
RESEARCHER PERCEPTIONS AND PRACTICES IN OPEN ACCESS PUBLICATION IN TWO ENGINEERING POSTGRADUATE PROGRAMS

AS PERCEPÇÕES E PRÁTICAS DE PUBLICAÇÃO EM ACESSO ABERTO DOS
PESQUISADORES DE DOIS PROGRAMAS DE PÓS-GRADUAÇÃO EM ENGENHARIA

PERCEPCIONES Y PRÁCTICAS PARA SU PUBLICACIÓN EN ACCESO ABIERTO DE
INVESTIGADORES DOS PROGRAMAS DE POSGRADO EN INGENIERÍA

¹Ariadne Chloe Furnival, ¹Daniel Andre Rigo Guirra

¹ Universidade Federal de São Carlos

Correspondência

¹ Ariadne Chloe Furnival
Universidade Federal de São Carlos
São Carlos, SP
Email: chloefscar@gmail.com
ORCID: <http://orcid.org/0000-0002-2344-4400>

Submitted: 26-08-2016

Accepted: 26-02-2017

Published: 12-04-2017



JITA: IM. Open data

RESUMO: Apresenta os resultados de um estudo que buscou conhecer as práticas e percepções de pesquisadores credenciados em dois programas de pós-graduação em engenharia, avaliados com nota sete pela CAPES, em relação à publicação em canais de acesso aberto comparado com aqueles restritos, de assinatura. Foi usada uma abordagem metodológica mista. Dados de publicação referentes ao período 2008 a 2016 foram levantados e tabulados dos currículos Lattes dos pesquisadores usando os softwares *ScriptLattes* e *VantagePoint* respectivamente. Os 244 títulos de revistas científicas e os 970 artigos identificados dessa fase foram então pesquisados em outras plataformas como o SciELO, DOAJ, SHERPA/ROMEO, Google Acadêmico e o próprio Repositório Institucional da instituição dos programas de pós-graduação para verificar a proporção de artigos publicados disponíveis em acesso aberto. Concomitantemente a esses levantamentos quantitativos, foram realizadas entrevistas semiestruturadas com 11 pesquisadores dos dois programas para identificar suas percepções em relação à publicação em geral e especificamente em acesso aberto, e em repositórios abertos. Dos dados quantitativos levantados, foi verificado que apenas 21,9% dos artigos publicados do primeiro programa de pós e 29,8% do segundo estão disponíveis em acesso aberto. No entanto, nas entrevistas a maioria dos autores apoia o acesso aberto, embora manifestem pouco conhecimento sobre o acesso aberto em repositórios, incluindo aquele da sua própria instituição, como também desconhecem o caminho complementar de se publicar primeiramente em revistas de alto impacto com o posterior arquivamento do mesmo artigo no repositório, ao término do eventual tempo de embargo da editora.

PALAVRAS-CHAVE: Acesso aberto. Comunicação científica. Práticas de publicação. Percepções de pesquisadores.

ABSTRACT: The article presents the results and discussion of a study that aimed to identify the practices and perceptions of researchers concerning making their publications available in open access, as compared to publishing in subscription, paywalled journals. The study subjects were researchers accredited in two postgraduate Engineering programs that have obtained the maximum evaluation by CAPES. A mixed methodological approach was used. Quantitative publication data for the period 2008 to 2016 were collected and tabulated from the researchers' Lattes curricula using the *ScriptLattes* and *VantagePoint* software respectively. The 244 periodicals and 970 published articles identified in this phase were then researched in other platforms such as SciELO, DOAJ, SHERPA/ ROMEO, Google Scholar and in the postgraduate programs' Institutional Repository itself to verify the proportion of published articles available in open access. Concomitantly, semi-structured interviews were conducted with 11 researchers of the two postgraduate programs to identify their perceptions regarding publication in general and specifically, in open access and self-archiving in open repositories. From the quantitative data collected, it was found that only 21.9% of the articles published in the first postgraduate program and 29.8% of the second are available in open access. However, in the interviews, most authors support open access while at the same time demonstrating little knowledge of open access repositories, including that of their own institution. They are also unaware of the complementary possibility of first publishing in high-impact journals, and subsequently depositing the same article in the open access repository on the completion of an eventual publisher embargo period.

KEYWORDS: Open access. Scientific communication. Publishing practices. Researcher perceptions.

RESUMEN: Presenta los resultados de un estudio que tuvo como objetivo conocer las prácticas y percepciones de los investigadores acreditados en dos programas posgrado en ingeniería evaluados por la CAPES nota 7, en relación con la publicación en los canales de acceso abierto en comparación con aquellos restringida, suscripción. Se utilizó un enfoque metodológico mixto. Datos de publicación para el período de 2008 a 2016, fueron reunidos y tabulados de los curriculum Lattes de los investigadores, usando-se los softwares *ScriptLattes* y *VantagePoint* respectivamente. Con los 244 títulos de revistas y 970 artículos identificados en esta fase, fueron luego buscados en otras plataformas como SciELO, DOAJ, SHERPA / ROMEO, Google Scholar y en el Repositorio Institucional de la institución de los programas de postgrado en sí, para verificar la proporción de artículos publicados disponibles en acceso abierto. Simultáneamente a estas encuestas, entrevistas semi-estructuradas se realizaron con 11 investigadores de los dos programas para identificar sus percepciones con respecto a la publicación en general y específicamente, con el acceso abierto y repositorios abiertos. De los datos cuantitativos planteados, se encontró que sólo el 21,9% de los artículos publicados en el primer programa y el 29,8% del segundo están disponibles en acceso abierto. Sin embargo, en las entrevistas, la mayoría de los autores es de acuerdo con el acceso abierto antes de que se muestran poco conocimiento de la vía verde, o repositorios abiertos, incluyendo el de su propia institución, pero tampoco conscientes de la manera que se complementen publicar primero en revistas de alto impacto con la presentación posterior del mismo artículo en el repositorio, al final de cualquier tiempo de embargo ejecutado por la editora.

PALABRAS CLAVE: Acceso abierto. Comunicación científica. Práticas de publicación. Percepciones de investigadores.

1 INTRODUCTION

In recent years, there has been the proliferation and consolidation of initiatives and practices to strengthen and expand open access (OA) to the scientific literature in Brazil. Adhering to the renowned Budapest Open Access Initiative's (BOAI) definition from 2002, open access to the literature refers to:

[...] its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited (BOAI, 2002).

Today, the OA landscape in Brazil is made up of 92 digital repositories (OpenDOAR, 2017) and 944 OA scientific journals registered in the respected Directory of Open Access Journals (DOAJ, 2017). In the OA literature, these two forms of OA – repositories and journals – are called, respectively, “green OA” and “gold OA” (SUBER, 2012).

It is highly likely that the outstanding performance of gold OA in Brazil is due to the existence, since 1997, of SciELO (Scientific Electronic Library Online), a software platform which curates collections of journals published in eleven South American and Caribbean countries, as well as some journals from South Africa, Portugal and Spain. SciELO is not merely a portal, platform or repository, but a *modus operandi* for publishing and scientific dissemination that maintains quality standards through rigorous criteria for inclusion and retention of journal titles. This has resulted in the inclusion of many of its titles in important international databases like Web of Science, PubMed, SCOPUS, CROSSREF, AGRIS, DOAJ (PACKER, 2009). Archambault et al. (2013) attribute the success of SciELO to the fact that 63% of articles published by Brazilian authors are available in OA, so that, according to these authors, Brazil has already passed the “tipping point” (ARCHAMBAULT et al., 2013, p. 2) in terms of availability of its production in OA, meaning that more than 50% of published papers by researchers from Brazil are freely available in OA.

It has been observed in surveys collecting scientists' opinions and perceptions of OA, that for the majority of publishing scientists, the expression "to make publications available in open access" connotes, first and foremost, publishing articles in OA journals. That is, the use of digital repositories to achieve open access is usually disregarded and quite often, unknown to scientists. In addition (and a point of concern for those who advocate for more initiatives to extend open access), OA journals are frequently seen by scientists as scientific journals that adopt the bad practice of sending out unsolicited emails, inviting scientists to submit their

manuscripts for publication in the journal in question, which may occur after an extremely rapid and therefore dubious peer review process. This practice has become so routine and ill-regarded that there is even a register, prepared by Jeffery Beall, a librarian from the University of Colorado, of these journals which he calls "predatory" (BUTLER et al., 2013). Of course, in order to publish, authors will have to pay a fee to cover the article processing charges (APCs), but it should be noted that this business model is also adopted by good quality and respected OA journals (and so registered in the abovementioned DOAJ), as well as the "hybrid" journals. The latter are journals that, in the same issue of a given volume, publish some articles in OA, and others in non-OA, with the authors of the former having opted for immediate publication in open access, while having to pay "article processing charges" (APCs) for this choice.

Scientists are renowned for aspiring to publish in journals that have the highest possible Impact Factor (IF), given its supposed potential to boost citations. The Qualis CAPES system maintained by the Brazilian government's Coordination for the Improvement of Higher Education Personnel (CAPES), also takes the IF and prestige of the scientific journals into consideration when designating the strata A1 (considered the most significant weight) to C (zero weight). Given that the intellectual output of scientists who teach and supervise on postgraduate programs in Brazil is monitored by CAPES for the ranking of the postgraduate programs, it is natural that these scientists choose to publish in journals accredited by Qualis CAPES, amongst which are both non-OA subscription-access journals as well as OA journals.

In view of the above, the research described here had as its main aim the mapping of publishing practices and the collection of the perceptions of a group of scientists regarding publication in open access compared to non-open access journals. These scientists are currently accredited in one of the two Postgraduate programs (PGPs) in Engineering at the University of São Paulo (USP) that have attained the highest ranking (7) in the excellence assessment of postgraduate programs carried out at four-yearly intervals by CAPES. In order to characterize the publishing practices of these scientists, data was collected, covering the five-year period (2008-2016), to identify the quantity of articles published in non-OA (subscription or paywall) and OA journals respectively, and whether published articles were also available in open access, online and in repositories, such as USP's own institutional repository, the Digital Library of Intellectual Output (in Portuguese, BDPI – *Biblioteca Digital de Produção Intelectual*). The collection and analysis of the quantitative data were complemented by the analysis of the transcripts of recordings of semi-structured interviews carried out with eleven scientists belonging to the two chosen postgraduate programs (PGPs), which aimed to assess their familiarity and perceptions of open access principles and practices.

2 SCHOLARLY COMMUNICATION AND OPEN ACCESS

There already exists a substantial amount of literature on OA, and it is not the purpose of this article to resume it here, but rather to focus on aspects that directly relate to our research

aims. The concept of the two ways to attain OA is much-discussed in the earlier literature on OA. Put simply, these refer to the publication in scientific journals that use the open access model ("gold OA"), or alternatively, to the deposit – through self-archiving or mediated archiving (and henceforth called “archiving”) – of scientific articles in institutional or subject repositories ("green OA"). The most famous and oldest repository, arXiv.org (hosted at Cornell University Library in the USA) is an example of the latter, which allows researchers from several fields in the Exact Sciences to upload their digital preprints (e-prints) of articles, frequently submitted simultaneously to subscription (and therefore non-OA journals). Other well-known subject repositories are repec.org which contains Economics papers, and cogprints.org for papers from the Cognitive Sciences and Psychology. Institutional repositories exist predominantly in universities and research institutes, although there are some in government organs. In Brazil, according to data collected from OpenDOAR in March, 2017, 80% of the 92 open access repositories are institutional, of which 11% are subject repositories covering specific disciplines, and 7% are governmental. Many of the Brazilian institutional repositories are called Digital Libraries (most of which contain full-text files of theses and dissertations), but the classification of these as repositories is correct. This is because they provide open access to the full-text files, as well as using the Open Access Initiative Protocol for Metadata Harvesting (OAI-PMH) to guarantee repository interoperability and exposure of metadata for third party information aggregation services like web search engines.

Since the inception of the OA movement, gold OA has usually been regarded as easier to attain as it did not represent a potential "threat" to publishers (GUEDON, 2006), given that the existence of OA journals merely constitutes more choice of journal titles to which authors may submit papers. Furthermore, with the growth of OA, many of the commercial publishers have established their own open access titles (e.g. Springer Open Choice, Nature Open, Palgrave Macmillan Open), the business model of which maintains the principle of open access for the reader at no cost, while charging Article Processing Costs (APCs) to authors (or their universities or research funding agencies), costs that cover publishing production overheads like graphic design and layout, dissemination, proofreading, language review etc. They do not include the costs of peer review, which has always been voluntary and unpaid work carried out by scientists from the wider scientific community. In the case of the Nature Publishing Group, APCs can range from \$2,000 to \$5,000 per article, depending on the journal. There are some well-known non-profit OA publishers, such as the Public Library of Science (PLoS) and BioMed Central, which use a variety of business models, sometimes even waiving the need for authors from certain countries to pay the APCs. With the emergence of the Public Knowledge Project's Open Journal System (OJS) software (adapted and promoted by IBICT in Brazil under the name of SEER), a myriad of OA journals starting in university departments and postgraduate programs have emerged. According to Bailey (2006, p. 23), OA journals have the following characteristics:

- 1) They are scholarly;
- 2) They utilize quality control mechanisms like those of conventional journals (such as editorial oversight and copy editing);
- 3) They are

digital; 4) They are freely available; 5) They may allow authors to retain their copyrights; 6) They may use Creative Commons or similar licenses.

Regarding the growth of OA, two changes, of a technical and political nature, have made green OA more feasible and attractive. The first was the development of the technical protocol for the exchange of metadata that allows the systematic harvesting of content found in repositories: the OAI-PMH, which in turn implied that Web browsers could collect and expose such metadata, thus making visible the content of collections in repositories scattered throughout the internet. The other major change was that several of the major commercial publishers accepted – some reluctantly – that authors of scientific articles routinely self-archive versions of their articles in an institutional or subject repository (GUEDON, 2006). The SHERPA-RoMEO service that collects and systematizes "publisher copyright and self-archiving policies", emerged precisely to clarify which journal for each commercial publisher permits such self-archiving in repositories, and if so, of which version (preprint, postprint, or publisher's copy) of the article. The SHERPA-RoMEO colour scheme communicates this spectrum of permissions, as presented in Chart 1 below:

Chart 1. SHERPA-RoMEO colour classification scheme

RoMEO colour	Archiving policy / permission
Green	can archive pre-print and post-print or publisher's version/PDF
Blue	can archive post-print (ie final draft post-refereeing) or publisher's version/PDF
Yellow	can archive pre-print (ie pre-refereeing)
White	archiving not formally supported

Source: SHERPA: Definitions and terms¹

For Harnad (2006), attaining a scenario in which 100% of scientific publications are available in OA will only be possible through green OA: that is, via the expansion of initiatives in universities and research institutes around the world in implanting their institutional repositories. Total open access via gold OA will not be possible, says Harnad (2006), because this will mean waiting for all scientific journals to be converted to open access, a highly impractical and unlikely scenario. More relevant is the fact that the expansion of gold OA lies beyond the control of institutions and individual researchers: it depends on a network of professionals, publishing expertise and funding to set up and maintain an OA journal. Green OA, on the other hand, depends mainly on whether scientists are motivated to archive their articles in their institution's repository, or in a subject repository.

What is important to highlight is that the archiving of an article in an OA repository can be entirely *complementary* to the publication of the same article in a non-OA (subscription) scientific journal: making an article OA in a repository and publishing in a high impact non-

¹

Available in:
<<http://www.sherpa.ac.uk/romeoinfo.html>>

OA journal are not mutually exclusive (SUBER, 2012; HARNAD, 2015). This may not be possible simultaneously, since many journals – both non-OA and OA – require an embargo period, which refers to the “waiting time” required by the journal between the publication date in the journal and making the article available in an OA repository. It should be emphasized that 866 out of 2,214 publishers (or 39%) registered in the SHERPA-RoMEO database permit the archiving in a repository of the pre-print or post-print of the published article, and a further 738 publishers (33%) allow the archiving of the post-print. It is important to note that SHERPA-RoMEO registers data on both non-OA (commercial) and OA publishers. For example, the database shows that Springer-Verlag's journal, *Engineering with Computers*, permits the archiving of both pre-print and post-print versions of the article in an OA repository, 12 months after its publication (and also allows the publishing of the post-print on the author's personal website immediately following publication). *Nature*, on the other hand, only allows archiving of the pre-print in an OA repository, and of the post-print six months after publication. In both journals, the final diagrammed, publisher's version (usually the PDF file), cannot be archived in an OA repository, technically making them both SHERPA-RoMEO “yellow” publishers (although the former is categorized in the database as a “green” publisher).

It would be understandable to question the need to archive copies of articles that have been previously published in OA journals in an institutional OA repository as well, since they can be freely accessed and downloaded from the journal website. But the function of an OA repository goes beyond that of merely providing access to the articles archived in it. In the case of institutional repositories, for example, one primary goal is to preserve the institutional memory, and another can be to generate and track the productivity indicators of institutional staff. Thus, archiving a copy of an article already available in an OA journal hosted by the SciELO platform, for example, should not be regarded as unnecessary duplication of the article, but rather a form of institutional digital curation and preservation. For OA journals, the SHERPA-RoMEO usually shows the statement: “This is a RoMEO ungraded journal” and signals that it is “Listed in DOAJ as an open access journal”, with the clear proviso being that despite being an OA journal, this does not necessarily mean that reuse policies are also guaranteed. It should be remembered that the focus of SHERPA-RoMEO's work is to clarify (and sometimes demystify) the policies and permissions of non-OA, subscription-based journals and their publishers regarding archiving in OA repositories. However, it is also worth remembering that, just by making articles available in OA to read does not necessarily mean that an OA journal allows their archiving in an OA repository: SHERPA-RoMEO (2017) itself alerts that:

This journal's policies have not been checked by RoMEO. DOAJ says this is an open access journal, but this may only mean that it is freely available to read. Most open access journals also permit self-archiving and re-use, but some do not. Do not assume that self-archiving is allowed, unless it is published under a Creative Commons license. Please contact the publisher for further information if necessary. Please contact us if you wish to suggest adding this publisher properly to RoMEO.

The term "open access" emphasizes *access* to the scientific literature; that is, considering primarily the end user, the reader, accessing that literature. Of course, it is relevant to appreciate

that every publishing scientist is also a reader and user of the scientific literature. That is, open access undeniably helps scientists during the research process, when they search for information. Today, it would be difficult to find a researcher who does not know how to mount a search expression in Google Academic, thereby immediately connecting that scientist to a network of publications whose metadata have been harvested from thousands of OA repositories around the world. As Guédon (2006, p. 28) observes, in the absence of OA, scientists lose access to potentially very relevant information if their institution does not pay a subscription to the relevant journal, which in turn means that many good ideas are not circulating within scientific communities, which could generate new advances for the scientific field in question.

Scientists are motivated to produce articles to be read, quoted and mainly used as a basis for further scientific advances. Such uses constitute the potential “impact” of their research results (published in the accessed article), and that is why research funding agencies reward scientists not merely for the number of articles published, but for their influence and scope, indicated in the measurement of the number of subsequent articles that use and cite them – their citation impact (GARFIELD, 2006). This means that publishing scientists are incentivised to be more interested in publishing in those journals with a greater “impact factor” – the Journal Impact Factor (JIF) – which measures the yearly average number of citations to articles recently published in the journal (THOMPSON REUTERS, 2012). The Qualis Journals system deployed by CAPES reinforces this evaluative approach at the journal level, classifying the journals where researchers working in Brazil publish into quality strata with A1 at the top, followed by A2, B1, B2, B3, B4, B5 and C being considered the worst quality. These quality measures form an integral part of the four-yearly post-graduate program assessment exercise, making such programs stipulate that its researchers (teachers and students) publish in journals highly evaluated by the Qualis CAPES measure.

There is now a body of literature that demonstrates a citation advantage that can be achieved by making articles available in open access, be it through a full OA journal, a hybrid journal (non-OA, subscription journals that offer the author the choice to make the article OA on publication, by paying APCs), or through archiving in a repository. This advantage is known as the Open Access Citation Advantage (OACA), which has already been shown to exist in some scientific fields in studies that compare citations of non-OA and OA articles published in the same journal volume and number (HARNAD; BRODY, 2004; ARCHAMBAULT et al., 2016). This advantage is an article-level metric (or ALM) metric rather than journal level (such as the JIF). As already mentioned, citations of OA publications can be made from the published version in a journal or from the same article found in an OA repository, the latter often discovered using internet search engines such as Google Scholar. Even so, it can be seen from article referencing practices that authors tend to reference and cite the bibliographic data of the published version from the journal, even if the article was indeed downloaded from a repository; it is unusual to see the URL of the repository being listed in references, and so the visibility of the journal tends to be maintained in references.

3 METHODOLOGICAL PROCEDURES

The aim of this research was to identify the practices and perceptions of researchers of two high-performing Engineering Postgraduate Programs (PGPs) in relation to publication in non-OA vehicles (journals) or OA ones (journals, repositories and personal webpages). An exploratory mixed-methods research approach was adopted, using both quantitative and qualitative data collection methods (CRESWELL, 2010). Creswell (2010) observes that researchers can enrich analysis if they seek to combine and connect data collected using quantitative and qualitative methods, and he refers to the “incorporation” of one set of data into another, to support and augment analysis. In the study reported here, the quantitative data collected pertained to details of publications by the researchers studied, while the qualitative data originated from interviews conducted with a group of the researchers, as discussed below.

Data were first collected from the University institutional webpages to identify research subjects from the two highest-performing (according to the CAPES evaluation exercise) Engineering PGPs. Postgraduate programs in Engineering were chosen because the University campus in question has a tradition, since 1948, of renowned high-quality Engineering teaching and research in Brazil. A total of 47 researchers (22 from the first Postgraduate program, 25 from the second) were identified who had published in the period 2008 to 2016, and with the names of these researchers, details of their publications were then extracted from their curriculum vitae published on the Lattes Platform, using the open source software ScriptLattes (MENA-CHALCO; CESAR JUNIOR, 2009).

The journal titles in which the researchers published in the period studied were consulted in DOAJ to determine if they were OA journals or not. Then, for those articles published in non-OA, subscription journals, the SHERPA-RoMEO database was consulted to verify the permission policies for archiving in OA repositories for each journal in question. Given that for many of the articles published in non-OA journals, the embargo period had already elapsed, or the journal’s policy permitted archiving of a version of the article in an OA repository, we then set out to discover how many of these articles were, in fact, fully available in open access on the Internet. To that end, we checked for the availability of the articles either in the institutional repository (BDPI), as well as carrying out searches in Google Scholar (one of the most frequently-used and comprehensive academic search engines), because Google Scholar retrieves the metadata and links to copies of articles archived in OA repositories worldwide. All quantitative data collected in these methodological steps were tabulated and stored in a Microsoft Office Excel® spreadsheet.

Simultaneously, semi-structured interviews (RICHARDSON, 2012) were carried out with a subset of 11 researchers from the two PGPs studied. In semi-structured interviews, the interviewer uses a question guide to guarantee that the same topics were dealt with in each interview, and the questions asked are content-focussed, aiming to deal with issues and areas judged by the researchers to be relevant to the research question. All of the 47 scientists who

teach, research and publish on the two PGPs were invited to participate in the interview stage of the research, so this subset of just 11 corresponds to those who replied in the affirmative to the e-mail sent out (twice), inviting them to participate in our study. In the interviews, which took approximately 35 minutes, participants were asked to give their impressions of the scholarly publishing system in general, and specifically of OA journals and OA repositories, as well as to comment on their degree of familiarity with their own institution’s repository and policy. The interviews were recorded with the permission of each interviewee, and later transcribed. The fact that the interviews had been guided by a topic guide used by the interviewer helped in the identification of primary topic categories in the transcriptions (RICHARDSON, 2012). The units of analysis of this textual content and its subsequent representation and discussion in the results related to each group of data obtained and treated in the quantitative phase of the study, was thus facilitated. For some of the questions asked in the interview, to which a simple “yes”/”no” answer could be given, we tabulated the results (as seen in Table 2), even if the interviewee had spoken in more detail on the topic, by illustrating with examples that had occurred, for example.

4 RESULTS AND DISCUSSION

Using ScriptLattes, the extraction of publication data from the researchers’ Lattes curricula revealed that a total of 970 articles were published between the two Postgraduate Programs (PGPs) for the period of 2008 to 2016. The 466 articles published by researchers from PGP-1 were distributed among 145 journals, and the 504 articles from the PGP-2 were distributed among 108 journals, as shown in Figure 1.

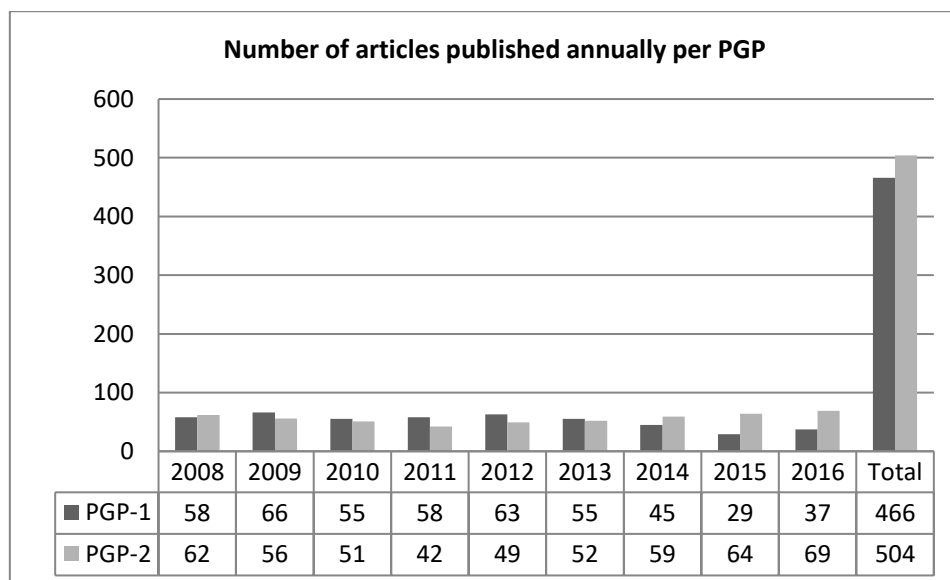


Figure 1. Number of articles published by researchers from the 2 PGPs between 2008 and 2016
Source: Authors

It was also observed that of the titles chosen for publication, 260 (55.8%) of those belonging to PGP-1, and 376 (74.6%) of those of PGP-2 fall into under the A1 to B2 strata of

the Qualis CAPES journal evaluation system. This underlines researcher concern to publish in the more highly-rated journals to maintain the high score (also attributed by CAPES) of the respective PGP. It was notable that in all the interviews, participants referred to the Qualis CAPES journal classification as the relevant criterion when choosing a journal title for manuscript submission. The explanation below given by one of the interviewees reflects such a preoccupation, together with the need to publish results seen to be more relevant to a national, Brazilian audience in the professional Engineering community:

“We chose fundamentally in view of at least two parameters. First parameter: it has to be a journal that has a CAPES classification, right? In our area of activity, for example, we are here in Engineering. So, we chose the best possible classification. We are looking for that journal which we know has the greatest visibility in the country, in Brazil. Why? Because we understand that being a Brazilian publicly-funded institution, we owe it to Brazilian society... (...) So we give the preference to publish in Brazil. It is obvious that sometimes we have some situations when this is not possible: we have an article whose approach is a little more generalized, so from there, we are looking for a journal that is, say, AI, because it presents more generalized information for many situations. These are the criteria that we have used.” (P9)²

After conferring in DOAJ to ascertain whether the journals in which the researchers from the two PGPs had published were OA or non-OA journals, the results presented in Figure 2 were obtained, where it can be seen that 102 (21.9%) of the articles published by authors from PGP-1, and 150 (29.8%) of articles published by authors from PGP-2 are published in OA journals.

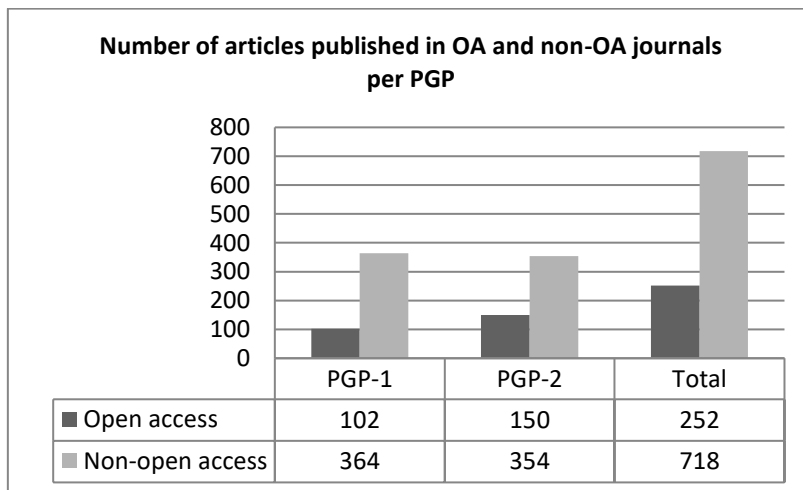


FIGURE 2. Quantity of articles published in OA and non-OA journals
Source: Authors

These results reflect those opinions expressed by the researchers in the interviews, when they were asked about OA journals. Five of the eleven interviewees spoke in favour of OA journals, but five others declined to comment, and one spoke out against them, stating: *“I am*

² Interviews conducted with two groups of accredited scientists in one of two postgraduate programs in Engineering at the University of São Paulo - USP.

against them. I do not think you should pay to publish. I, for example, do not practice this because I do not think it makes sense for you to pay to publish." (P8). In fact, the five interviewees in favour of them had also expressed this reservation of having to pay to publish in OA journals. Here, it is relevant to remember that OA literature, like non-OA, is not free to produce or publish. As Suber (2012, p. 21) clarifies:

OA isn't an attempt to deny the reality of costs. No serious OA advocate has ever said that OA literature is costless to produce, although many argue that it is much less expensive to produce than conventionally published literature [...].

Because all scientific journals – non-OA and OA – are not free to produce, some OA journals charge Article Processing Charges (APCs) to authors (or research funding agencies) precisely because they do not charge readers and users access charges, which non-OA journals do in the form of subscriptions (or article-level access charges). Suber (2012) observes that most researchers are more aware of gold OA than green, to the point that the concept of OA for many actually means OA *journals* – gold OA – and not the repositories of green OA. Suber (2012, p. 137-8) expands on this, noting:

Apart from the myth that all OA is gold OA, the most common myth about gold OA is that all OA journals charge "author fees" or use an "author-pays" business model. There are three mistakes here. The first is to assume that there is only one business model for OA journals, when there are many. The second is to assume that charging an upfront fee means authors are the ones expected to pay it. The third is to assume that all or even most OA journals charge upfront fees. In fact, most OA journals (70 percent) charge no upfront or author-side fees at all.

Despite the fact that some top quality OA journals waive publication fees for authors without research funding, this had evidently not been the case for the interviewees in our study. Even those expressing opinions in favour of OA in principle, complained about the need to pay APCs, as seen in the following excerpts from the interviews:

*"In our area, it's the payment that restricts us. For example, we now publish in an open access journal, only it is open access for the user, but the author has to pay and it really is expensive! I don't remember if it was \$600 or \$ 500 we paid. (P11)
What is happening is that open access journals charge and closed access journals do not charge. As we generally do not have the funds to publish, or have little, we end up publishing in the closed access journals, even if we want the article to be available in open access."* (P4)

In fact, only three of the respondents had paid to publish in an OA journal, compared to two that said they never had; the other six researchers expressed serious doubts about it. It was also notable that only researcher perceived that OA journals adhere to the same peer review quality standards as non-OA, subscription journals, commenting that *"[...] They are open access but they still set high, demanding standards for publication."* (P6).

After having identified the quantity of articles published in OA and non-OA journals by researchers from the two Programs (as presented in Figure 2), the following step was to

identify, in the SHERPA-RoMEO database (at <http://www.sherpa.ac.uk/romeo/>), the archiving policies for the non-OA journals for each Program, to verify if any version of the article could, in fact, be made available in OA. The results demonstrate (in Table 2 below) that, for the articles published in non-OA journals by the researchers of the two PPGs, most (391) were published in journals that deploy a green archiving policy. This means that they permit the archiving of a version of the article in an OA repository, often after an embargo period, and often in the final PDF format, using the publisher layout. In addition, it was found that for the articles identified that had been published in journals with this green archiving policy, the embargo period had already expired, meaning that these articles can be made available in an OA repository.

Table 1. Distribution of the archiving policies of the non-OA journals and the number of articles published in these, by PGP

RoMEO archiving policy	PPG-1	PPG-2	
Green	159	232	391
Blue	14	11	25
Yellow	9	19	28
White	10	19	29
RoMEO ungraded journal	20	13	33
Not found in RoMEO	152	60	212
TOTAL	364	354	718

Source: Authors

For each of the 718 articles published in non-OA journals that each have different archiving policies for OA repositories (as shown in Table 1), a title search of the article was conducted in Google Scholar and then in USP’s institutional repository, BDPI, to verify if they are available in OA. The premise here was that, if the publisher’s embargo period had already elapsed, then these articles would already be available in OA. The results are shown in Figure 3. It should be noted here that these searches were carried out using a private, residential internet connection, free of access proxies either to the CAPES Journal Portal or to any USP intranet, with none of the access privileges that academic users sometimes have, thereby simulating searches made by a user from the wider general public, for whom, as taxpayers subsidising scientific research carried out in public universities, open access to the scientific literature constitutes a gainful right. (ZUCCALA, 2010; WILLINSKY, 2006).

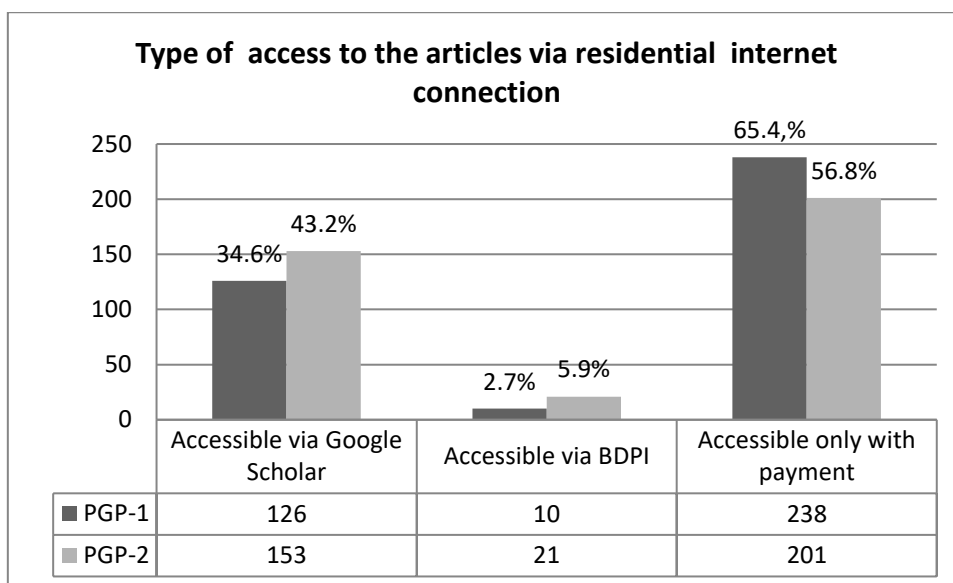


Figure 3. Types of access to articles published in non-OA journals
 Source: Authors

Through these searches, it was observed that in PGP-1, there were 126 articles (34.6%) with the PDF file available in OA through the Google Scholar search, compared to 238 (65.4%) whose access demanded the payment of fees, charged in US dollars or Euros. For PGP-2, these figures were 43.2% (153 articles) and 56.8% (201 articles) respectively. It should be remembered here that when it is stated that articles are “available through Google Scholar”, this is because they have been archived in an OA repository somewhere in the world, with the Google Scholar search engine harvesting and exposing the article’s metadata for the user to download from the repository (or repositories, if it has been archived in more than one) in question.

On verifying the number of articles available in OA in the university’s institutional repository, using the repository’s own search engine, the proportions of articles available in open access fall dramatically, with only 2.7% for PGP-1 and 5.9% for PGP-2 of the total amount of articles for each PGP available in OA. It is disconcerting that out of a total of 391 articles from both PGPs that could potentially be available in OA when we consider that they were published in journals with "green" archiving policies, only a significantly low amount – 10 articles for PGP-1 and 21 for PGP-2, are indeed available in open access through a residential connection to the Internet.

During the interviews, when the subject of archiving in the USP institutional repository arose, only three of the interviewees were aware of this repository’s existence, and these had only a vague notion of the possibility of copies of their publications being archived there. During the interviews, the interviewer cited Resolution USP 6444/2012, which establishes the institutional repository and recommends that “all members of the USP community publish their research results, preferably in OA vehicles or in those that permit, in their publication

agreements, the archiving of a copy of the publication in the BDPI.” Some interviewees were surprised to discover this “freedom” to archive copies of their published articles in the BDPI repository. Reflecting the results of similar surveys conducted outside Brazil (MOORE, 2011; SWAN, 2006; PARK; QIN, 2007), this possibility runs up against scientists’ concerns that they may be violating copyright agreements that they have already accepted at the moment of submission of the manuscript, as expressed by two interviewees:

“We’ve always been a little afraid of the legal part, because in fact, when you publish in a journal of certain impact, they ask you to sign a document giving the copyright to them. So, I do not know to what extent (...) right? But we get scared, so I really do not know, I do not know this directive [referring to the Resolution] there. I know there is a digital bank of theses, dissertations, but those things that do not have copyright, right? Student’s copyright and the student in general have the interest and divulge so he puts everything else that involves the copyright we have a bit of fear. (P3)”

“For the interest of the researcher is interesting to the maximum disclosure, the maximum disclosure of the product, now, really did not know. I always thought it was a conflict, because the paper arrives to sign and we sign it here and it’s ready, now we cannot, I cannot anymore, that was my idea, right? It’s news to me. It doesn’t make sense.” (P10)

Despite the fact that the more specific details of self-archiving or mediated archiving was not explicitly raised in the interviews, one interviewee added that *“I think the librarians themselves could do for us, right?”* (P11). However, considering that more than half of all the publications published here are registered with the BDPI (as shown in Figure 3), it is deduced that, for now, USP librarians are responsible for In relation to the other questions asked to the scientists interviewed, we have the following results in tabulated:

Table 2. Participant answers on open access topics raised in the interviews

	Yes	No	Uncertain
Have you ever used an OA repository? Do you know any?	2	3	6
Have you ever received emails from people outside USP requesting a copy of one of your papers?	6	3	2
Have you ever experienced difficulty in accessing a paper that you needed?	4	6	1

From these results, we observed that knowledge of OA repositories is scarce among the participants, with only two of them claiming to have used or accessed them. Such a knowledge gap is perhaps not surprising: although Google Scholar was mentioned at several points in all of the participants’ interviews, there is not necessarily the clarity that this platform collects, aggregates, and displays publication metadata from the worldwide network of OA repositories. And the fact that 6 of the 11 scientists interviewed responded to requests for their articles from people outside the University not only exemplifies the continuity of the millennial tradition intrinsic to the scientific community of sharing the scientific knowledge produced by them, but it also points to a dimension that can be leveraged in awareness-raising campaigns on the benefits of scientists ensuring that their own articles are available in OA repositories. Indeed, it was on this issue of such requests coming from people in countries in a less favourable

situation than Brazil, that the participants spoke more intuitively on the need to institute OA to the scientific literature. The fact that a little more than half expressed that they themselves had never experienced difficulty in accessing an article they needed, attests to the comparatively privileged situation of the group of scientists who took part in our study in which access to the literature is not deemed to be problematical: in addition to having access to the CAPES Journals Platform, as is the case for all scientists employed in public universities in Brazil, they also teach and research on highly-rated PGPs at one of the most prestigious and respected university in Latin America (USP). These access privileges diminish a real need for OA, a factor that can sometimes translate into lack of a need to even know about OA.

5 CONCLUSION

The aim of the study presented here was to collect and analyse the perceptions and publishing practises in open access (OA) and non-OA of a group of Brazilian researchers who teach and research on highly-ranked postgraduate programs. To that end, we carried out quantitative data collection, comparing their publishing practices in OA and non-OA journals, as well as mapping the free and open availability of their published papers in repositories. These aims were attained in the course of our study.

In the light of the literature on OA that describes the notable growth of publications available in OA over recent years, we expected a greater percentage of articles published in OA by the researchers covered by our research. We deduce that these authors often fail to publish in OA not intentionally, but rather they are influenced by the policies of development and evaluation agencies (such as CAPES, for example), which prioritize publications in journals with a high impact factor (IF). For the participants of this research, publication in non-OA journals of high IF is not compatible with OA, mainly because they are unaware of the option they have of self-archiving copies of non-OA published papers in their institutional repository. That is, the fact that self-archiving in a repository and publishing in a high IF journal can be paths that complement each other is not clear to the research participants: it is hardly known that non-OA subscription journals allow authors to make their articles available in OA repositories after the embargo period has elapsed. Although SHERPA-RoMEO has proved to be of extreme importance as a tool for consulting copyright and self-archiving policies, the site has a significant number of journals whose self-archiving policies have not yet been verified (starting with all of those journals already classified as OA), which could generate insecurity for some authors (or librarians acting as intermediaries in the archiving process) who intend to archive their articles in repositories. We suggest a greater dissemination of the SHERPA-RoMEO platform here (which also has a Portuguese interface) so that more authors can use it and with this, make more and more articles available to the whole community.

Taking into account the embargo periods and the archiving policies of the journals surveyed, we observed a significant amount of articles unavailable in OA repositories, and specifically in the institutional repository of USP, even after four or five years of publication.

Therefore, the university could foster more OA advocacy for the use of its institutional repository among researchers and teachers, and even to use its network of librarians to archive these articles in the repository to ensure access to them, and to make visible an optimally-populated institutional repository for the university. It will be by these means that USP, as a prestigious, highly-regarded university in Latin America, will reap the benefits, along with society in general, of making the results of its research more widely accessible.

REFERENCES

- ARCHAMBAULT, Éric; CÔTÉ, Grégoire; STRUCK, Brooke; VOORONS, Matthieu. **Research impact of paywalled versus open access papers**. Disponível em: <<http://www.lscience.com/oanumbr.html>> Acesso em: 24 ago. 2016.
- BAILEY, Charles William Jr. What is open access? In: JACOBS, N. **Open access: key strategic, technical and economic aspects**. Oxford: Chandos, 2006. p. 13-26.
- BUTLER, Declan. The dark side of publishing. **Nature**, v. 495, n. 7442, p. 433-435, 2013.
- CAPES. **Histórico**. Disponível em: < https://www.periodicos.capes.gov.br/?option=com_pcontent&view=pcontent&alias=historico&mn=69&smn=87 >. Acesso em: 09 abr. 2017.
- CRESWELL, John. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. Porto Alegre: Artmed, 2010.
- GARFIELD, Eugene. Citation indexes to science: a new dimension in documentation through association of ideas. **International Journal of Epidemiology**, v. 35, n. 5, p. 1123-1127, 2006.
- GUEDON, Jean Claude. Open access: a symptom and a promise. In: JACOBS, N. **Open access: key strategic, technical and economic aspects**. Oxford: Chandos, p. 27-38, 2006
- HARNAD, Stevan; BRODY, Tim. Comparing the impact of Open Access (OA) vs. Non-OA Articles in the same journals. **D-Lib Magazine**, v. 10, n. 4, 2004.
- HARNAD, Stevan. Optimizing open access policy. **The Serials Librarian**, v. 69, n. 2, p.133-141, 2015. Disponível em: < <http://eprints.soton.ac.uk/381526/> >. Acesso em: 24 out. 2016.
- HARNAD, Stevan. Opening access by overcoming Zeno's paralysis. In: JACOBS, N. **Open access: key strategic, technical and economic aspects**. Oxford: Chandos, 2006. p. 73-98.

MENA-CHALCO, Jesus Pascual; CESAR JUNIOR, Roberto Marcondes. ScriptLattes: an open-source knowledge extraction system from the Lattes platform. **Journal of the Brazilian Computer Society**, v.15, n.4, p.31-39, 2009.

MOORE, Gale. **Survey of University of Toronto Faculty awareness, attitudes and practises regarding scholarly communication**: a preliminary report. Toronto: University of Toronto, 2011. Disponível em: < <https://tspace.library.utoronto.ca/handle/1807/26446> >. Acesso em: 15 set. 2014.

PACKER, Abel. The SciELO open access: a gold way from the South. **Canadian Journal of Higher Education**, v. 39, n. 3, p. 111-126, 2009.

PARK, Ji-Hong; QIN, Jian. Exploring the willingness of scholars to accept open access: a grounded theory approach. **Journal of Scholarly Publishing**, v. 38, n. 2, jan. 2007. p. 55-84.

RICHARDSON, Roberto Jarry. **Pesquisa social**. Métodos e técnicas. São Paulo: Atlas, 2012.

SHERPA-RoMEO. Disponível em: < <http://www.sherpa.ac.uk/> >. Acesso em: 08 maio 2016.

SWAN, Alma. The culture of open access: researchers' views and responses. In: JACOBS, N. **Open access**: key strategic, technical and economic aspects. Oxford: Chandos, 2006. p. 65-72.

SUBER, Peter. **Open access**. Massachusetts: MIT Press, 2012.

THOMPSON REUTERS. **Journal Citation Reports**. Atualizado em 2012. Disponível em: < http://admin-apps.webofknowledge.com/JCR/help/h_impfact.htm >. Acesso em: 12 maio 2016.

WILLINSKY, John. **The access principle**. The case for open access to research and scholarship. Cambridge, Mass.: MIT Press, 2006.

ZUCCALA, Alisia. Open access and civic scientific information literacy. **Information Research: An International Electronic Journal**, v. 15, n. 1, 2010.

AGRADECIMENTOS

Agradecemos ao CNPq/CAPES pelo apoio financeiro fornecido à primeira autora, processo número 401875/2011-3.



