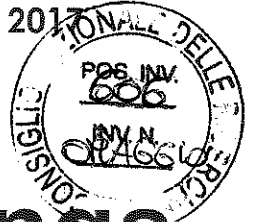


Nineteenth International Conference on Grey Literature

Public Awareness and Access to Grey Literature

National Research Council of Italy, CNR Rome • October 23-24, 2017



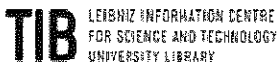
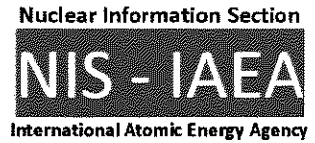
Proceedings

ISSN 1386-2316

ESTI
BIBLIOTECA

Collec.

Conference Host and Sponsors:



Inist



Istituto di Scienza e Tecnologie
dell'Informazione "A. Faedo"

NTK

50°6'14.083"N, 14°23'26.385"E
Národní technická knihovna
National Technical Library



ISTIT-BIB

GL19(2017) PRE

CNR PISA



WILEY



GL19 Program and Conference Bureau

TextRelease

Javastraat 194-HS, 1095 CP Amsterdam, Netherlands
www.textrelease.com • conference@textrelease.com
Tel/Fax +31-20-331.2420



Grey Literature and Research Assessment exercises: From the current criteria to the Open Science models

Silvia Giannini, Rosaria Deluca, Anna Molino, Stefania Biagioni
CNR, Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", Italy

1. Introduction

In the recent years the application of strategies, procedures and tools to evaluate the work of researchers have become subject of interest and their application is currently matter of discussion.

This topic is of major importance and will probably have a greater influence, since this type of exercise has strong political implications and determines a significant economic impact on the future of Universities and Research Institutions. Research assessment is a complicated business, as the design of a practical, informative process requires making decisions about which methodology should be used, which indicators calculated, and which data collected (Moed 2011). The evaluation procedure is also time-consuming and expensive, since the comprehension of the mechanisms underlying the research activity might be quite complex, and the results are not wholly predictable in some cases (Bianco 2010). Moreover, finding the right balance between two different kinds of approach makes the assessment exercise more problematic. On the one hand, the *quantitative* approach deals with the impact of the research, measuring the degree of diffusion of a certain idea in the scientific community. On the other hand, the *qualitative* approach determines the value of the research in terms of authenticity, relevance and clarity in the exposition of the results (Baccini 2011). Finally, the social and economic impacts have to be taken into account. Indeed, the evaluation procedures has recently acquired a greater importance due to the shortage of economic resources, therefore becoming a strategic instrument for the quality assessment of Universities and Research bodies.

The assessment exercises are regulated at national level and are carried out in different European countries such as France, United Kingdom and The Netherlands. The English RAE - Research Assessment Exercise is the oldest performance-based research funding system (Rebora, Turri 2013). The Research Assessment Exercises (RAE) have been held in the UK since 1986.

The RAE has been a benchmark for the relatively recent Italian assessment exercise, as the first steps in this direction were taken in Italy at the beginning of the '90s. In 1993 the *Osservatorio per la valutazione del sistema universitario* was created, becoming active only in 1996. In 1998 the *Comitato di Indirizzo per la Valutazione della Ricerca* (CIVR) was set up, and in 1999 the *Comitato Nazionale per la Valutazione del Sistema Universitario* (CNVSU) was created as successor of the *Osservatorio*, becoming officially operative in 2000 (Rubele 2012). The first research assessment exercise has been legislated in 2003 and entrusted to CIVR. The Committee analyzed the research products of 2001-2003 in order to evaluate the scientific performance of Universities, as well as state and private Research Institutions. Three years later the CIVR and other committees have been replaced by a specific agency named National Agency for the Evaluation of Universities and Research Institutes [*Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca* (ANVUR)].

The Agency aims to « [...] rationalize the system of assessment of the quality of Universities, state and private Research Institutions beneficiary of public funds [...] » « The results of these activities managed by ANVUR represent one of the criteria to assign the state funds to Universities and Research Institutions ».

More in detail, the Agency evaluates the quality of the processes, the results, and the products released by Universities and Research Bodies and defines criteria and methods for the evaluation of the university branches and of the course of studies.

It cooperates with other scientific international committees operating in the field of research assessment and with the European Union.

At the present time, ANVUR has completed two evaluation exercises of the quality of the research named Evaluation of Research Quality [*Valutazione della Qualità della Ricerca* (VQR)]: the first one spans the years 2004 – 2010 (VQR1); the second from 2011 to 2014 (VQR2).

Despite the shared belief among the different scientific communities about the importance of the research assessment, there is no such agreement regarding the purposes and the procedures for its realization (Galimberti 2012). The international debate on methods and critical issues of the



research assessment practices has become more intense, registering an increased number of negative judgements about the procedures currently applied. Although the ultimate goal is obtaining excellence, quality and the greater impact on society, the parameters currently in use to evaluate the products of the research and the consequences of such measuring practices have a negative influence on the attitude of the researchers towards this topic (Galimberti 2017). Criticisms are mainly concentrated on quantitative measurements, because of their improper use of a range of commercial bibliometric indicators. The *ROARS – Return On Academic Research* association, whose aim is particularly focused on the policies for the evaluation of the research, dedicates large part of its blog to the major issues concerning the ANVUR evaluation procedures, promoting various initiatives that encourage fairer practices and more responsible behaviors in the research assessment. Proposals like *DORA – Declaration on Research Assessment*, as well as the *Leiden Manifesto for the Research Metrics* aim at defining criteria that would represent more widely the complexity of the research analysis.

The conceptual challenges taken on by the Open Science (OS) movement may be crucial for the evolution of these matters. The OS is multidimensional; approaches and skills of various kinds are necessary for its fulfilment and for the achievement of objectives such as openness, sharing, transparency and quality¹. As a matter of fact, the term Open Science indicates the opportunity to freely contribute to the knowledge production, sharing the outcomes and being inclined to the cooperation with the whole scientific community (Delfanti 2008). The expression encompasses subject matters like the *Open Access (OA)*, the free access to scientific publications; the *Open Data*, as it promotes the dissemination of the research data; the *Open-Notebook Science*, encouraging the online sharing of lab notes and raw data (Stafford 2010). Moreover, the OS introduces the ideas of *Open Learning*, a new, customizable teaching plan, independent of time and space; and the *Open Research and Citizen Science*, making science available to all the citizens of the *knowledge-society*, where they have the right to access the most advanced researches (Gioè 2016).

The work looks at the environment of VQR in order to understand the organizational set-up, the operational models, the scientific Areas involved in the process and the selection and evaluation criteria of the research products. More in detail, our work analyzes and compares the evaluation exercises conducted in Italy with the aim of verifying if and how *Grey Literature (GL)* is involved in the research evaluation processes. The article checked the types of products admissible for the research assessment and those actually presented by the researchers of Universities and Research Institutions. We measured the products from a quantitative point of view and observed their ramification in the different disciplinary fields rather than their transformation during the period of time taken into consideration. At the same time, we focused on the OS movement in order to understand what could be its role within the research assessment exercises and how it could affect the future of scholarly scientific communication.

2. The VQRs framework

2.1 Organization and methods

We consulted the public documentation provided by ANVUR, consisting in a set of preliminary documents and a set of documents produced as final reports for the two evaluations².

The two exercises done in Italy were addressed to the assessment of the research conducted in both state and private universities, as well as in public research bodies and other public and private subjects whose research activities are funded by the government. Researchers, assistant professors, associate professors, full professors, all being on duty at the time the evaluation

¹ <https://sites.google.com/site/scienzaapertaricercamiigliore/programma>

² Announcement_VQR 2004-2010, 17 July 2011,

http://www.anvur.org/index.php?option=com_content&view=article&id=122&Itemid=305&lang=it.

Specific criteria by the expert groups selected by the Agency,

http://www.anvur.org/index.php?option=com_content&view=article&id=32&Itemid=372&lang=it.

Area Reports VQR_2004-2010, 30 June 2010, <http://www.anvur.org/rapporto/>.

Announcement_VQR 2011-2014, 11 November 2015,

http://www.anvur.org/index.php?option=com_content&view=article&id=825&Itemid=599&lang=it

Specific criteria by the expert groups selected by the Agency,

http://www.anvur.org/index.php?option=com_content&view=article&id=841&Itemid=601&lang=it

Area Reports VQR_2011-2014, 21 February 2017, <http://www.anvur.org/rapporto-2016/>.



started, have been appraised. The number of research products to be assessed was indicated with reference to each individual evaluated.

In both exercises, a taxonomy based on macro disciplinary areas (Tables 1-2) was used, each subdivided into Scientific Disciplinary Sectors [*Settori Scientifico-Disciplinari (SSD)*]³.

| Macro-Areas VQR1 | | Macro-Areas VQR2 | |
|------------------|---|------------------|---|
| Area | Description | Area | Description |
| 1 | Computer science and Mathematics | 1 | Computer science and Mathematics |
| 2 | Physics | 2 | Physics |
| 3 | Chemistry | 3 | Chemistry |
| 4 | Earth sciences | 4 | Earth sciences |
| 5 | Biology | 5 | Biology |
| 6 | Medicine | 6 | Medicine |
| 7 | Agricultural and veterinary sciences | 7 | Agricultural and veterinary sciences |
| 8 | Civil engineering and Architecture | 8a | Architecture |
| 9 | Industrial and computer engineering | 8b | Civil engineering |
| 10 | Antiquity, philological-literary and historical-artistic sciences | 9 | Industrial and computer engineering |
| 11 | Historical, philosophical, psychological and pedagogical sciences | 10 | Antiquity, philological-literary and historical-artistic sciences |
| 12 | Legal sciences | 11a | Historical, philosophical and pedagogical sciences |
| 13 | Economics and Statistics sciences | 11b | Psychology |
| 14 | Social and political sciences | 12 | Legal science |
| | | 13 | Economics and Statistics sciences |
| | | 14 | Social and political sciences |

Tables 1-2

From the comparison between the tables is clear the substantial overlapping between macro-areas in the two exercises, with the exception of Areas 8 and 11, split in two sub-categories in VQR2, the number of Areas going from 14 to 16.

The ANVUR constituted Groups of experts for the evaluation⁴ [*Gruppi di Esperti della Valutazione (GEV)*] for each macro-area, composed of both Italian and foreign qualified experts.

Sub-groups of specialists were created within those GEVs dedicated to highly heterogeneous disciplinary areas and characterized by a large number of products to be evaluated. Each GEV has a nominated Coordinator.

The judgment on the quality of the products was based on the following general criteria:

1. VQR 2004 – 2010: Relevance – Originality/Innovation – Internationalization
2. VQR 2011 – 2014: Originality – Methodological rigor – Impact (recognized or potential)

Among the general principles established by the Agency, each GEV set its own criteria, which allowed the Groups to define different quality steps, as illustrated in Table 3. It is evident that the quality steps and their related scores were partially revisited in the second exercise, where the class *Plagiarism/Fraud* was deleted.

| VQR 2004-2010 | | VQR 2011-2014 | |
|---------------------|-------|------------------|-------|
| Class of merit | Score | Class of merit | Score |
| A. Excellent | 1 | A. Excellent | 1 |
| B. Good | 0.8 | B. High-level | 0.7 |
| C. Acceptable | 0.5 | C. Fair | 0.4 |
| D. Limited | 0 | D. Acceptable | 0.1 |
| E. Not evaluable | -1 | E. Limited | 0 |
| F. Plagiarism/Fraud | -2 | F. Not evaluable | 0 |

Table 3 - Quality steps

The two assessments were based on an *informed peer-review*. In the technical areas, this technique was based on a combination of bibliometric parameters and peer-review. The papers

³Cfr. "Evaluation of Research Quality 2011 – 2014 (VQR 2011 – 2014). ANVUR final report".

For example: Area 1 - *Computer science, Logic, Algebra, Geometry, Complementary maths, Mathematical analysis...* Area 2 - *Experimental physics, Theoric physics, (mathematical models and methods), Physics of matter, Nuclear and subnuclear physics, Astronomy and Astrophysics, Physics for the Earth system...*

⁴cfr. "Evaluation of Research Quality 2011 – 2014 (VQR 2011 – 2014). ANVUR final report".



in journals were evaluated through indicators derived from commercial citation databases⁵. An algorithm taking into account the quality of the journal as well as the number of the article's citations was exploited during the assessment exercise. Other categories or more recent products with a lower number of citations were evaluated by *peer-review*. In the Humanities and Law disciplines the bibliometric analysis is less frequent, as databases such as WoS or Scopus do not have a widespread coverage. In these cases, the evaluation carried on by the GEVs was only peer-reviewed, or alternatively based on a technique named *informed peer review*, which relies on different evaluation methods that make use of various information tools, like editorial peculiarities, reviews, translations, awards.

The GEVs can entrust the *peer-review* to external reviewers or conduct it inside the group itself, and it must define specific assessment criteria to harmonize the different evaluation methods. Moreover, the GEV is responsible of the final result of the assessment.

2.2 The VQRs "objects"

The specifications given by ANVUR for the assessment of the research products by the GEVs were enhanced in the VQR2.

In VQR1 the categories listed below were defined for the classification of all identified research products, not providing further details on the characteristics and subject matters⁶:

Documentary categories in VQR1

- a) Papers in journals
- b) Books, chapters of books, and conference proceedings provided with ISBN
- c) Critical editions, translations, and scientific comments
- d) Patents
- e) Compositions, drawings, design, performances, exhibitions and organized expositions, handwork, prototypes, artworks and related designs, databases and software, thematic maps

Documentary categories in VQR2

1) *Scientific monographs*

Research monograph, Concordance, Scientific comment, Annotated bibliography, Critical editions of texts, Critical editions of excavations, Publication of unedited sources, Critical manuals (not for educational purpose only), Grammars and science dictionaries, Translations of books (upon GEV's decision).

2) *Articles in journals*

Scientific paper, Review essays, Letters, Contribution to a Forum upon invitation of the editorial staff, Case notes, Translations in journal.

3) *Contributions to books*

Scientific articles in peer-reviewed conference proceedings, Foreword and afterword in the form of essay, Curatorship of books with introductory essay, Catalogues with introductory essay, Critical entries in dictionaries or encyclopedias, Translations in book (upon GEV's decision), Catalographic records, bibliography or corpora.

4) *Other types of scientific products*

Compositions, Drawings, Architectonic projects, Performances, Exhibitions, Prototypes of art and related projects/designs, Database and software, Thematic maps, Psychological evaluations, Audiovisual material.

5) *Patents*

The category Patents is always considered as evaluable, but it may be attributed to class A or B only if internationally renowned or licensed.

Not admissible products in VQR1

- *Editorial and curatorial activities*
- *Conference abstracts (even if published in journals)*
- *Texts or software used for educational and dissemination purpose only*
- *Routine or laboratory tests*
- *Internal technical reports*

⁵ Web of Science (WoS) by Clarivate Analytics; Scopus by Elsevier; Mathematical Reviews on the web (MathSciNet) by The American Mathematical Society.

⁶ Cit. Announcement_VQR 2004-2010, 17 July 2011.



Not admissible products in VQR2

- *Manuals and texts for educational purpose only*
- *Review of a single article not showing any critical analysis of the literature on the topic*
- *Short, non-original encyclopedia or dictionary entries*
- *Short, non-original case notes*
- *Short catalographic records*

The merger of the research products in documentary classes highlights some differences in their own composition. The second assessment exercise was more inclined towards categories such as articles in journals and contributions in books, as alternative to the mere scientific articles.

Indeed, in the first class of VQR2 we find sub-categories like *review essays, letters, case notes, contribution to a forum, and translations*, while the second includes *essay collections, concordances, bibliographies, critical manuals, grammars, corpora, and catalographic records*, as well as the entries *critical editions, translations, and scientific comments*, considered as autonomous in the previous evaluation.

The category *e)* of VQR1 was merged into *Other types of scientific products*, which includes the additional entries *architectonic projects, psychological evaluations, audiovisual material*, while *design and handwork* were excluded.

3. Tracking the Grey Literature

The list of documentary typologies evaluated does not allow a thorough classification of the *grey* products. The Implementation Announcements of the two VQRs do not specify any requirement concerning the manners of publication; they only define the categories containing works circulated through the conventional distribution channels as well as products disseminated out of the traditional publishing chains. The fact that the documents have ISBN and/or ISSN code, or that they have been successfully peer-reviewed, does not grant their publication by a commercial publishing company. The procedures, the evaluation principles and the assessment tools only are able to determine if the majority of the products belong to the traditional literature.

The analysis of the GEVs' assessment criteria and of the FAQs published by ANVUR contributed to the understanding of some concepts and supported the interpretation of some results, facilitating only partially the process of identification of the *Grey Literature* inside the various documentary typologies. For these reasons, the process of identification of the Grey Literature was based on the following considerations:

- I. The evaluation exercises mainly founded their bibliometric analysis on the contents of the two commercial databases WoS and Scopus⁷. This is due to the fact that the international scientific community makes extensive use of them for the assessment of the scientific levels of the journals. Although the databases found their bibliometric indicators on different parameters, they are both based on the calculation of the number of citations. The majority of the literature indexed by the two databases is published by commercial publishing companies; only a small percentage of products ascribable to the *Grey Literature* are indexed in Scopus.
- II. The use of the databases restricts the contents only to the references indexed (based on ownership criteria). Especially for the *papers in journals*, the algorithm for the assessment takes into account the number of citations of a paper and the corresponding bibliometric indicator of the journal inside one or more disciplinary classes defined in the two databases⁸. The more commonly evaluated research products like *books, papers in books and papers in proceedings* are assessed taking into account their occurrence in the databases. Moreover, the GEVs reserve the *peer-review* only to the products not indexed in the databases.
- III. In both evaluation exercises, it is made reference frequently to the products published by commercial publishing companies, especially if they are renown at international levels. In some of the GEVs' criteria it is specified that self-published products are not evaluated. Moreover, it is clearly stated that products accepted but not yet published are not taken into consideration.
- IV. The GEVs' criteria specify that the products listed in *e)* in the VQR1 Announcement would be evaluated making reference to their characteristics, not to their formal publication. In VQR2 Announcement these products are grouped in *Other types of scientific products*. Therefore,

⁷ For the evaluation of Areas 12 and 13 in VQR1 a reference is made for the use of Google Scholar as assessment parameter for the evaluation of the journal, if not indexed in WoS or Scopus. This is not taken into account in VQR2.

⁸ WoS Subject Category (SC) and Scopus All Science Journal Classification (ASJC).



the eligibility conditions of these products are clearly expressed in the criteria used for specific disciplinary areas. For instance, in some cases it is specified that products like *drawings, prototypes of art, or architectonic projects* may be evaluated if they have been published or worthy of mention/winner of prizes in a competition. At the same time, the *thematic maps* may be assessed if their theme is evidenced using particular procedures and graphic adaptations, allowing the immediate understanding of the distribution, differentiations, and correlations of one or more phenomena. In the Area 02 (Physical sciences) the item *composition* was evaluated by the GEV with its products *handwork, devices and prototypes*, along with the entries *exhibition, database and software*, whereas the types *drawings, architectonic projects, performance, prototypes of art* and related projects, as well as the *thematic maps*, were not assessed. More details for the software products have been given in Area 9 (Industrial and computer engineering).

With specific reference to what listed above, we did not identify *grey products* within categories such as *papers in journals, books or proceedings* as well as products belonging to other assimilated categories such as *curatorship, critical editions, and translations*. On the other hand, we agreed on ascribing some groups of products to the *non-conventional literature*, including in this range also entries with a certain degree of uncertainty with respect to their arrangements for publication. This is the case of products measured in VQR2, such as: *concordance, publication of unedited sources, entries, catalographic record*. As a matter of fact, the indications given by the GEVs neither completely clarify the nature and the characteristics of these products, nor explicate their inclusion in the list of works conventionally published.

4. Analysis of data and results

Tables 4 and 5 show the whole range of products evaluated in the two exercises, sub-divided in the 14 disciplinary areas of reference⁹. Some of the original tables of VQR1 do not list the number of the products but the percentage calculated by the GEVs only, here reported. This is the case of Areas 2 (Physics), 4 (Earth sciences), 11 (Historical, philosophical, psychological and pedagogical sciences), and 12 (Legal sciences). Moreover, in the same exercise the data relative to the Area 11 are sub-divided between products evaluated through bibliometric analysis and those *peer-reviewed*. In VQR2 this sub-division is not present.

In both exercises and for each disciplinary area, the most significant numbers are referred to the entries *papers in journals, papers in books e papers in proceedings*.

| Categories | VQR 2004-2010 | | | | | | | | | | | | | | | |
|--|---------------|--------------|--------------|-------------|--------------|--------------|--------------|-------------|--------------|--------------|-------------------|-----------------|--------------|--------------|-------------|--|
| | Area 01 | Area 02 | Area 03 | Area 04 | Area 05 | Area 06 | Area 07 | Area 08 | Area 09 | Area 10 | Area 11- bibli | Area 11- Bib | Area 12 | Area 13 | Area 14 | |
| | % | % | % | % | % | % | % | % | % | % | % | % | % | % | % | |
| Abstract (in journals or in proceedings) | 0.06 | | 0.09 | | 0.21 | 0.07 | | | | | | | | | | |
| Case notes | | | | | | | | | | | | | 0.56 | | | |
| Composition | | | | | | | | 0.04 | | | | | | | | |
| Critical edition | 0.03 | | | | 0.01 | | | 0.07 | | 0.92 | 0.56 | | | | | |
| Curatorship | 0.07 | | 0.04 | | 0.02 | | 0.26 | 2.69 | | 1.76 | 1.84 | 0.52 | 0.73 | | 2.80 | |
| Database | 0.01 | | | 0.20 | 0.01 | | | 0.01 | | | | | | | | |
| Design | 0.02 | | | | | | | | | | | | | | 0.13 | |
| Entry (in dictionary o encyclopedia) | | | | | 0.01 | | | 0.03 | | | | | 0.29 | | | |
| Exhibition | 0.01 | | | | | | | 0.02 | | | | | | | | |
| Foreword/Afterword | | | | | | | | 0.02 | | | | | | | | |
| Handwork | 0.02 | 0.70 | 0.01 | | | | | 0.08 | | | | | | | | |
| Maps | | | | 0.20 | | | | | | | | | | | | |
| Monograph or scientific treaty | 1.27 | 0.20 | 0.21 | 1.03 | 0.38 | 0.52 | 1.22 | 14.56 | 0.75 | 22.69 | 33.06 | 9.26 | 25.88 | 12.75 | 33.63 | |
| Other | 0.13 | 0.30 | 0.01 | 2.28 | 0.04 | 0.07 | 0.82 | 1.09 | 0.44 | 0.80 | 0.61 | 0.22 | 0.64 | 1.01 | 0.46 | |
| Paper in books | 3.27 | 0.60 | 0.39 | 5.06 | 1.26 | 1.53 | 4.65 | 23.60 | 2.46 | 32.80 | 32.86 | 11.57 | 36.00 | 19.88 | 32.59 | |
| Paper in journals | 86.11 | 93.40 | 98.45 | 85.96 | 96.92 | 96.94 | 87.54 | 43.69 | 81.68 | 26.50 | 23.99 | 77.22 | 32.76 | 62.45 | 28.63 | |
| Paper in proceedings | 8.84 | 4.70 | 0.40 | 4.80 | 0.89 | 0.74 | 5.51 | 13.82 | 14.16 | 14.19 | 7.09 | 1.21 | 3.13 | 3.91 | 1.90 | |
| Patent | 0.10 | 0.20 | 0.39 | 0.18 | 0.26 | 0.13 | | 0.14 | 0.51 | | | | | | | |
| Prototype of art and related project | | | | | | | | 0.01 | | | | | | | | |
| Software | 0.07 | | 0.02 | 0.28 | 0.01 | | | | | | | | | | | |
| Translation | | | | | | | | | | 0.33 | | | | | | |
| Total | 10685 | 19773 | 11608 | 8433 | 16407 | 26713 | 10004 | 9533 | 16347 | 14073 | 9513 | 3639 | 11882 | 11941 | 4327 | |

Table 4 - Research products in VQR1 1

In VQR1 the percentage indicates that the *papers in journals* prevail in almost all disciplinary areas, in some cases reaching nearly 100%.

⁹ The tables are a rework of those contained in the area reports produced by the GEV. These can be viewed at URLs <http://www.anvur.org/rapporto/> (VQR1) and <http://www.anvur.org/rapporto-2016/> (VQR2).



In Areas 10 (Antiquity, philological-literary and historical-artistic sciences), 11nb (Historical, philosophical, psychological and pedagogical sciences), 12 (Legal sciences), and 14 (Social and political sciences) only the *monographs* and the *papers in books* show significant percentages, proving the most widely used modalities in the scientific communication in these sectors. The *papers in proceedings* represent the largest number in Areas 8 (Civil engineering and Architecture), 9 (Industrial and computer engineering), and 10 (Antiquity, philological-literary and historical-artistic sciences).

| VQR 2011-2014 | | | | | | | | | | | | | | | | |
|--|---------|---------|---------|---------|---------|---------|---------|---------|----------|---------|---------|----------|----------|---------|---------|---------|
| Categories | Area 01 | Area 02 | Area 03 | Area 04 | Area 05 | Area 06 | Area 07 | Area 08 | Area 08b | Area 09 | Area 10 | Area 11a | Area 11b | Area 12 | Area 13 | Area 14 |
| | % | % | % | % | % | % | % | % | % | % | % | % | % | % | % | % |
| Abstract (in journals or in proceedings) | | | | | | | | | | | | | | | | 0.02 |
| Architectonic project | | | | | | | | 1.45 | | | | | | | | |
| Bibliographic/Catalographic record, corpus | | 0.09 | | | 0.06 | 0.03 | 0.03 | | | | 0.09 | | | | | |
| Bibliography | | | | | | | 0.01 | | | | 0.03 | 0.05 | | | | |
| Case notes | | | | | 0.04 | 0.01 | | 0.03 | | | | | 0.04 | 1.52 | | |
| Composition | | 0.75 | 0.01 | | | | | 0.03 | | 0.05 | | | | | 0.01 | |
| Concordance | | | | | | | | | | | 0.05 | | | | | |
| Critical edition | | 0.02 | | | | | | 0.09 | | | 2.58 | 0.80 | | 0.09 | 0.05 | 0.27 |
| Curatorship | 0.15 | | 0.03 | 0.11 | 0.02 | 0.02 | 0.09 | 5.01 | | 0.02 | 0.11 | 1.45 | 0.35 | 0.47 | 0.38 | 1.88 |
| Database | | 0.04 | | 0.18 | 0.03 | | 0.08 | 0.03 | | 0.04 | 0.05 | 0.02 | | 0.02 | | |
| Design | | | | | | | | 0.12 | | | | | | | | |
| Entry (in dictionary or encyclopedia) | | | | | 0.01 | | | 0.20 | | 0.09 | 0.23 | 0.23 | 0.13 | 1.37 | 0.10 | 0.03 |
| Exhibition | | 0.11 | | 0.07 | 0.06 | 0.01 | 0.03 | 0.14 | 0.04 | 0.03 | 0.05 | | | | | |
| Handwork | | | | | | | | | | | | | | | | |
| Maps | | | | 0.23 | | | 0.04 | | | | | | | | | |
| Monograph or scientific treaty | 1.35 | 0.22 | 0.26 | 0.88 | 0.20 | 0.36 | 0.89 | 23.29 | 0.60 | 0.63 | 19.64 | 27.24 | 3.82 | 26.17 | 8.67 | 24.13 |
| Other | 0.21 | | | 0.18 | 0.01 | 0.01 | | 0.52 | 0.04 | | 0.17 | 0.07 | 0.09 | | 0.55 | |
| Paper in books | 2.95 | 0.73 | 0.58 | 3.18 | 0.79 | 1.01 | 3.69 | 28.91 | 2.82 | 1.66 | 35.91 | 30.48 | 6.50 | 30.57 | 14.68 | 32.01 |
| Paper in journals | 87.92 | 96.85 | 98.12 | 91.78 | 97.23 | 97.92 | 91.31 | 26.42 | 88.63 | 88.71 | 32.12 | 34.79 | 88.09 | 38.04 | 72.71 | 40.83 |
| Paper in proceedings | 7.32 | 0.70 | 0.41 | 3.23 | 0.62 | 0.52 | 3.39 | 12.96 | 7.66 | 8.28 | 6.78 | 3.94 | 0.75 | 1.61 | 2.68 | 0.27 |
| Patent | 0.02 | 0.18 | 0.33 | 0.07 | 0.13 | 0.04 | 0.19 | 0.29 | 0.18 | 0.38 | | | | | 0.02 | |
| Performance | | | | | 0.03 | | | 0.03 | | | 0.07 | | | | | 0.07 |
| Preface/Postface | | | 0.01 | | | | | 0.20 | | 0.01 | 0.37 | 0.38 | 0.04 | 0.05 | 0.04 | 0.30 |
| Prototype of art and related project | | | | | | | | 0.12 | | 0.03 | | | | | | |
| Publication of unedited sources | | | | | | | | 0.09 | | | 0.15 | 0.24 | | 0.01 | 0.01 | 0.07 |
| Review (in book or in journal) | | 0.12 | 0.25 | | 0.76 | 0.07 | 0.32 | | 0.04 | 0.10 | 0.10 | 0.15 | 0.09 | | 0.06 | 0.03 |
| Scientific comment | | | | | | | | | | | 0.56 | 0.10 | | | 0.02 | 0.07 |
| Software | 0.07 | 0.21 | | 0.09 | | | | 0.09 | | 0.04 | 0.01 | | 0.09 | | | |
| Translation | | | | | 0.01 | 0.02 | | | | | 0.83 | 0.08 | | 0.05 | 0.01 | 0.03 |
| Totale | 6062 | 10588 | 6897 | 4430 | 10986 | 16693 | 7541 | 3456 | 2832 | 11564 | 8744 | 6123 | 2276 | 8488 | 8385 | 2971 |

Table 5 - Research products in VQR2

In VQR2 the *papers in journals* still represents the more widely evaluated category, with the exception of the Areas 8a (Architecture) and 10 (Antiquity, philological-literary and historical-artistic sciences), where the *papers in books* dominate.

In the other documentary typologies, we find limited percentages of specific products for each disciplinary area, especially in VQR2. Among the products most frequently submitted for the evaluation are *curatorship* and *critical editions* in VQR1, *reviews* and *translations* in VQR2. However, the presence of products such as *curatorship*, *critical editions*, *abstracts*, *reviews*, *entries*, *catalographic record*, *translations*, *corpora*, etc. is determined by the criteria adopted by the GEVs. Indeed, each GEV had the possibility of defining more in detail the criteria determining the admission of the products to the evaluation, considering also the relevance in each research area, and the procedures applied to value their judgments. In some cases, the GEVs made different choices, including, for instance, in the entries *papers in journals* and *papers in books* products like *forwards/afterwards*, *lexicons*, *catalogues*, *guides*, *concordances*, *critical edition* and *publication of unedited sources* so.

Table 6 shows the frequency of GL by Area in VQR1 and VQR2.



| VQR 2004-2010 | | | | VQR 2011-2014 | | | |
|---------------|----------------|-------------------|--------------------------|---------------|----------------|-------------------|--------------------------|
| Areas | N. of products | N. of GL products | Frequency of GL products | Areas | N. of products | N. of GL products | Frequency of GL products |
| Area01 | 10685 | 38 | 0.36 | Area01 | 6062 | 18 | 0.30 |
| Area02 | | na | | Area02 | 10588 | 145 | 1.37 |
| Area03 | 11608 | 25 | 0.22 | Area03 | 6897 | 24 | 0.35 |
| Area04 | na | na | | Area04 | 4430 | 36 | 0.81 |
| Area05 | 16407 | 51 | 0.31 | Area05 | 10986 | 40 | 0.36 |
| Area06 | 26713 | 54 | 0.20 | Area06 | 16693 | 14 | 0.08 |
| Area07 | 10004 | 82 | 0.82 | Area07 | 7541 | 27 | 0.36 |
| Area08 | 9533 | 148 | 1.55 | Area08a | 3456 | 105 | 3.04 |
| Area09 | 16347 | 156 | 0.95 | Area08b | 2832 | 7 | 0.25 |
| Area 10 | 14073 | 112 | 0.80 | Area09 | 11564 | 69 | 0.60 |
| Area 11-nbib | | na | | Area 10 | 8744 | 58 | 0.66 |
| Area 11-bib | | na | | Area 11a | 6123 | 19 | 0.31 |
| Area 12 | | na | | Area 11b | 2276 | 8 | 0.35 |
| Area 13 | 11941 | 121 | 1.01 | Area 12 | 8488 | 247 | 2.91 |
| Area 14 | 4327 | 20 | 0.46 | Area 13 | 8385 | 57 | 0.68 |
| | | | | Area 14 | 2971 | 3 | 0.17 |
| Total | 131638 | 807 | 0.61 | Total | 118036 | 877 | 0.74 |

Table 6- Frequency by Areas

The frequency of the *Grey Literature* is 0.61 in VQR1 and 0.74 in VQR2. However, this is a rough estimation, as in the VQR1 calculations some Areas 2 (Physics), 4 (Earth sciences), 11 (Historical, philosophical, psychological and pedagogical sciences), and 12 (Legal sciences) had to be excluded, because the frequency of each product had not been stated in the GEVs' final reports.

Tables 7 and 8 show the products of GL by Area in VQR1 and VQR2.

| GL categories | VQR 2004-2010 | | | | | | | | | | | | | | |
|--------------------------------------|---------------|--------------|--------------|-------------|--------------|--------------|--------------|-------------|--------------|--------------|--------------|-------------|--------------|--------------|-------------|
| | Area 01 | Area 02 | Area 03 | Area 04 | Area 05 | Area 06 | Area 07 | Area 08 | Area 09 | Area 10 | Area 11-nbib | Area 11-bib | Area 12 | Area 13 | Area 14 |
| | % | % | % | % | % | % | % | % | % | % | % | % | % | % | % |
| Case notes | | | | | | | | | | | | | | | 0.56 |
| Composition | | | | | | | | 0.04 | | | | | | | |
| Database | 0.01 | | | 0.20 | 0.01 | | | 0.01 | | | | | | | |
| Design | 0.02 | | | | | | | 0.13 | | | | | | | |
| Entry (in dictionary o encyclopedia) | | | | | 0.01 | | | 0.03 | | | | | 0.29 | | |
| Exhibition | 0.01 | | | | | | | 0.02 | | | | | | | |
| Handwork | 0.02 | 0.70 | 0.01 | | | | | 0.08 | | | | | | | |
| Maps | | | | 0.20 | | | | | | | | | | | |
| Other | 0.13 | 0.30 | 0.01 | 2.28 | 0.04 | 0.07 | 0.82 | 1.09 | 0.44 | 0.80 | 0.61 | 0.22 | 0.64 | 1.01 | 0.46 |
| Patent | 0.10 | 0.20 | 0.59 | 0.18 | 0.26 | 0.13 | | 0.14 | 0.51 | | | | | | |
| Prototype of art and related project | | | | | | | | 0.01 | | | | | | | |
| Software | 0.07 | | 0.02 | 0.28 | 0.01 | | | | | | | | | | |
| Totale | 10685 | 19773 | 11608 | 8433 | 16407 | 26713 | 10004 | 9533 | 16347 | 14073 | 9513 | 3639 | 11682 | 11941 | 4327 |

Table 7 - Grey products in VQR1 by Areas

In VQR1 the most relevant percentages are those referred to:

- Area07 - Agricultural and veterinary science
- Area08 – Civil engineering and architecture
- Area10 – Antiquity, philological-literary and historical and artistic sciences
- Area 12 – Legal Sciences
- Area13 – Economics and statistics sciences



| GL categories | VQR 2011-2014 | | | | | | | | | | | | | | | |
|--|---------------|--------------|-------------|-------------|--------------|--------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | Area 01 | Area 02 | Area 03 | Area 04 | Area 05 | Area 06 | Area 07 | Area 08a | Area 08b | Area 09 | Area 10 | Area 11a | Area 11b | Area 12 | Area 13 | Area 14 |
| | % | % | % | % | % | % | % | % | % | % | % | % | % | % | % | % |
| Architectonic project | | | | | | | | 1.45 | | | | | | | | |
| Bibliographic/Catalographic record, corpus | | 0.09 | | | 0.06 | 0.03 | 0.03 | | | | 0.09 | | | | | |
| Case notes | | | | | 0.04 | 0.01 | | 0.03 | | | | | 0.04 | 1.52 | | |
| Composition | | 0.75 | 0.01 | | | | | 0.03 | | 0.05 | | | | | 0.01 | |
| Database | | 0.04 | | 0.18 | 0.03 | | 0.08 | 0.03 | | 0.04 | 0.05 | 0.02 | | 0.02 | | |
| Design | | | | | | | | 0.12 | | | | | | | | |
| Entry (in dictionary or encyclopedia) | | | | | 0.01 | | | 0.20 | | 0.03 | 0.23 | 0.23 | 0.13 | 1.37 | 0.10 | 0.03 |
| Exhibition | | 0.11 | | 0.07 | 0.06 | 0.01 | 0.03 | 0.14 | 0.04 | 0.03 | 0.05 | | | | | |
| Handwork | | | | | | | | | | | | | | | | |
| Maps | | | | 0.23 | | | 0.04 | | | | | | | | | |
| Other | 0.21 | | | 0.18 | 0.01 | 0.01 | | 0.52 | 0.04 | | 0.17 | 0.07 | 0.09 | | 0.55 | |
| Patent | 0.02 | 0.18 | 0.33 | 0.07 | 0.13 | 0.04 | 0.19 | 0.29 | 0.18 | 0.38 | | | | | 0.02 | |
| Performance | | | | | 0.02 | | | 0.03 | | | 0.07 | | | | | 0.07 |
| Prototype of art and related project | | | | | | | | 0.12 | | 0.03 | | | | | | |
| Software | 0.07 | 0.21 | | 0.09 | | | | 0.09 | | 0.04 | 0.01 | | 0.09 | | | |
| Totale | 6062 | 10588 | 6897 | 4430 | 10986 | 16693 | 7541 | 3436 | 2832 | 11564 | 8744 | 6123 | 2276 | 8488 | 8385 | 2971 |

Table 8 - Grey products in VQR2 by Areas

In VQR2 the most relevant percentages are those referred to:

- Area 2 – Physics
- Area8a – Architecture
- Area 12 – Legal Sciences
- Area 13 – Economics and statistics sciences

A specific reference must be made to the category *other*, as it is not really clear which products includes. In GEVs' final reports, the entry is defined as the incorporation of a collection of different products, their number too small to be treated separately and impossible to merge into other categories. This item is present in all disciplinary areas in VQR1 and in multiple areas in VQR2, where the documentary categories are wider. Therefore, it is possible that some products have been classified more properly.

The following table shows the percentage distribution of GL by Areas and years.

| Areas | VQR 2004-2010 | | | | | | | VQR 2011-2014 | | | |
|---|---------------|------|------|------|------|------|------|---------------|------|------|------|
| | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| Area01 - Computer science and Mathematics | 0.15 | 0.36 | 0.45 | 0.31 | 0.31 | 0.61 | 0.25 | 0.22 | 0.27 | 0.44 | 0.25 |
| Area 02 - Physics | | | | | | | | 0.88 | 1.24 | 1.03 | 2.65 |
| Area03 - Chemistry | 0.34 | 0.18 | | 0.50 | 0.06 | 0.27 | 0.17 | 0.18 | 0.29 | 0.22 | 0.71 |
| Area 04 - Earth sciences | | | | | | | | 0.58 | 0.95 | 0.78 | 0.91 |
| Area05 - Biology | 0.05 | 0.19 | 0.27 | 0.22 | 0.45 | 0.27 | 0.63 | 0.34 | 0.39 | 0.27 | 0.46 |
| Area06 - Medicine | 0.13 | 0.12 | 0.08 | 0.13 | 0.12 | 0.09 | 0.22 | 0.05 | 0.13 | 0.07 | 0.08 |
| Area07 - Agricultural and veterinary sciences | 1.43 | 0.81 | 0.58 | 1.19 | 0.68 | 0.45 | 0.84 | 0.35 | 0.32 | 0.15 | 0.63 |
| Area 8 - Civil engineering and Architecture | 1.39 | 2.04 | 1.47 | 2.25 | 1.05 | 0.96 | 1.95 | | | | |
| Area 8a - Architecture | | | | | | | | 5.42 | 2.62 | 3.54 | 2.09 |
| Area 8b - Civil engineering | | | | | | | | 0.18 | 0.30 | 0.39 | 0.12 |
| Area 09 - Industrial and computer engineering | | | | | | | | 0.56 | 0.51 | 0.50 | 0.81 |
| Area 10 - Antiquity, philological-literary and historical-artistic sciences | 0.64 | 0.64 | 0.61 | 0.99 | 0.86 | 0.74 | 0.93 | 0.69 | 0.71 | 0.87 | 1.12 |
| Area 11 - Historical, philosophical, psychological and pedagogical sciences | | | | | | | | | | | |
| Area 11a - Historical, philosophical and pedagogical sciences | | | | | | | | 0.49 | 0.65 | 0.58 | 0.48 |
| Area 11b - Psychology | | | | | | | | 0.40 | 0.36 | 0.33 | 0.32 |
| Area 12 - Legal sciences | | | | | | | | 3.19 | 2.43 | 3.08 | 3.16 |
| Area 13 - Economics and Statistics sciences | | | | | | | | 0.78 | 0.47 | 0.76 | 0.77 |
| Area 14 - Social and political sciences | 0.85 | 0.20 | | 0.15 | 0.29 | 0.93 | | 0.17 | | 0.13 | 0.36 |

Table 9 - GL percentages over the years

In VQR1 the annual trend is steadier for the following disciplinary Areas:

- Area01 – Computer science and Mathematics
- Area07 – Agricultural and veterinary sciences
- Area08 – Civil engineering and Architecture
- Area10 – Antiquity, philological-literary and historical-artistic sciences

For the other Areas, the annual trend is not steady because the values increase and decrease over the years.

In VQR2 the annual trend is steadier for most of the Areas.

Table 10 shows the different grey products and their annual distribution.



| Grey categories | VQR 2004-2010 | | | | | | | VQR 2011-2014 | | | |
|--|---------------|-------|-------|-------|-------|-------|--------|---------------|-------|-------|-------|
| | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |
| Architectonic project | | | | | | | | 38.00 | 24.00 | 20.00 | 18.00 |
| Bibliographic/Catalographic record, corpus | | | | | | | | 25.81 | 32.26 | 19.35 | 22.58 |
| Case notes | | | | | | | | 17.65 | 25.74 | 23.53 | 33.09 |
| Composition | 25.00 | 25.00 | | 25.00 | 25.00 | | | 18.18 | 15.91 | 25.00 | 40.91 |
| Database | | | 33.33 | | | | | 23.53 | 20.59 | 20.59 | 35.29 |
| Design | 14.29 | 14.29 | 7.14 | 28.57 | | 21.43 | 14.29 | 25.00 | 25.00 | 25.00 | 25.00 |
| Entry (in dictionary or encyclopedia) | | 25.00 | | | 75.00 | | | 23.12 | 26.01 | 26.59 | 24.28 |
| Exhibition | | 33.33 | | | | | 66.67 | 28.95 | 28.95 | 18.42 | 23.68 |
| Handwork | 9.09 | | 9.09 | | | 27.27 | 27.27 | | | | |
| Map | | | | | | | | 7.69 | 38.46 | 30.77 | 23.08 |
| Other | 11.80 | 10.91 | 10.62 | 20.94 | 15.04 | 14.16 | 22.12 | 20.18 | 20.18 | 29.36 | 30.28 |
| Patent | 5.74 | 7.38 | 7.38 | 12.30 | 14.75 | 12.30 | 24.59 | 17.02 | 21.99 | 22.70 | 38.30 |
| Performance | | | | | | | | 66.67 | 33.33 | 33.33 | 50.00 |
| Prototype of art and related project | | | | | | | 100.00 | 14.29 | 42.86 | | 28.57 |
| Software | | 10.00 | 20.00 | 20.00 | | 40.00 | 10.00 | 9.76 | 39.02 | 21.95 | 29.27 |

Table 10 - Grey products in the VQRs by year

It seems evident that the extension of the documentary categories in VQR2 supported the presence of a wider range of research products, encouraging the submission of different typologies in each disciplinary area. This influenced also the incidence of the Grey categories in VQR2, where we find products that did not appear in the previous evaluation, such as *architectonic project*, *bibliographic/catalographic record* and *corpora*, *case notes*, *map*, *performance* and *publication of unedited sources*. The calculation of the total number of products per year in each category clearly shows that in VQR1 only *patents*¹⁰ and *other* are registered every year. Significant percentages are registered for *entries* (both *in dictionary* or *encyclopedia*) in 2008, *exhibition* in 2010 and *handwork*, permanent in 2004/2006 (9.09) and 2009/2010 (27.27). The *handwork* is completely absent in VQR2.

As far as the annual trend is concerned, some categories are more stable while others are more fluctuating through time. Only the percentages referred to the products in VQR2 are not conflicting. Two peaks for the entry *performance* are evident in 2012 and 2014, as well as a much higher value of the item *concordance* in 2012.

At the same time, some of the products present in both evaluations do not show significant annual variations, with the only exception of the category *entries* (both *in dictionary* or *encyclopedia*), presenting a very high value in 2008, as well as *exhibition* and *prototype* in 2010. In addition, there is a substantial increase of the entries *patents* and *other*, although not annually.

Among the non-admissible products, we find: *educational material*, *technical reports*, *commentary*, *obituary*, *erratum*. No mention has been made to entries such as *preliminary studies*, *progress reports*, *accounts*, *search results*, *dossier*, *market researches*, *normative documents*, *feasibility studies*, etc. The *genetic studies* or the *clinical trials* are admitted to the assessment only if the author is the person who carried on the work, whereas the participation of experimenters or collaborators in the study is not taken into account. Other indications inferred through the GEVs' criteria concern the *case notes*, admitted to the evaluation only if drafted in the form of an article. The entry *working papers* is not present, although it has been declared as admissible by all GEVs, with the exception of GEV13 – Economics and Statistics¹¹. It is not clear if they may have been included in the category *papers in journals* or *papers in books*.

5. Open Science and Grey Literature... a perfect marriage

As mentioned in the introduction, Open Science is a composite idea, promoting different approaches to research and scientific communication. We may represent OS using keywords such as *network*, *data*, *collaboration* and *transparency* (Adams 2015). The focus is on the cooperation and the distribution of information through advanced technology networks (Salmi 2015). We may also observe how the main contents of the OS are tightly related to *Grey Literature* (GL). Indeed, it is evident that the focal notions of the GL, from contents to production and distribution procedures, are shared by the OS, as it differs from traditional scientific methods of knowledge

¹⁰The category *patents* is always considered as evaluable, but it may be attributed to class A or B only if internationally renowned or licensed.

¹¹ The GEV13 considers these products as designed for future publication, therefore evaluable in forthcoming assessment exercises. For the other GEVs, if the *working papers* have a ISSN code, they are considered as articles; if they have ISBN code, and a) they are open access; b) the author divests the IP rights to the working papers' series; c) every manuscript is peer-reviewed, then they are related to *monographies*.



acquisition, generation and dissemination. Moreover, if we focus our attention on the principles of OS, we will identify numerous shared contents and goals between GL and Open Access. The latter certainly faces the drawbacks of the present publication system, proposing different rationales, opposed to those currently governing the editorial market. The open and unrestricted access to the research results aims at going beyond the logic of “publish or perish” and its difficulties, such as the substantial delays in the publication, the disconnection between what is scientifically relevant and what is relevant for the career of the researchers, with specific reference to the pressures on researchers to publish (Giglia 2017). GL and OA both promote the view of knowledge as a common good, as well as the immediate, detailed and in-depth circulation of the research products in every form, not exclusively those referable to scientific articles issued by publishing companies.

At the same time, the OS supports Open Data as a mean for sharing the so-called “computational products”, i.e. protocols, procedures and/or software used by researchers to analyse and share the data along with their experimental and practical context (Candela, Manghi 2017). The Open Notebook aims at voicing documents with a low degree of visibility and availability such as research notes and laboratory tests, encouraging new forms of open peer-review based on shared protocols for the evaluation of the products in the moment they are posted online.

Therefore, the common principle of GL and OA is distributing information not conventionally disseminated, like results of failed or ambiguous experiments (Stafford 2010). Moreover, the principle of the OA is based on the advantage of obtaining and co-generating new knowledge through the interaction with citizens and local communities (Reale 2017). The terms “Open Research” and “Citizen Science” make reference to the active civic participation to the collection of data, scientific experiments, and problem solving. The involvement of the population may bring to light urgent social needs and priorities, as well as drive the attention to issues of great interest, such as environmental protection, better health, more justice, a more equitable (not necessarily equal) distribution of income (Weber et al. 2015).

A further aspect of the OS is the Open Learning and Open Education, which consists in the introduction of innovative, interactive and collaborative education policies. Their main features are the openness, the digital dimension, and the innovative spirit (Reggiani 2017). In this way, the range of available resources, contents, and products of learning materials becomes wider and more shareable.

Likewise, in the relationship between OS and GL the availability of shared tools such as repositories is of crucial importance, particularly referring to those named *next generation repositories and infrastructures*. Indeed, institutional and disciplinary repositories are commonly accepted as the most suitable sites to disseminate and store the scientific production. They traditionally host different categories of *Grey Literature*, from dissertations to sketches, from conference objects to reports, etc. The OS encourages the creation of *next generation repositories*, archives based on a series of principles and requirements including value-added services.

The new repositories would be provided with a wider range of roles and functionalities, allowing the integration in an online distributed infrastructure. This view is sponsored by associations like COAR – *Towards a global knowledge commons* that promotes the implementation of next generation repositories provided with *11 new behaviours, as well as the technologies, standards and protocols that will facilitate the development of new services on top of the collective network, including social networking, peer review, notifications, and usage assessment*¹².

Another tool shared by both the Open Science movement and the “grey” community is the use of controlled vocabularies, as *the use of controlled vocabularies for bibliographic metadata “ensures that everyone is using the same word to mean the same thing”*¹³. The results published in institutional and thematic repositories are described by bibliographic metadata. The open access and metadata exchange requires a standardized description of specific properties concerning publications and research data. The review, updating, and curation of controlled vocabularies guarantee the semantic interoperability between repositories and linked archives.

Our study highlights the need of improving new-generation metrics, as this is one of the most debated topic in the OS movement. The results of our analysis outline two complementary paths to assess the research results: one is the *peer review*, a qualitative evaluation among scientists; the other consisting in a quantitative evaluation based on the use of bibliometric indicators that counts the number of publications and the number of citations received. As previously observed,

¹² Cfr. <https://www.coar-repositories.org/activities/advocacy-leadership/working-group-next-generation-repositories/>.

¹³ Cfr. <https://www.coar-repositories.org/activities/repository-interoperability/coar-vocabularies/>.



in Italy most of the documentary typologies undergoing evaluation are represented by *papers in journals*, *papers in books* and *papers in proceedings*, i.e. products that can be easily indexed in citation databases. The most important numbers generally refer to articles published in prestigious journals with high impact factors. Indeed, traditional metrics are based on indicators like Impact Factor and H-Index, this leading to the misapplication of principles born to count the number of citations, which are actually used to evaluate the value of researchers. Moreover, these indicators are the result of complex algorithms designed by commercial operators that select and organize information following criteria not always scientifically supported, often producing incorrect data (Galimberti 2017).

Bibliometrics has not been conceived for research evaluation, but to lead librarians in purchases and to measure progresses in scientific disciplines, in spite of the extensive use of its means.

Bibliometrists agree on the assumption that the indicators are misused, as the assessment of the research impact cannot be exclusively based on the calculation of citations, yet taking into account different aspects and dimensions. At the present time, *peer review* is the only and most effective type of qualitative analysis. However, although retrospective peer-review does not present the same level of criticality of perspective peer-review¹⁴, it is an onerous and subjective procedure, not completely free from inaccuracies and not completely objective.

The analysis of new research evaluation instruments and criteria is of basic importance for the fulfilment of the OS, as it is fundamental to transform the assessment models and methods along with the science itself (Cassella 2017). The realization of OS implies that the quality of a researcher and its publications would be evaluated through other parameters, such as: accuracy, reproducibility of the results, coherence of the methods, coherence with the ethical code, openness, and participation to editorial committees (Giglia 2015). Various international initiatives promote the application of new assessment standards for the evaluation of the scientific production¹⁵. The OS movement examines a series of alternative metrics to monitor the development of the scientific system and to measure both individual and group work. Among them we find the *usage metrics*, which are based on the number of views and downloads of a product. The usage measurement differs from the citations measurement since it involves a broader range of users who are potentially interested in reading and downloading articles and data. The metrics based on the usage appraise the interest and the degree of absorption of a work and may be quite relevant for the Open Science, not only for the use made of the publications, but also for the monitoring of non-traditional publications (e.g. posts, blogs) and for the reuse of open data and open software.

The *altmetrics* represents a further sub-system of new generation metrics mostly based on social media applications. More in detail, the *altmetrics* make reference to downloads, blog posts, social media interactions, citations, comments, tweets, opinions expressed by the users through means such as *likes* on Facebook. The *altmetrics* may contribute to the evaluation of the impact of a study. Indeed, nowadays researchers are increasingly exploiting the web in their studies; therefore, discussions among experts have shifted from laboratories' hallways to blogs and social networks, as well as the "raw science" (datasets, code, and experimental designs) finds place in blogging, microblogging, and annotations available and shareable online¹⁶.

The advantages of the use of *altmetrics* are quite clear: citations can be retrieved faster; authoritative, but not frequently cited works are not disregarded; the operating environment of

¹⁴ Among the most debated aspects, we find the fact that the reviewer is not always more competent than the author of the paper reviewed, in addition to potential conflicts of interest, the lack of accuracy, long times for the review process, implying consistent delays in the publication.

¹⁵ The *DORA Declaration* provides recommendations to funding agencies, institutions, publishers, organizations supplying metrics, and to researchers; it underlines the need of avoiding the use of indicators like the IF, instead considering the intrinsic value rather than the journal, and taking vantage of the opportunities provided by new digital indicators. The *Leiden Manifesto* contains 10 principles indicating possible solutions to the issues created by the inappropriate introduction of Bibliometrics, and suggests the usage of valid statistics along with a correct evaluation of the objectives and the nature of the research assessed. The *Science in Transition Position paper* highlights the continuous growth of exchanges in scientific information out of the traditional channels and documentation (e.g. journals and books), preferring more informal, fast and open modalities and self-production systems like blogs and microblogs. *The Metric Tide* examines the role of metrics in the evaluation and management of the research in the British system, emphasizing the limits of the quantitative measures and indicating *peer review* as the only possible yardstick. The report lists a series of recommendations for the design of metrics based on *robustness, humility, transparency, diversity, and reflexivity*.

¹⁶ *Altmetrics: a manifesto*, <http://altmetrics.org/manifesto/>.



the researcher is considered properly; different metrics for the evaluation of a study may be aggregated.

Although *altmetrics* are still in their experimental stage, the crisis of the traditional evaluation systems is so serious that the use of alternative indicators is expected to be promoted and improved in a very near future (Schöpfel, Prost 2017).

6. Conclusions

The Italian Research Assessment Processes do not completely exclude *Grey Literature*. However, they are almost exclusively based on the analysis of commercially distributed products.

This is due to:

- the non-eligibility of some research products (e.g. *educational material, technical reports, commentary, obituary, erratum...*);
- the lack of interest in items such as *preliminary studies, progress reports, accounts, search results, dossier, market researches, normative documents, feasibility studies, etc.*;
- the disadvantage in submitting scientific products other than articles in journals;
- the impact of the evaluation criteria on researchers leads to the philosophy of *Public or Perish*: the researchers publish only scientific articles in prestigious journals.

The combination of principles and tools of the OS may become a primary channel of dissemination for the GL. On the other hand, GL may evolve into a primary source for the OS. GL has a long-standing tradition, it is a mosaic of different documentary typologies including various areas of interest: from documentary research, to a wide range of materials produced by local, national and international private or public institutions, industry associations, foundations, private individuals, etc. Both GL and OS meet the need of faster, more efficient, economical, focused dissemination channels, insisting on the urgency of making available all the documentary forms currently not circulated and inaccessible. These documentary and procedural demands may be fulfilled by the tools of the OS. The new research scenarios offer considerable opportunities for the collection, description, identification, and dissemination of literature and data. Any kind of product may be identifiable and accessible using tools like repositories or new generation infrastructures, which supports all the components of the research activity: objects, people, technology, procedures (Candela, Manghi 2017). The use of these tools opens the status of scientific product to a wide range of documents: drafts, software, pictures, diagrams, tables, experimental protocols, then creating communication patterns of major interest either to those who would actively collaborate with researchers in their studies, or to those who would simply collect information or gain knowledge. GL, currently left out from traditional metrics, may be involved in the application of new typologies of metrics including aspects of scientific products usually labelled as *Grey Literature*. The *altmetrics* cross the borders of traditional *research results*, limited to databases like WoS and Scopus, taking into account non-traditional and non-commercial products. Exploiting such tools, the assessment exercises may finally turn into transparent, comprehensible, and shared processes.

The logic that moves the current scientific communication implies the risk of producing *fashionable* research rather than *quality* research. The concept of OS includes the necessary human skills, resources, standards, best practices and technical infrastructures necessary to realize an innovative *entire research enterprise*.

In this new ecosystem of the scientific communication *Grey Literature* might find its ideal collocation, but it is necessary that scientific institutions and politics exchange experiences and build networks across national borders in order to realize this new system, then allowing the growth of a new dialogue between science and society.

The developments in technology and the opportunities offered by the semantic web may have not led to the advancements expected about ten years ago. However, it has been widely demonstrated that the technologies available are ready to support substantial and complex goals. Cultural, political and economic changes are necessary in order to realize the Open Science. Europe plays a key role in supporting greater openness and in redefining the research processes.

**Bibliography¹⁷**

- 1) Adams J. (2015). *Impact of Open Science methods and practices on the economics of research and science. Case Studies from Life, Mathematical and Social Sciences*. European Commission, 2015.
- 2) Aliprandi S. (ed.) (2017). *Fare Open Access. La libera diffusione del sapere scientifico nell'era digitale*. LeEdizioni (ebook), <https://aliprandi.org/books/fare-openaccess/>.
- 3) *Altmetrics Status Quo*. OpenUP - OPENING UP new methods, indicators and tools for peer review, impact measurement and dissemination of research results, Deliverable D5.1. European Commission, 2016.
- 4) ANVUR – Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca, <http://www.anvur.org/>.
- 5) Baccini A. (2011). *Valutare la ricerca scientifica. Uso e abuso degli indicatori bibliometrici*. Il Mulino (e-book).
- 6) Biagetti M.T. (2017). Validità e limiti della *library catalog analysis* per la valutazione della ricerca nelle scienze umane e sociali. *AIB studi*, 57 (1)(gennaio/aprile 2017), 7-22.
- 7) Bianco L. (2010). *Riflessioni sulla valutazione della ricerca e l'istituzione dell'ANVUR*, <http://matematica.unibocconi.it/articoli/riflessioni-sulla-valutazione-della-ricerca-e-istituzione-dellanvur>.
- 8) Candela L., Manghi P. (2017). La comunicazione scientifica i tempi della Open Science. *Paradoxa*, XI (1) 69-93.
- 9) Caso R. (ed.) (2009). *Atti del Convegno Pubblicazioni Scientifiche, Diritti d'autore e Open Access*. Università degli Studi di Trento.
- 10) Cassella M. (2013). Dati aperti e ricerca scientifica: aspetti gestionali e normativi nel contesto dell'e-science. *AIB studi*, 53 (3) 223-238.
- 11) Cassella M. (2014). Bibliometria sì, bibliometria no: la valutazione della ricerca nelle scienze umane e sociali al bivio. *AIB studi*, 54 (2/3) 295-304.
- 12) Cassella M. (2017). La valutazione alternativa: *altmetrics* e dintorni. *AIB studi*, 57 (1) 79-90.
- 13) Castellucci P. (2017). Formiche virtuali o virtuose? Verso un'etica dell'accesso. *AIB studi*, 57 (1) 51-62.
- 14) Ćirković S. (2017). Transition to Open Access and its Implications on Grey Literature Resources. *Proceedings of the Eighteen International Conference on Grey Literature (GL18)*, 81-90. TextRelease, Amsterdam.
- 15) CRUI – Commissione Biblioteche, Gruppo Open Access. *L'Open Access e la valutazione dei prodotti della ricerca scientifica. Raccomandazioni*. Roma, aprile 2009, <https://www.cru.it/oa.html>.
- 16) De Robbio A. (2007). Analisi citazionale e indicatori bibliometrici nel modello Open Access. *Bollettino AIB*, 3, 257-288.
- 17) Delfanti A. (2008). Collaborative Web between open and closed science. *Journal of Science Communication*, 7 (2), <https://jcom.sissa.it/archive/07/02/!com0702%282008%29C01>.
- 18) European Commission. *Next-generation metrics: Responsible metrics and evaluation for open science*. Report of the European Commission Expert Group on Altmetrics, <https://ec.europa.eu/research/openscience/pdf/report.pdf>.
- 19) Faggiolani C., Solimine G. (2012). La valutazione della ricerca, la bibliometria e l'albero di Bertoldo. *AIB Studi*, 52 (1) 57-63.
- 20) FORCE11 – The Future of Research Communications and e-Scholarship (2012). *Improving Future Research Communication and e-Scholarship. White Paper*. February 19, 2012. <https://www.force11.org/sites/default/files/files/Force11Manifesto20120219.pdf>
- 21) *The future of scholarly scientific communication*. Report, <https://royalsociety.org/science-events-and-lectures/2015/04/future-of-scholarly-scientific-communication-part-1/>.
- 22) Galimberti P. (2012). Qualità e quantità: stato dell'arte della valutazione della ricerca nelle scienze umane in Italia. *JLIS.it*, 3 (1), <https://www.jlis.it/issue/view/326>.
- 23) Galimberti P. (2017). Open Science: altmetrics, impatto e controllo della qualità. *Paradoxa*, XI (1) 29-38.
- 24) Galimberti P. (2017). Open access, open science. L'Italia, un paese in grave ritardo. *Intervento sul sito ROARS – Return On Academic Research*, <https://www.roars.it/online/?p=57268>.
- 25) *Interim Report on Researcher Impact*. OPENING UP new methods, indicators and tools for peer review, impact measurement and dissemination of research results – OpenUp Deliverable D5.3. European Commission, 2017.
- 26) Giglia, E. (2009). Più citazioni in Open Access? Panorama della letteratura con uno studio sull'Impact Factor delle riviste Open Access. *CIBER 1999-2009*, 125-145. Ledizioni, Milano.
- 27) Gioè L. (2016). Open Science. Per una democrazia della conoscenza. *Scienza in rete*, <http://www.scienzainrete.it/contenuto/partner/open-science-democrazia-della-conoscenza/marzo-2016>.
- 28) Guerrini M. (2009). Nuovi strumenti per la valutazione della ricerca scientifica. Il movimento dell'open access e gli archivi istituzionali. *Biblioteche oggi*, 27 (8) 7-17.
- 29) Leiden Manifesto for research metrics, <https://www.roars.it/online/leiden-manifesto-for-research-metrics/>.
- 30) The Metric Tide, <http://www.hefce.ac.uk/pubs/rereports/year/2015/metrictide/>.
- 31) Moed H.F. (2005). *Citation Analysis in Research Evaluation*. Springer, 2005.
- 32) Moed H.F. (2011). The multi-dimensional research assessment matrix. *Research Trends*, 23 (May 2011), <https://www.researchtrends.com/issue23-may-2011/the-multi-dimensional-research-assessment-matrix/>.
- 33) Pozzo R. (ed.) (2017). Scienziati, giù dalla torre d'avorio! *Paradoxa*, XI (1).
- 34) Reale E. (2017). Promuovere una scienza aperta: risorse, incentivi e valutazione. *Paradoxa*, XI (1) 15-27.
- 35) Rebora G., Turri M. (2013) - *The UK and Italian research assessment exercises face to face*. *Research Policies*, 42 (9) 1657– 1666.
- 36) ICSU – International Council for Science (2014). *Open access to scientific data and literature and the assessment of research by metrics*, <http://www.roars.it/online/wp-content/uploads/2014/09/ICSU-Report-on-Open-Access.pdf>.
- 37) Reggiani L. (2017). In vista d'alti cieli: l'Open Education tra conoscenza scientifica e società democratica. *Paradoxa*, XI (1) 95-106.

¹⁷ URL last access: December 2017.



- 38) ROARS – Return On Academic Research, <https://www.roars.it/>.
- 39) Rubele R. (2012). Appunti per una storia dell'ANVUR (I). *Intervento sul sito ROARS – Return On Academic Research*, <https://www.roars.it/online/appunti-per-una-storia-dellanvur-i/>.
- 40) Salmi J. (2015). *Study on Open Science. Impact, Implications and Policy Options*. European Commission, 2015.
- 41) San Francisco Declaration on Research Assessment, <http://www.ascb.org/dora/>.
- 42) Serini P. (2003). Attualità della letteratura grigia: il ruolo delle biblioteche nella sua valorizzazione. *Biblioteche oggi*, 21 (1) 61-72.
- 43) Stafford N. (2010). Science in the digital age. *Nature*, 467 (7317) S19-S21.
- 44) Scholarly Publishing and Academic Resources Coalition Advocating change in scholarly communications for the benefit of researchers and society. *Better ways to evaluate research and researchers A SPARC Europe Briefing Paper*, <https://sparceurope.org/new-sparc-europe-briefing-paper-better-ways-evaluate-research-researchers/>.
- 45) Schöpfel J., Prost H. (2017). Altmetrics and Grey Literature: Perspectives and Challenges. *The Grey Journal*, 13 (1) 5-11.
- 46) Turbanti S. (2016). La visibilità – e l'impatto? – nel Web ai tempi dei social: i principali strumenti di altmetrics. *AIB studi*, 56 (1) 41-58.
- 47) Weber M., André D., Llerena P. (2015). *A new role for EU Research and Innovation in the benefit of citizens: towards an open and transformative R&I policy. Policy Paper by the Research, Innovation, and Science Policy Experts (RISE)*. European Commission, 2015.
- 48) <https://opensource.com/resources/open-science>.
- 49) <http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/open-science-movement/>.
- 50) http://www.anvur.org/index.php?option=com_content&view=article&id=122&Itemid=305&lang=it.
- 51) http://www.anvur.org/index.php?option=com_content&view=article&id=32&Itemid=372&lang=it.
- 52) <http://www.anvur.org/rapporto/>.
- 53) http://www.anvur.org/index.php?option=com_content&view=article&id=825&Itemid=599&lang=it.
- 54) http://www.anvur.org/index.php?option=com_content&view=article&id=841&Itemid=601&lang=it.
- 55) <http://www.anvur.org/rapporto-2016/>.
- 56) <https://sites.google.com/site/scienzaapertaricercamigliore/programma>.
- 57) <http://www.anvur.org/rapporto/> (VQR1) and <http://www.anvur.org/rapporto-2016/> (VQR2).