

Curators of scientific publications

Anastásios Perdicóulis

Assistant Professor, ECT, UTAD (<http://www.tasso.utad.pt>)

Senior Researcher, CITTA, FEUP (<http://www.fe.up.pt/~tasso>)

Visiting Researcher, Oxford Institute for Sustainable Development, OBU, UK

Abstract

The vast volume of scientific publications, yield of an ever-accelerating production process, creates requirements for appropriate information management. The attribution of curators, in turn, raises questions of utility, responsibility, and power.

1 Introduction

The proliferation of scholarly works — i.e. articles, journals, and books that involve or relate to serious academic study — is generally welcome and hailed as an indicator of scientific progress. From the point of view of researchers, though, having to explore and keep up with the ever-increasing volume of scientific publications (Munroe, 2013) is no less than a daunting task.

The global pool of scientific publications is useful to researchers to the extent that they can *find* the information they need, at the *quality* they need (e.g. resolution, objectivity), they are able to *buy* it (as most information carries a price tag while it is useful), and have the time to *read* it within the confines of their research projects — otherwise, abundance is no better than scarcity (Figure 1).

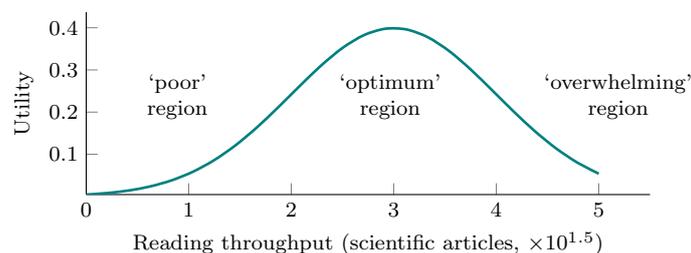


FIGURE 1 Abundance is bounded by human or project limitations; adapted from Perdicóulis (2012a)

A global and un-biased management of scientific publications (e.g. organisation and retrieval) could facilitate research, but *who* could *safely* be the curator of scientific publications? The function deserves scientific objectivity and transparency (Perdicóulis, 2012b, 2013a), and is a public service rather than a money-making opportunity (Perdicóulis, 2014a).

2 Sources and pointers

2.1 Publishers

Whether independent or corporations, publishers are the *par excellence* sources of scholarly publications. Their databases often have online search and retrieval interfaces, which range from basic to very advanced — for instance, [ScienceDirect \(website\)](#), which is service of [Elsevier \(website\)](#), provides a powerful search facility for the contents of the publisher.

2.2 Authors

Authors of scholarly publications usually aggregate the bibliographical references of their work — and of their team members — in classic print forms such as the *curriculum vitae*. Recent technological advances facilitate, or even stimulate online versions of that information, including search interfaces, as well as digital copies of the work itself. However, due to copyright issues with the publishers, author databases usually feature either (a) pre-final versions of their work, or (b) mere pointers (i.e. hyperlinks) to the publisher websites. At any rate, authors are disperse sources, or relatively small centres of information.

2.3 Institutions

Academic departments and/ or research centres may maintain their own collections — often referred to as ‘repositories’ — with the publications of their employees or affiliates. Such collections may contain (a) physical or electronic copies of the work, and/ or (b) pointers (i.e. hyperlinks) to the actual work at the publisher or author websites. For the same reasons as with the author collections, institutional collections usually feature ‘pre-final’ versions of their work, and are also disperse sources — albeit more aggregated than authors.

3 Registers

3.1 National libraries

National libraries — e.g. the [Library of Congress \(website\)](#) in the USA, the [Bodleian Libraries \(website\)](#) in the UK, and the [Bibliothèque Nationale \(website\)](#) in France — function as central repositories for all new publications produced in their territory, and maintain their own physical collections as well as catalogues or databases. Although it is possible to visit these libraries and obtain access to the publications, it is not practicable for *all* researchers to be doing so — especially those who live far. To this extent, some libraries employ services such as [EBSCO \(website\)](#) or [ProQuest \(website\)](#) to provide external access to their bibliographic database ([Library of Congress, website](#)).

3.2 ISO-based registers

For organisation purposes, a private business initiative in the 1960s initiated a ‘standardised’ registration of books ([Bowker, website](#)), which later became the *International Standard Book Number* ISO norm ([ISBN, website](#), ISO 2108:2005). Although international in scope, ISBN

registration¹ is operated by national agencies, which can be public or private entities — for instance, the ISBN agency in the USA is operated by [Bowker \(website\)](#), which is an affiliated business of [ProQuest \(website\)](#), while in Portugal it is operated by the Portuguese association of publishers and booksellers ([APEL, website](#)).

In the same manner, ‘open’ volumes such as journals and websites can be registered for an *International Standard Serial Number*, which follows a different ISO norm ([ISSN, website](#), ISO 3297:2007). Many ISSN agencies are national, although countries with relatively low production of journals are served by the international centre.

In more recent years, individual articles² can be registered for a *Digital Object Identifier* (DOI), which is subject to a specific ISO norm ([International DOI Foundation, website](#), ISO 26324:2012). In the same manner as ISBN and ISSN, there are specific registration agencies (RA) responsible for issuing DOI numbers ([International DOI Foundation, website](#)).

4 Services

4.1 Abstracting and indexing

While librarians can be credited with the first cataloguing of publications ([Perdicoulis, 2014b](#)), their mission is nowadays overshadowed by a number of private companies that offer abstracting and indexing services — for instance, [Scopus \(website\)](#), which is a registered trademark of [Elsevier \(website\)](#), as well as [EBSCO \(website\)](#) and [ProQuest \(website\)](#) maintain publication databases, complete with abstracts, keywords, and full bibliographic information, as well as associated search facilities.

The indexing of scholarly publications carried out as a specialisation of popular internet search engines — for instance, as done by [Google Scholar \(website\)](#) — provides a facility appreciated by researchers: the convenient handling of bibliographic information, commonly known as ‘metadata’, such as information about the author, publisher, and copyright of the publications. By using publication metadata with reference management software such as [Zotero \(website\)](#) or $\text{BIB}\text{T}\text{E}\text{X}$, researchers can maintain their own indexed bibliographies.

Although the abstracting and indexing services facilitate the search for scientific publications, most databases are likely to be *partial*, either (a) due to the commercial interests of their holding companies, or (b) because they do not have the resources (and often neither the scope nor ambition) to be as extensive as the registers (§ 2).

4.2 Rating and ranking

The early attempt of Eugene Garfield — the founder of the *Institute for Scientific Information* (ISI) — to trace the relations between scientific publications ([Perdicoulis, 2013a](#)) is continued by the [Thomson Reuters Corporation \(website\)](#), and has brought about new contenders such as [CrossRef \(website\)](#). In the domain of scientific research, Thomson is currently popular for (a) the wider *Web of Knowledge* (WoK) electronic platform, and (b) the *Web of Science* (WoS), a subscription-based scientific citation indexing service. Such services use citations not only for pattern analyses, but also

¹A single publication obtains a different registration number for each medium in which it is distributed — e.g. hardback print, paperback print, digital on CD-ROM, or digital online; hence, there are more ISBNs than books.

²DOI can also apply to books, journals, and other publication media — therefore, there may exist multiple registration numbers for certain ‘objects’.

for the construction of complex *indices* of popularity such as the ‘impact factor’, and subsequent rating (e.g. 1.657) and ranking (e.g. top, ‘highly cited’) of scientific publications. Thomson is also popular for its *citation indexes* (Example 4.1), created by selecting scholarly journals through a number of assessments (Example 4.2).

EXAMPLE 4.1 WOS CITATION INDEXES (TESTA, 2006, 2008)

- Science Citation Index/ Science Citation Index Expanded
- Social Sciences Citation Index
- Arts and Humanities Citation Index
- Index Chemicus
- Current Chemical Reactions

EXAMPLE 4.2 WOS ASSESSMENTS AND CRITERIA (TESTA, 2006, 2008)

Publishing standards:

- timeliness — according to stated publication schedules
- international editorial conventions — e.g. informative/ descriptive article titles, full author addresses, complete bibliographic information
- English-language bibliographic information
- peer review

Editorial content: must meet Thomson’s customer needs

International diversity: representation of many countries

Citation analysis: based on cited references; includes the ‘impact factor’

5 Discussion

5.1 Reliance (sources and registers vs. services)

Since they do not purport to be assessing the scientific publications they include, sources and pointers (§ 2) as well as registers (§ 3) create an *ambivalence*: they may include either (a) *bona fide* scholarly publications, regardless of their popularity, or (b) material such as ‘vanity’ publications, replications of already published material, or works produced under ‘predatory’ conditions — i.e. when publishers solicit articles from authors for a fee (Perdicoulis, 2014a). In most cases repositories and registers are compiled with care and caution by trained professionals (usually librarians) who hardly see themselves as censors. Therefore, it is the researchers who must make the distinction with their own capacity and criteria (Perdicoulis, 2013a).

On the other hand, services (§ 4) make a more explicit selection of publications (Example 4.2). However, as it follows from Example 4.2, a number of *bona fide* scientific journals can be excluded from these collections and ranking lists for reasons that are not related to quality — for instance: (a) journals with no periodicity; (b) journals of national scope; (c) journals with bibliographic information not in English; (d) journals with low reading or citation rates. On the other hand, a number of ‘predatory’ publications may be added if they just fulfil the admission criteria — or, it could be argued, could even see an opportunity to be created, since there is a market with many authors pressured to publish and everyone is after popularity ratings. Contrary to the ‘wild territory’ of the sources and registers, in the case of the services there is a false confidence which

justifies issuing a *caveat emptor*: a system that appears to be verified with journal selection and ranking procedures may be concealing useless or even false publications, while many good ones are still unaccounted for.

5.2 Decision (abstracting and indexing vs. rating and ranking)

Through keywords and summaries, abstracting and indexing services (§ 4) help the researchers locate the articles they need — although this often comes at a price. Hence, researchers in principle select the most relevant and apparently trustworthy publications, read, use them in their work, and cite them. Considering the complementary track of research, regarding the selection of an appropriate journal to submit their work for publication, authors would decide by subject, quality, or even what they might see as the journal’s ‘image’ — Figure 2(a).

On the other hand, rating and ranking services offering journal or article popularity statistics influence the decision-making of the readers: (a) as researchers, which publications to consult, use, and cite; (b) as authors, in which journals to submit their work for obtaining more projection themselves — Figure 2(b). Hence, rating and ranking services have an active role in the circuit of scientific publications: they actually strengthen the stronger journals and eliminate the weaker ones — including small-scale journals or independent publications.

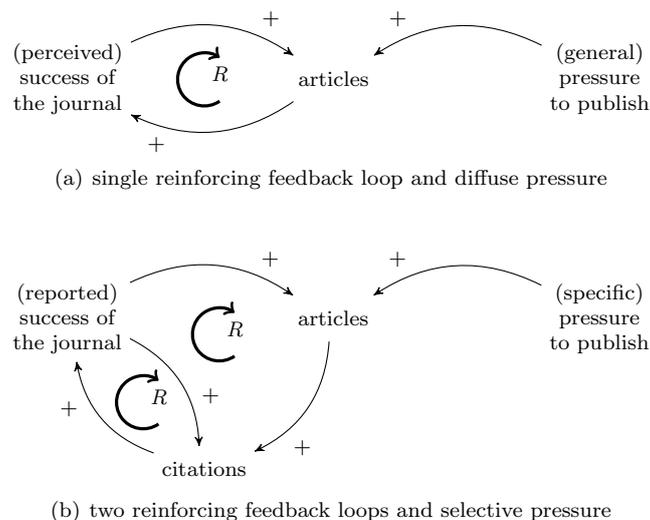


FIGURE 2 The calculated and announced popularity (or unpopularity) of a journal catapults it either to stardom or to termination

6 Challenges

Curators of scientific publications can facilitate the work of researchers in different ways, but their participation is not meant to be intrusive. Exempt curators such as the ISO-based registers and national libraries are perhaps the best choice of researchers for being informed of ‘what exists’ — i.e. the big picture, without comments —, but researchers must form their own opinions and decide for themselves, which requires proper education and training (Perdicoulis, 2013a).

For intricate motives such as business interests (Perdicóúlis, 2014a, 2012c) and the need for reassurance or ambition of researchers and institutions (Perdicóúlis, 2013a), the curators offering rating and ranking services appear particularly appealing. Rating and ranking agencies thus gain considerable power to dictate what is to be read and cited — and, by exclusion, what is not. This tendency raises a number of serious concerns such as ‘is scientific research carried out in democracy?’ or ‘to what extent are researchers allowed to decide for themselves?’

If indeed scientific publications are all about popularity, then marketing and lobbying may provide interesting boost effects at the ‘strategic’ level — but where does this take the *ethics* of research? In the meantime, the race for accumulating citations entices researchers to ‘peccadillos’ such as replicating publications (e.g. citations in Examples 4.1 and 4.2). In such a context, the conscientious researcher must be in a healthy state of ‘puzzled’.

References

APEL — Associação Portuguesa de Editores e Livreros (website) <http://www.apel.pt>

Bibliothèque Nationale de France (website) <http://www.bnf.fr>

Bodleian Libraries (website) <http://www.bodleian.ox.ac.uk>

Bowker — US ISBN agency (website) <http://www.isbn.org>

CrossRef (website) <http://www.crossref.org>

EBSCO Information Services (website) <https://search.ebscohost.com>

Elsevier (website) <http://www.elsevier.com>

Google Scholar (website) <http://scholar.google.com>

International DOI Foundation (website) <http://www.doi.org>

ISBN — International (website) <https://www.isbn-international.org>

ISSN — International Centre (website) <http://www.issn.org>

Library of Congress (website) <http://www.loc.gov>

Munroe, R. (2013) How much science is there? *Science*, **342**:58.

Perdicóúlis, A. (2014b) *Cataloguing*. Systems Planning Guidance Folio.

Perdicóúlis, A. (2014a) Sharing the science. *oestros*, **16**

Perdicóúlis, A. (2013b) The scientific qualifier. *oestros*, **11**

Perdicóúlis, A. (2013a) Educated readership. *oestros*, **7**

Perdicóúlis, A. (2012c) System to strategy. *Systems Planner*, **11**

Perdicóúlis, A. (2012b) Scientific writing. *oestros*, **5**

Perdicóúlis, A. (2012a) Information and understanding. *oestros*, **2**

ProQuest (website) <http://www.proquest.com>

ScienceDirect (website) <http://www.sciencedirect.com>

Scopus (website) <http://www.scopus.com>

Testa, J. (2008) The Thomson Scientific journal selection process. *Contributions to Science*, **4**(1):69–73.

Testa, J. (2006) The Thomson Scientific journal selection process. *International Microbiology*, **9**:135–138.

Thomson Reuters Corporation (website) <http://thomsonreuters.com>

Zotero (website) <https://www.zotero.org>

