

# Methodologies for the Characterization of the Publication Output of Higher Education Institutions Using Institutional Databases

Emanuela Reale<sup>1</sup>, Bianca Maria Poti<sup>2</sup>, Emilia Primeri<sup>3</sup>, Benedetto Lepori<sup>4</sup>, Carole Probst<sup>5</sup>, Isabel Gomez Caridad<sup>6</sup>, Daniela Alejandra De Filippo<sup>7</sup>, Elias Sanz Casado<sup>8</sup> and Peter van den Besselaar<sup>9</sup>

<sup>1</sup>[e.reale@ceris.cnr.it](mailto:e.reale@ceris.cnr.it), <sup>2</sup>[b.poti@ceris.cnr.it](mailto:b.poti@ceris.cnr.it), <sup>3</sup>[e.primeri@ceris.cnr.it](mailto:e.primeri@ceris.cnr.it)  
CNR- Ceris (Institute for Research on Firms and growth), Via dei Taurini 19, 00185 Rome (Italy)

<sup>4</sup>[benedetto.lepori@unisi.ch](mailto:benedetto.lepori@unisi.ch), <sup>5</sup>[carole.probst@lu.unisi.ch](mailto:carole.probst@lu.unisi.ch)  
Università della Svizzera Italiana (USI) Via Lambertenghi 10, CH-6904 Lugano (Switzerland)

<sup>6</sup>[igomez@cindoc.csic.es](mailto:igomez@cindoc.csic.es), <sup>7</sup>[dfilippo@cindoc.csic.es](mailto:dfilippo@cindoc.csic.es)

IEDCYT (Instituto de Estudios Documentales sobre Ciencia y Tecnología)-CSIC, Joaquín Costa, 22, 28002, Madrid (Spain)

<sup>8</sup>[elias.sanz@uc3m.es](mailto:elias.sanz@uc3m.es)  
Universidad Carlos III de Madrid, Calle Madrid, 126 -28903 Getafe, Madrid (Spain)

<sup>9</sup>[p.vandenbesselaar@rathenau.nl](mailto:p.vandenbesselaar@rathenau.nl)  
Netherlands Center for Science System Assessment Rathenau Instituut PO Box 95366 2509 CJ Den Haag, (The Netherlands)

## Introduction

UNIPUB is a project developed within the PRIME Network of Excellence (VI EUFP) which aims at experimenting the exploitation of institutional databases for producing indicators of publications at the level of whole higher education institutions (HEIs). Publications constitute one of the major means for transmitting results of research of higher education institutions, alongside formal teaching, direct (largely tacit) transmission of knowledge, mobility of people and technological outputs. Nevertheless it is extremely difficult to characterize the complete publication output of an institution through the most widely used databases for the evaluation of publication outputs of universities, like the Web of Science and, more recently, Scopus, which represent adequately only scholarly publications addressed to an international audience, mostly in English (Larédo, 2003). They allow for quite fine-grained analyses and they have been used widely to evaluate the output of universities as a whole, to produce international bibliometric rankings (Van Raan, 2004) and to evaluate productivity and efficiency of higher education institutions (Bonaccorsi et al., 2007). This despite methodological limitations as the limited coverage of most domains in social sciences and humanities, extremely problematic when assessing the publication activities of generalist universities, and the absence of coverage of national publications and of non-journal publications as books, reports, communications, which are likely to be not so relevant for academic international visibility but are extremely important at national and regional level (Gomez et al., 2007; Hicks, 2004; Nederhof, 2006).

## Background and purpose

Experiments for characterizing publication outputs using different sources have been carried out and a promising trend in this respect is the development of databases internal to higher education institutions. The Spanish GREC, developed by University of Barcelona, and UNIVERSITAS XXI, used by Carlos III University of Madrid as well as other 10 Spanish universities, UNISCIENCE database of the University of Lausanne, and the DARE, the Digital Academic Repository, and METIS developed in the Netherlands are some examples. So far the UNIPUB project would allow to assess the extent to which institutional databases can be used to produce simple indicators of publications at the level of whole higher education institutions. The methodology would also allow to point out other dimensions of universities analysis, as the academic vs. third-mission orientation, the international vs. national vs. regional orientation and variations of these orientations for the scientific domains considered. *The purpose is to build profiles of activities rather than to evaluate quality of research coherently with the positioning indicators rationale and with a policy orientation promoting the differentiation of higher education institutions according to different missions and social and geographical spaces rather than their competition along the unique dimension of the international academic reputation (Lepori, Barré & Filliatreau, 2008).*

## Methodology

The team is composed of four research groups: CERIS CNR in Italy, the University of Lugano in

Switzerland, IEDCYT and University Carlos III of Madrid, Spain, and the Rathenau Instituut in the Netherlands.

For the project purposes a sample of 8 universities, two for each country partner, have been selected according to the following criteria: presence of a good database, adequate time and disciplinary coverage, size and location. A standard description of selected institutional databases, in order to assess their usability for producing indicators on the publication activity, has been realized according to a common framework focusing on three main items: -institutional information (University, Department, individuals); -bibliographical information (author's information, scientific domain, condition and type of publication, language, aim of publication and type of audience); -procedures for construction and updating databases and guidelines for access and exploitation

Six main categories of indicators have been developed according to available data. The first concerns the orientation of the institution (national, international); the second focuses of the reference community (academics, policy makers, society); the third addresses differences in research productivity among research units, while the fourth refers to the differentiation of the publication within disciplines. The fifth category of indicators aims at characterizing the scientific production considering the type of knowledge (basic, applied, technology, education, communication, practical), and the type of publications (refereed or not, with Impact Factor or not). Finally the last category focuses on collaboration' dimension in research production. These indicators have been tested through a pilot test covering a one year period publications. This would allow to refine the methodology and the set of indicators.

Further ongoing activities are:

1. Large test on the datasets for a 2 years period publications;
2. Final definition of a set of indicators and methodological observations about their exploitation according to datasets characteristics;
3. Country based survey to assess the state of the art, the level of information, the use and main trends of institutional publications databases.

Information gathered both by the large test and the survey will be finalized in a methodological manual and a country report pointing out the state and perspectives of institutional publication databases and providing also recommendations and best practices for the organization. Finally, in the exploitation phase, results of the test phase will be discussed with representatives of involved universities and their finalization in papers or other publication will also be considered.

#### **Expected results**

The produced indicators should help higher education institutions to evaluate their own position and to assess their strengths. Indicators will also aim at covering different dimensions (as well as their combination) as the national vs. international orientation of the institution, the differentiations in research productions among and within research units and disciplinary areas, the type of knowledge produced and the reference community (academic vs. policy vs. companies). Expected outputs are: a methodological manual, a descriptive report of institutional databases and their development and two scholarly papers, the first on methodological issues and the second focused on the comparative analysis of the institutions of the sample.

#### **References**

- Bonaccorsi, A., Daraio, C., Lepori, B. & Slipersaeter, S. (2007). Indicators for the analysis of Higher Education Systems: some methodological reflections. *Research Evaluation*, 16(2), 66-78.
- Gomez, I. et al., (2007a). Structure and Research Performance of Spanish Universities. In D. Torres-Salinas & H.F. Moed (Eds.), *Proceedings of ISSI 2007* (pp. 334-345). Madrid: ISSI.
- Hicks, D. (2004). The Four Literatures of Social Science. In H. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 473-496). Dordrecht: Kluwer Academic Pubs..
- Larédo, Ph. (2003). University Research Activities: On-going Transformations and New Challenges. *Higher Education Management and Policy*, 15(1), 105-123.
- Lepori, B., Barré, R., & Filliatreau, Gh. (2008). New Perspectives and Challenges for the Design of S&T Indicators. *Research evaluation*, 17, 33-44.
- Nederhof, A.J. (2006). Bibliometric Monitoring of Research Performance in the Social Sciences and the Humanities: a review. *Scientometrics*, 66(1), 81-100.
- van Raan, A. (2004). Measuring Science. In H. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 19-50). Dordrecht: Kluwer Academic Pub.