already have advanced documentation systems and are experienced in the production of lists of research outputs for various purposes, such as research evaluation or the regular monitoring of research performance.

While it is clearly not possible to include a researcher-representative from every SSH discipline, the standard-setting body must endeavour to establish close collaboration with SSH researchers not represented in the standard setting body. This could, for example, be done via research councils along similar lines to NORFACE and HERA, who could then incorporate the gathering of such inputs into their activities. The research councils would then communicate the inputs to the standard-setting body.

This standard-setting body will be responsible for setting the standard for the structure of SSH bibliometric databases and for establishing the criteria for the inclusion of articles and books (and, in due course, other research outputs). The purpose of this action is to expedite the formation of a standard based in large part on an examination of different information management systems, such as METIS, DRIVER or a commercial structure (for instance, ONIX in the case of

books), to see if some orchestrated convergence or adaptation of these systems is possible.

- That the standard-setting body should act autonomously, but should actively consult with a range of SSH scholars as well as with experts on bibliometrics and impact assessment techniques, commercial database suppliers and national repositories.
- That the standard-setting body will establish the **minimum criteria** for inclusion in the SSH bibliometric databases, which could, as a pragmatic measure, begin with (1) scholarly articles from peer-reviewed international and national journals; (2) books that have undergone a similar peer-review process prior to publication.<sup>39</sup> This exercise will be accompanied by extensive consultation with SSH scholars, including European and national scientific and research associations in the different SSH disciplines covered by the group of lead research councils. This consultation will be repeated as the SSH database expands to include other Member States.
- As a possible alternative to the above minimum criteria, the standard-setting body will identify a group of leading HEIs with extensive experience in setting up bibliographic databases/lists to help determine appropriate common criteria for the inclusion of SSH articles and books. As with Point (2) above, close consultation with the SSH communities is again strongly recommended.
- That the standard-setting body, in consultation with external bibliometric experts, commercial database suppliers such as Thomson-Reuters and Scopus, Proquest/Cambridge Scientific Abstracts (and non-

---

<sup>39</sup> Caution with regard to publishers will need to be exercised. According to some Spanish academics, certain leading Spanish publishers are charging authors for publication of their books. Authors who refuse to comply with the charges may therefore not be published. This raises the question of peer-review or the quality of the books.

commercial ones such as CABI) and SSH researchers (from the countries of the research councils leading the effort including those who join subsequently), will also consider what 'book metrics' are required.

- That the standard-setting body, in consultation with national repositories and SSH scholars associated with the research councils leading this initiative, will decide the time-frame for including journals – in particular, how far back they wish to go.<sup>40</sup>
- That the standard-setting body establish a realistic time-frame required for the identification of peer-reviewed journals and books for inclusion into the database. This time-frame will be established **after** the structure of the standard has been completed.
- That resources are sought for national institutions or repositories that currently lack the capability to develop a bibliometric database. Although this capability will be required for the SSH bibliometric database that will eventually include all European Union Member States, we **recommend** as a preparatory measure that
  - all EU national research councils and the ESF present a case to the European Commission to make 'structural funds' available for this task. Alternatively, a case for such funding could be made to the European Commission in conjunction with the European Strategy Forum for Research Infrastructure (ESFRI).
- To oversee and manage the project, and initially maintain the SSH bibliometric database, we **recommend**
  - that the group of lead research councils appoint a project manager from among themselves;

---

<sup>40</sup> Over time, some journals cease publications while other new ones appear, so a pragmatic decision will be needed as to how far back the SSH database should extend, at least in the first instance.

- that the project manager set up a committee comprising representatives from the group of lead research councils and assign tasks to each committee member;
- that the project manager and committee submit progress reports to their respective research councils (or equivalent) and funders on a half-yearly basis;
- that the project manager and committee consult with experts in impact assessment techniques and methodologies;
- that the project manager should initially maintain the database but should, with the committee, determine the resources required for the maintenance of this in the longer term, where the funds should be obtained from, and time frame for making this transition.

#### *Bottom-up*

- National institutions and institutional repositories of the countries of the lead research councils will be responsible for applying the agreed minimum criteria to their bibliographic databases as they begin to transform them into full bibliometric databases.
- National institutions and institutional repositories, in close consultation with the national SSH communities, will be responsible for the identification of high quality peer-reviewed national or regional journals and books.
- Each national institution or repository of the countries represented by the lead research councils will be responsible for ensuring that the standard decided by the standard-setting body for structuring the bibliographic databases is then implemented so that over time they are transformed into comparable bibliometric databases.

## **Recommendation 2. Explore the option of involving a commercial supplier in the construction of a single international SSH bibliometric database**

The construction of the SSH bibliometric database by a commercial supplier may prove to be a particularly cost-effective measure, given that Thomson-Reuters (publishers of the Web of Science) and Elsevier (publishers of Scopus) are both established bibliometric database suppliers, while Google Scholar/Google Books already covers a range of books in its database.

This recommendation is underpinned by three key components:

1. obtaining the necessary 'buy-in' from national organisations, in particular research councils, to provide the impetus and funding (either directly from themselves, or indirectly, for example, through the European Commission) for such an initiative;
2. ensuring strong coordination between those national organisations so that the commercial suppliers can be approached with a clear and common goal;
3. stimulating the bottom-up development of lists and bibliographic databases by national institutions or repositories in a process in which national bibliometric experts and users and SSH scholars are all closely involved in providing and validating the content of these lists, with those lists then being passed over to the commercial developer of the SSH bibliometric database.

### **Actions required**

#### *Top down*

- Decide on how best to approach Thomson-Reuters (Web of Science), Elsevier (Scopus) and perhaps also Google with a view to exploring a possible deal. We **recommend** that
  - that the chair of the standard-setting body be the chief negotiator for this action, under the auspices of the lead research councils;



- that the chair of the standard-setting body consult with institutions or individuals that have previously dealt with Thomson-Reuters, Elsevier and Google, such as Tony van Raan (CWTS), Felix de Moya (SCImago)<sup>41</sup>, Lorraine Estelle (Joint Information Systems Committee – JISC Collections)<sup>42</sup>, Graeme Rosenberg (Higher Education Funding Council of England, HEFCE), and Ana Maria Prat (the National Commission for Scientific and Technological Research, CONICYT)<sup>43</sup>, for information on their experiences in dealing with these commercial suppliers.
- Decide whether the commercial suppliers should be asked to ‘clean up’ existing bibliographic databases and lists and incorporate them into their existing bibliometric databases; OR whether commercial suppliers should instead be invited to construct a new database (focusing on national journals and different languages, and books) to complement their existing database (focusing on international journals). We **recommend** that
  - **both options** be presented to commercial suppliers for pricing. Depending on the price difference between the two options (that is, if it is not large), it is preferable to select the incorporation option as many institutions already have extensive bibliographic databases/lists.

#### *Bottom up*

- National institutions and institutional repositories will be wholly responsible for the creation of their respective bibliographic databases/lists and possibly the eventual

---

<sup>41</sup> Felix de Moya was one of the authors of the commissioned study led by Henk Moed, and he works extensively on Scopus.

<sup>42</sup> JISC is funded by the UK HE and FE funding bodies to provide world-class leadership in the innovative use of ICT to support education and research. JISC manages and funds more than 200 projects within 28 programmes. Outputs and lessons are made available to the HE and FE community. JISC also supports 49 Services that provide expertise, advice, guidance and resources to address the needs of all users in HE and FE. See <http://www.jisc.ac.uk/whatwedo.aspx>.

<sup>43</sup> Ana Maria Prat attended the project workshop held at SPRU on 18 March 2009. Her institution has dealt with Google on a bibliographic database for her institution.

transformation of these into bibliometric databases. Note that even if the option eventually chosen is for the commercial supplier instead to construct a new bibliometric database, bibliographic lists will still be required for handover to that commercial supplier.

- National institutions/repositories should consult with a broad range of SSH researchers to ensure the quality and validity of their respective bibliographic lists.

### **Recommendation 3. Conduct a pilot study on one or more selected SSH disciplines**

As SSH disciplines exhibit quite different communication modes, there would be merit in conducting a pilot study focusing on one or more selected SSH disciplines and collecting data on the relevant research outputs that should be included in a bibliometric database aimed at serving the five main purposes highlighted in this scoping study. The task of how best to construct appropriate quality or impact indicators could also be addressed. The pilot study should try to reflect ideas on the standard emerging from the implementation of Recommendation 1, and could be undertaken in parallel with Recommendation 2.

#### **Actions required**

##### *Top down*

- The group of lead research councils should decide on the **SSH discipline(s)** and **countries** that will be the focus of the pilot study. Examples of SSH disciplines that might be candidates include history, geography, linguistics and philosophy.<sup>44</sup> The choice of the disciplines will be made by the research councils within the lead group that have the necessary resources to fund the pilot studies.

---

<sup>44</sup> Several participants at the Berlin workshop argued that there needed to be a particular focus on the Humanities, where more work needs to be done on bibliometrics, given that their communication and publishing modes (for instance, the much greater emphasis on books) differ more from the Social Sciences and from STEM subjects.

- The group of lead research councils should decide whether this pilot study should be carried out and, if so, how it is to be funded. For example, it might be undertaken by 'volunteer' institutions in the selected countries. We **recommend** that
  - a fee' be made available to these institutions;
  - an invitation to research institutions/HEIs of the selected countries be issued, after which the group of lead research councils will select the candidates to conduct the pilot study;
  - if the funding for such a fee is not available, then an invitation for 'volunteers' will be issued to **selected leading HEIs across Europe** that possess substantial experience in constructing bibliographic databases;
  - a timescale for delivery of the bibliometric database will be set of around 12 months.

*Bottom up*

- The institutions selected to conduct the pilot study will engage in full consultation with SSH researchers in determining the data coverage and appropriate performance indicators.
- The institutions selected will consult experts on impact assessment techniques and/or studies dealing with these techniques.
- The institutions selected will have the capacity to construct the bibliometric database and to deliver it within a timescale set by the group of lead research councils.

## **Recommendation 4. Longer-term expansion and enhancement of the SSH bibliometric database**

This recommendation corresponds to the last part in the long journey toward the construction of a fully inclusive international SSH bibliometric database. It focuses on the gradual inclusion of other SSH outputs (that is, in addition to peer-reviewed articles and books).

As with the above three recommendations, this recommendation involves a hybrid approach, and it will build on what has been achieved in Recommendations 1 and/or 2 and 3.

This recommendation is underpinned by four key components:

- ensuring that there is a consensus among the lead research councils as to what other SSH research outputs are to be included in the expanded SSH bibliometric database;
- working in close communication with experts in impact assessment techniques for SSH so that the process of development of the SSH bibliometric database is not undertaken in isolation from other work on SSH impact assessment techniques and methodologies;
- deciding on what research outputs can best capture economic and social impacts;
- agreeing on other impacts that they wish to capture from the full range of SSH research outputs in order to inform the construction of appropriate research output data and indicators that best reflect the needs and interests of the full range of SSH disciplines.

### **Actions required**

#### *Top down*

- A preliminary decision needs to be made on who/what institution is to be responsible for subsequently maintaining the SSH bibliometric database (as this implies a long-term commitment of significant resources). We suggest<sup>45</sup> that:
  - collective funding from national research councils be used to underwrite the maintenance of the SSH bibliometric database;
  - alternatively, the consortium of lead research councils and the ESF should approach the European Commission (perhaps in conjunction with the European Strategy

---

<sup>45</sup> Unfortunately, the scope of the work carried out by the Project Board is such that we are unable to make very specific recommendations here, particularly with regard to the likely costs.

Forum for Research Infrastructure) to seek long-term funding for the SSH bibliometric database;

- then an Invitation to Tender should be issued by the lead research councils for the further development of the international bibliometric SSH database (this presumes that Recommendation 2 has not been pursued with commercial providers or has proved unsuccessful).
- The new standard-setting body (see Recommendation 1) should develop research output indicators and criteria for inclusion of a range of other SSH outputs (i.e. other than scholarly articles and books). We **recommend**
  - that monographs and 'grey' and 'enlightenment' literature should be among the first items to be included in this expanded SSH bibliometric database, as well as other important research outputs identified from the pilot study<sup>46</sup>;
  - that the standard-setting body identify a select **group of leading HEIs experienced in the production of bibliographic databases** to help determine appropriate criteria and indicators, including those suggested from the pilot study, as they will have valuable experience in dealing with the unforeseen problems that will inevitably occur in the long-term development process of database construction;
  - that the standard-setting body also seek inputs from **HEIs experienced in collecting systematic information on the production of non-textual outputs**, as they should have valuable insights into what non-textual outputs should/could be included in the inclusive SSH bibliometric database;
  - that the standard-setting body also consult with commercial suppliers and bibliometrics experts about

---

<sup>46</sup> This presumes that Recommendation 3 for a pilot study on specific SSH disciplines has been implemented.

the creation of a wider range of SSH research output indicators.

- that the standard-setting body consult with experts on impact assessment techniques for SSH, who could provide valuable suggestions for appropriate impact (or 'quality') indicators, such as those for social and economic impact.

#### *Bottom up*

- National institutions and institutional repositories will adopt the various SSH research outputs identified by the standard-setting body (see above) and produce systematic databases/lists of these outputs;
- National institutions and institutional repositories will comply with the criteria established for inclusion of the identified SSH outputs in the production of their databases/lists of these outputs, which will then be supplied to the developer of the international SSH bibliometric database.

For a graphical presentation of the recommendations and timescales, see Annex 3.

The above four recommendations would appear to offer the best way forward in exploring and then establishing an international bibliometric database for the Social Sciences and Humanities. The hybrid approach outlined, because it involves various stages, provides the opportunity to curtail the process at any point if the problems prove to be intractable or excessively expensive to overcome. It also assumes that the prospects of success will be greater if the initiative starts with a relatively small group of research councils and countries, allowing others to join in subsequently as momentum builds and as the necessary resources become available.

## **Selected References**

- Burnhill, P.M. and M.E. Tubby-Hille (2003) 'On Measuring the Relation between Social Science Research Activity and Research Publication'. *Research Evaluation*, 4, 3, 130-152..
- CWTS (Centre for Science and Technology Studies), Leiden University, (2007) *Scoping study on the use of bibliometric analysis to measure the quality of research in UK higher education institutions*. A Report to HEFCE.
- Hicks, D. and J. Wang (2009) *Toward a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project*. A Report to the Project Board of the Scoping Study 'Towards a Bibliometric Database for the Social Sciences and the Humanities'. Under the auspices of the European Science Foundation and funded by ESRC, AHRC, ANR, NWO and DFG
- Kyvik, S (2003) 'Changing trends in publishing behaviour among university faculty, 1980-2000'. *Scientometrics*, 58, 1, 35-48.
- Moed, H., J. Linmans, A. Nederhof, A. Zuccala, C. Lopez Illescas and F. de Moya (2009) *Options for a Comprehensive Database of Research Outputs in Social Sciences and the Humanities*. A Report to the Project Board of the Scoping Study 'Towards a Bibliometric Database for the Social Sciences and the Humanities'. Under the auspices of the European Science Foundation and funded by ESRC, AHRC, ANR, NWO and DFG.
- White, B.D. (2006) Examining the claims of Google Scholar as a serious information source. *New Zealand Library & Information Management Journal*, 50 no 1, pp. 11-24.

## **Annex 1**

### **Towards a Bibliometric Database for the Social Sciences and Humanities – A European Scoping Project**

Diana Hicks & Jian Wang

School of Public Policy

Georgia Institute of Technology



## **Annex 2**

### **Options for a Comprehensive Database of Research Outputs in Social Sciences and Humanities**

**Henk F. Moed, Janus Linmans, Anton Nederhof and Alesia Zuccala**

Centre for Science and Technology Studies (CWTS), Leiden University,  
The Netherlands

and

**Carmen López Illescas and Felix de Moya Anegón,**

SCIMago Research Group, CSIC Madrid and University of Granada,  
Spain

Research report to the Project Board of the Scoping Study 'Towards a Bibliometric Database for the Social Sciences and the Humanities' set up by the Standing Committees for the Social Sciences and the Humanities of the European Science Foundation (ESF)

## Annex 3 – Roadmaps









