

OPEN SCHOLARLY COMMUNICATION IN THE  
EUROPEAN RESEARCH AREA FOR SSH -  
PREPARATION

WP6 Innovation

D6.4 Report on innovative models of  
bibliodiversity in scholarly publications

30 April 2021

DRAFT



## Deliverable 6.4

### Report on innovative models of bibliodiversity in scholarly publications

<b>Grant Agreement number</b>	: 871069
<b>Project acronym</b>	: OPERAS-P
<b>Project title</b>	: Open Scholarly Communication in the European Research Area for SSH - Preparation
<b>Funding Scheme</b>	: INFRADEV-02-2019-2020
<b>Project's coordinator Organization</b>	: CNRS-OpenEdition
<b>E-mail address</b>	: pierre.mounier@openedition.org
<b>Website</b>	: <a href="https://www.operas.unito.it/projects/operas-p/">https://www.operas.unito.it/projects/operas-p/</a>
<b>WP and tasks contributing</b>	: WP6, Task 6.4
<b>WP leader</b>	: IBL Pan
<b>Dissemination level</b>	: PU
<b>Due date</b>	: 30/04/2021
<b>Delivery date</b>	: 30/04/2021
<b>Authors</b>	: <b>Authors:</b> Delfim Leão, Ana Balula, Carlos Costa, Cornelia Plag, Lorena Caliman, Maria João Padez, Mickael Silva, Susana Jarmelo (Coimbra UC) <b>Contributors:</b> Pierre Mounier (OpenEdition) with the collaboration of the Multilingualism SIG (including Representatives of the Helsinki initiative, Federation of Finnish Learned Societies, ENRESSH, AEUP, Academic Book Publishers, PKP, FORCE11)



**Table of content**

<b>Introduction</b>	<b>4</b>
<b>Phase 1</b>	<b>5</b>
<b>Phase 2</b>	<b>15</b>
<b>Phase 3</b>	<b>34</b>
<b>References</b>	<b>36</b>

DRAFT

## Introduction

Research questions that this task aims to answer:

Objectives: Given the growing need to strengthen the bonds between stakeholders involved in scholarly communication and multilingualism, this WP has a three-fold purpose: (i) synthesize evidence in the literature as to innovative dynamics of knowledge-sharing and scholarly communication within linguistically diverse scholarly contexts and research networks; (ii) have a better understanding of the role of multilingualism within bibliodiversity in scholarly communication, through the lens of publishers and translators/researchers; and (iii) present the conceptual design of a future OPERAS Translation Platform aiming at supporting translation services at the scholarly communication level (involving publishers, translators, researchers).

Goals:

1. to prepare theoretical background to discuss the usage of multilingualism in scholarly communication
2. to identify, analyse, and understand innovative dynamics of working practices and knowledge-sharing within linguistically diverse scholarly contexts and research networks
3. to identify and analyse the motivations behind these practices (questionnaires/focus groups - how tools may answer to needs)
4. to formulate recommendations/guidelines for OPERAS regarding the future implementation of a service aiming at enhancing multilingualism
5. to prepare the conceptual design of a platform prototype for shared translation service at the scholarly communication level (involving publishers, translators, researchers).

## Phase 1

### **Phase 1: Literature review: multilingualism and scholarly communication**

**[corresponding to goals 1 and 2 ]** Initial draft of the literature review by Ana Balula, with the collaboration of Delfim Leão, who have published a paper (Balula and Leão, 2021) based on this review, but prior to the publication of the OPERAS-P report. This part of the report was revised within the OPERAS Special Interest Group on Multilingualism, with concrete contributions by Gimena del Rio Riande, Ivo Volt, Jadranka Stojanovski, Janne Pölönen, Jean-François Nomine, and Pierre Mounier.

### **Methods**

This is a qualitative study of exploratory nature and the method used is in the scope of an integrative literature review, summarising prior research to clarify research trends based on *in vivo* content analysis of the selected corpus. This method follows stages, starting with problem formulation, which frames data collection, selection, treatment and analysis, as well as the presentation of results. In what respects the problem, it reflects a gap in the recent literature, namely as to identifying factors that influence the dynamics underlying language selection and the use of multilingualism within scholarly communication. The database selected was Google Scholar and the search terms used were 'scholarly communication' 'language' and 'multilingualism' – combined with the Boolean operator AND. The search, undertaken on 6 April 2020, yielded 152 works. These results were reviewed to exclude duplicates, PhD and Master's dissertations, as well as works that did not meet this literature review's goals. In what concerns selection criteria, the works had to: i) have a DOI (Digital Object Identifier) code, ii) be published in open access, between 2019 and 2020, and iii) be written in English, French, German, Portuguese, Italian or Spanish. This resulted in the selection of 12 documents, which were analysed resorting to qualitative content analysis of the abstracts and conclusion sections. Subsequently, the final category framework reflects the *corpus* codification structure that emerged from the analysis, as presented below.

### **Results and Discussion**

The categorisation of the *corpus* regarding the dynamics between multilingualism and scholarly communication in SSH were identified *in vivo* and translates into: 1. Research relevance, 2. Content curation, 3. Reputation and 4. Balanced multilingualism (see Table 1).



Categories		Studies
1. Research relevance	1.1. Global Englishisation	Balula and Leão (2019), Costello (2020),  Giglia (2019), Guns, Eykens and Engels (2019), Jokić, Mervar and Mateljan (2019)  Kulczycki <i>et al.</i> (2020), Luzón (2019), Sivertsen, Guns, Kulczycki and Pölönen (2019)
	1.2. Language competence	Balula and Leão (2019), Costello (2020),  Jokić, Mervar and Mateljan (2019), Kulczycki <i>et al.</i> (2020)
2. Content curation		Balula and Leão (2019), Giglia (2019),  Ibrahim (2019), Koutsomitropoulos (2019)  del Rio Riande, Tóth-Czifra, Wuttke and Moranville (2020)
3. Reputation		Balula and Leão (2019), Costello (2020),  Guns, Eykens and Engels (2019),  Jokić, Mervar and Mateljan (2019),  Kulczycki <i>et al.</i> (2020), Luzón (2019),  Sivertsen, Guns, Kulczycki and Pölönen (2019)
4. Balanced multilingualism		Balula and Leão (2019), Guns, Eykens and Engels (2019), Kulczycki <i>et al.</i> , (2020)

*Table 1. Categorisation of the corpus regarding the dynamics between multilingualism and scholarly communication in SSH.*



## Research relevance

Starting with the category *Research relevance*, it refers to the language selected for scholarly communication purposes and subdivides into two sub-categories, i.e. 1.1. *Global Englishisation* and 1.2. *Language competence*. Regarding the former, English is assumed, worldwide and in multiple societal contexts, as language of communication, posing as important means to disseminate (share and search for) information. When it comes to research, the results of the literature review undertaken by Balula and Leão (2019:4) underline that “in terms of information availability, which underpins the co-construction of knowledge, the use of English as *lingua franca* promotes the dissemination of research outputs and breakthroughs”. In this respect, Luzón (2019) recalls that, in the Science and Technology fields, most authors use English as the default working language within the disciplinary community. Nevertheless, many SSH researchers frequently develop culturally and societally relevant studies in their local languages, in particular, because their aim is to contribute to the debate, decision-taking and innovation processes in specific areas, such as cultural heritage, education, migration, public administration, etc. (Kulczycki *et al.*, 2020; Giglia, 2019; Luzón, 2019).

Jokić, Mervar and Mateljan (2019:1007) points out that, although the “contemporary scientific communication pattern amongst CEE countries regarding social sciences is unknown”, it seems that SSH issues, in particular those focused on national or local realities, tend to be more relevant to national (rather than international) scientific communities and, thereof, scholarly communication is undertaken in native languages. Giglia (2019:143) also emphasises that “SSH research is often grounded in specific cultural or geographical areas, hence the persistence of native languages opposed to English as *lingua franca* in STEM”. This trend seems to reflect in institutional research websites, once, as Luzón (2019) concluded for the case of Spanish research groups, those presented only in native language seem to aim at having a local impact. Even though some researchers also provide links to works written in English to reach a wider (more international) community, most tend to write in their native languages because their funding and resources (including time) are limited (Luzón, 2019).

Within European scholarly publication in SSH, there is “a huge number of small size players of different types and quality serving local scientific communities and specializing in narrow fields of research” (Giglia, 2019:143). In this scenario, the sole use of national languages can pose as a relevant fragmentation element (specially for languages used by less population, e.g. such as Finnish), inhibiting international information-sharing and co-construction



of knowledge (Balula and Leão, 2019). Thus, authors as Kulczycki *et al.* (2020) and Balula and Leão (2019) underline that global Englishisation of scholarly communication can have two main consequences: on the one hand, it can facilitate the sharing of research data and results, but, on the other hand, it can also restrict important contributions from non-Anglophone researchers. Besides, the analysis of information produced in native language may contribute to enrich the readers' multicultural and multilingual awareness and competence, and allow the scientific community to skim and scan for information.

In his study, Costello (2020:3) concludes that “the pre-eminence of English hence appears to be edging out other languages in an increasing trend. It is assembling an unassailable claim to be the language of legitimation, the language of science and scholarship – the language of knowledge and ultimately, of truth”. In this respect, Guns, Eykens and Engels (2019) highlight that, although the number of works published in English by non-Anglophone authors is tending to increase (almost doubling), the number of publications in the non-English researchers' native language does not seem to be decreasing, at least so evidently.

SSH research is being published in both local and international journals (Sivertsen, Guns, Kulczycki and Pölönen, 2019). When it comes to language selection, several authors point out the reasons underpinning the choice of English, which are closely interconnected, namely: research outreach, the need for internationalization, collaboration practices and target audience (e.g. Kulczycki *et al.*, 2020; Luzón, 2019; Guns, Eykens and Engels, 2019; Balula and Leão, 2019). In this context, it is important to underline that there are non-English journals that reach a wide international community and that some international journals are increasingly accepting and promoting the use of English *pari passu* with other international languages (French, Spanish, Portuguese, etc.) or even multiple languages (Sivertsen, Guns, Kulczycki and Pölönen, 2019). In this respect, Kulczycki *et al.* (2020:2) highlight that “publishing in more than one language not only reaches a wider audience but supports a diverse perspective on research”. Thus, there seems to be a slow shift in the way language issues are being dealt with in research, valuing, for instance, the combination of “different languages to cater for various audiences” (Luzón, 2019:54) within scholarly communication platforms. Authors such as Sivertsen, Guns, Kulczycki and Pölönen (2019), and Guns, Eykens and Engels (2019:2) sustain this trend, once the results of their study reveal that SSH researchers “write a substantial share of publications in a local language, in addition to publications in English or other international languages”, which may evidence a shift of scholarly communication towards inclusive multilingualism,



as defined by the Helsinki Initiative on Multilingualism in Scholarly Communication (for more details, see: <https://www.helsinki-initiative.org/>).

Concerning the sub-category *Language competence* (1.2., see Table 1), it is important to start by underlining that, in scholarly communication, the selection of a certain language has an impact on the way information is conveyed and understood, as concepts and narratives are constructed and de-constructed within dynamic linguistic (and cultural) lines and linguascapes, even if language competence of researchers needed for reporting research results may be increasingly limited to the native language and English (Kulczycki *et al.*, 2020). Nevertheless, several authors (e.g. Balula and Leão, 2019; Sivertsen, Guns, Kulczycki and Pölönen, 2019; Costello, 2020) conclude that research not written in English seems to be undervalued.

Considering that culture is embedded in language itself, the production of research in a foreign language is necessarily demanding and complex for non-native speakers, since it implies “the transference between different conceptual mind-sets” (Balula and Leão, 2019:4). In the case of some researchers that have low foreign language proficiency (and in particular, in English), there are cases in which “some resort to practices of copying and pasting fragments of English text, before then attempting to edit and refactor these reproductions to new ends: trying to build a picture, but with pieces drawn from different jigsaws” (Costello, 2020:3) – and this tends to undermine publication itself. Thus, having a multilingual approach to research is not always easy for scholars, not only because of the lack of a good command of foreign languages, but also because the budget for translations is often scant or non-existent.

After having the English version of the research, authors face another great challenge – the peer review process. Regarding language, after submitting a research output written in English, many scholars often receive a variation of the following critical remark: “This manuscript could benefit from proofing by a native English speaker” (Costello, 2020:1). Although there are cases in which it fully applies, this also happens when works are written by (proficient) native speakers, a situation that definitely underlines the power of language competence within research, as well as the clear empowerment of reviewers as language specialists. In fact, on the one hand, studies written in English are often revised by non-Anglophone scholars and works written in other languages are not widely reviewed by Anglophone peers and, on the other hand, “not all ‘natives’ have high competence in English [or any other language], just as ‘non-natives’ do not necessarily lack this competence” (Costello, 2020:2). Costello (2020) even associates ‘native speakers’ with the so-called ‘digital



natives', i.e. the same way we should not expect the latter to know by instinct how to use digital technology to identify reliable information sources, or to use a specific software for learning or at work, we should also not assume that proficient native speakers master the academic jargon of several specific areas.

In closer look into research work, a study focussing on seven European countries (Czech Republic, Denmark, Finland, Flanders [Belgium], Norway, Poland, and Slovenia) points out that most “articles in English were more often written by multiple authors rather than one author. We also found that, for all countries, articles that were written in local and other languages were more often single-authored than multi-authored” (Kulczycki *et al.*, 2020:10). In addition, a commonality unveiled in the study developed by Jokić, Mervar and Mateljan (2019:1007), for the case of central and eastern European countries, is that “almost all quotes were from authors from their native countries”, which may also evidence the dramatic effect of language barriers in scholarly communication, and subsequent co-construction of knowledge. Thus, Englishisation does not seem to fully address research intended main goals, concerning information sharing and discussion, as well as co-construction of knowledge, for which multilingualism can pose as an important asset, while promoting inclusiveness and equity of researchers (Balula and Leão, 2019).

### Content curation

There are several initiatives trying to facilitate multilingual and multiformat interaction between researchers from several areas (e.g. European Commission's *European Open Science Cloud*, OPERAS, etc.), which may play a very important role in conferring societal relevance to research (Balula and Leão, 2019). Some initiatives, such as OPERAS (through the TRIPLE project: <https://operas.hypotheses.org/projects/triple>), are working on the creation of platforms, which allow for “all European researchers in SSH to discover, from a single point of entry, open resources (data, publications and other materials) relevant to their research. The added value consists in the feature of indexing resources with disciplinary ontologies and thesauri and to align them across several languages” (Giglia, 2019:151–152) and, in this scenario, *Content curation* (see 2., Table 1) plays a pivotal role.

Some authors report on their contribution in this area, i.e. the development of platforms, such as: *Isidore Discovery Platform* (<https://isidore.science/>) – a research discovery platform (Mounier, 2018 *cf.* Balula and Leão, 2019), or *OpenMethods* (<https://openmethods.dariah.eu/>) – which “is intentionally interdisciplinary and multilingual to facilitate a timely, global disclosure and dissemination of knowledge and to raise peer recognition for open Digital



Humanities tools and practices” (del Rio Riande, Tóth-Czifra, Wuttke and Moranville, 2020:3). In both cases, the embedment of multilingualism within the scholarly communication ecosystem seems to promote equity among scholars in and from different linguistic and cultural communities and allow for “raising awareness to the value of language diversity by enhancing visibility and recognition of languages and cultures other than English, thus weakening the hegemonic position of English” (del Rio Riande, Tóth-Czifra, Wuttke and Moranville, 2020:6).

Authors as del Rio Riande, Tóth-Czifra, Wuttke and Moranville (2020), as well as Koutsomitropoulos (2019), identify some traits/ functionalities/ tools that should be considered in the design and management of search platforms, including:

- the selection of a working team according to the “fields of expertise and language skills as one aim is to be able to cover as a group the selection and curation of relevant content in various languages” (del Rio Riande, Tóth-Czifra, Wuttke and Moranville, 2020:3);
- the creation of a multilingual landing page, thus, democratizing access to knowledge in the context of arts of humanities research/scholarship usually grounded in regional, national and language-specific communities” (del Rio Riande, Tóth-Czifra, Wuttke and Moranville, 2020:9);
- the possibility of filtering and searching for information, using “several content enrichment and categorization efforts” (del Rio Riande, Tóth-Czifra, Wuttke and Moranville, 2020:2) undertaken by content curators;
- the addition of a short English introduction, as well as a link to the original research, in the visualisation of search results, to enhance research outreach of non-English publications;
- the possibility to expand search using different criteria (author, language, topic, etc.);
- the use of automatic annotation, once it minimises the workload.

As to this last matter, Ibrahim, Fathalla, Yazdi, Lehmann and Jabeen (2019:216) also refer that most studies in ontology enrichment focus work upon English ontologies from English sources only; nevertheless, the authors sustain that there are advantages in focus on Ontology Enrichment using Cross-lingual



Matching (OECM), i.e. the use of “multilingual ontologies, where a class label is presented by several natural languages, from monolingual ones. Such approach supports the ontology matching process with multiple translations for a class label in order to enhance the matching results”. Consequently, the possibility of having “semi-automated approach to enrich ontologies from multilingual text or from other ontologies in different natural languages” (Ibrahim, Fathalla, Yazdi, Lehmann and Jabeen, 2019:216) should be considered.

The possibility of having reliable multilingual research information will definitely contribute to an efficient dissemination of research (and research data) produced in national languages, as well as communication among publishers and researchers – thus promoting the development of intercultural, comparative and/or complementary studies in SSH. In this context, content curation is crucial and very sensitive, once it can determine the way researchers interact with other works/authors, namely in terms of peer discussion, co-construction of new knowledge and integration of research outcomes (Koutsomitropoulos, 2019).

### Reputation

Regarding the category *Reputation* (see 3., Table 1), Luzón (2019:39) defines scholarly reputation as “the expert appraisal of a scholar’s standing in their collegial reference group, which is collectively determined on the basis of their research achievements”, which is closely interconnected with research visibility. In the scope of SSH, European researchers have a long tradition in resorting to national publishers, especially because their primary target audience is also national, which, to a certain extent, may pose as “a barrier in global scientific communication” (Jokić, Mervar and Mateljan, 2019:1007). The type of publication also appears to assume relevance, in particular because bibliographic and citation databases (e.g. Web of Science and Scopus) primarily address articles and journals, and books were just recently included, but still give clear preference to those written in English (Jokić, Mervar and Mateljan, 2019:1006).

Given that a considerable amount of SSH research is published as monographs and/or in local language, the use of these databases to evaluate research and establish the researchers’ or the institutions reputations is necessarily fallacious. In fact, Kulczycki *et al.* (2020) have shown that Web of Science and Scopus covered only 25.0% and 30.7%, respectively, of the 164,218 peer-reviewed journal articles published in 2013-2015 by SSH investigators from seven European countries (Czech Republic, Denmark, Finland, Flanders [Belgium], Norway, Poland, and Slovenia). Considering that this hurdle



precludes the citation analysis of books not written in English, in their study about role of prestigious international versus local/regional publishers in Eastern and Central Europe, Jokić, Mervar and Mateljan (2019:1006) conclude that in SSH “data sources such as Google Books and Google Scholar are, in comparison with the above-mentioned commercial databases, valuable sources for book bibliometric analysis”. In addition, Kulczycki *et al.* (2020:2) refer that “neglected or undervalued research is less likely, then, to fulfil its responsibilities toward society, or create localized impacts”. This stands out as an important issue that national publishers should address, especially because publication with rich cross-lingual ontologies and automatic algorithms potentially has a positive impact on the researchers’, the institutions’ and the publishers’ own reputation.

As Luzón (2019:38) underlines, “universities and funding agencies expect groups to produce research outcomes with social and international impact and to engage in outreach activities”. In this scenario, publication in English language in international journals seems to be deeply linked to indexation and metrics, which are of growing importance for authors (Guns, Eykens and Engels, 2019). In fact, most scholars are being strongly advised to focus, essentially, on bibliometrics associated to their publications and, consequently, “publishing in non-hegemonic journals has almost become an activist statement” (Levitt and Crul (2018:45) cited in Balula and Leão, 2019:4). In this respect, Costello (2020:3) stresses that “‘foreign’ is uncritically equated with ‘prestige’”, which results in holding researchers hostage within oligopolistic [English] publishing, and Kulczycki *et al.* (2020:2) state that “researchers may choose to move away from locally relevant research toward decontextualized approaches of interest to English-language audiences”.

Still regarding the language of publication, Kulczycki *et al.* (2020:13) argue that “one way of making research results published in English more readily accessible to citizens is to publish the same results in a local language, but in a more popularized format, for instance, via a blog or alternative news source. This practice could, however, be at odds with current regulations concerning self-plagiarism”. In facts, it seems that the European Code of Conduct for Research Integrity (ALLEA, 2017:8) ranks “re-publishing substantive parts of one’s own earlier publications, including translations, without duly acknowledging or citing the original (‘self-plagiarism’)”, under the label “research misconduct and other unacceptable practices”. Nevertheless, considering that multilingual publishing may help local research become more relevant with the potential added-value of reflecting in bibliometric indicators, Kulczycki *et al.* (2020:13) argue that “an international discussion is needed to determine more



clearly how this type of publication strategy could be seen as beneficial, rather than a violation of research integrity and publication counting”. This approach could result in a win-win balance between international excellence and local significance of research outputs.

### **Balanced multilingualism**

The concept of *Balanced multilingualism* was assumed for the final analysis category (see 4., Table 1) because of the comprehensive view it offers over the multifaceted dynamics generated by language use within scholarly communication. Balula and Leão (2019) recall the definition proposed by Sivertsen (2018:2), which is understood as a dynamic approach, encompassing “all the communication purposes in all different areas of research, and all the languages needed to fulfil these purposes, in a holistic manner without exclusions or priorities. Balanced multilingualism is also to establish instruments for documenting and measuring the use of language for all the different purposes in research, thereby providing the basis for the monitoring of further globalization of research in a more responsible direction”.

Authors such as Guns, Eykens and Engels (2019), as well as Kulczycki *et al.* (2020), consider that one of the goals should be finding solutions grounded in sustainable balanced multilingualism, supported by well-designed and robust digital infrastructures for local language publishing. Hence, it seems reasonable to conclude that, even though Englishisation has a key role in facilitating scholarly communication, “the use of English should not be seen as a sole linguistic option, since the need for communicating in a *lingua franca* does not necessarily imply the adoption of a *lingua unica* – being the combination of balanced multilingualism and bibliodiversity foreseen as a much more fertile approach, in cultural, identitarian, and even in economic terms” (Balula and Leão, 2019:8).

In conclusion, balanced multilingualism in scholarly communication seems to pose as a golden breakthrough to embrace information-sharing, collaborative knowledge construction and equity by enabling global interaction with multinational and multidisciplinary research (and researchers); thus, mitigating the hurdles underlying static, poor translations and bridging research worldwide (Balula and Leão, 2019:7).

## Phase 2

**Phase 2: Survey regarding “Multilingualism in Scholarly Communication” [corresponding to goals 2 and 3]** Initial draft of the survey analysis by Lorena Caliman, with the collaboration of Carlos Costa, Mickael Silva and Maria João Padez, and final revision by Ana Balula and Delfim Leão. This Phase 2 of the report was revised within the OPERAS Special Interest Group on Multilingualism.

### Introduction and purposes

This report presents the main results and an in-depth analysis of a survey on Multilingualism in Social Sciences and Humanities that was conducted during the Summer of 2020 (from the 19th June to the 20th August), in an online survey distributed among researchers, translators and publishers in the OPERAS network and other channels. The survey was responded by a total of 359 participants who had a common initial set of questions, and then could give their contribution according to three different perspectives (Researchers, Translators and Publishers), separated or combined, depending on the respondents' profiles.

Both the survey and this report are part of the OPERAS-P Task “Innovative models of bibliodiversity in scholarly publications”. In the beginning of 2020, a literature review on multilingualism and scholarly communication was undertaken (see above Phase 1), which served, among other purposes, to find evidence of potentialities and shortcomings of multilingualism in scholarly communication, helping to establish grounds for the design and implementation of a survey addressing researchers, translators and publishers. Following the first step of the literature review, the empirical survey paved the way to two main objectives: to collect evidence as to the role of multilingualism within bibliodiversity in scholarly communication; and to contribute to the conceptual design of a platform prototype for community-owned translation services at the scholarly communication level, involving the needs of publishers, translators and researchers.

From the total of 359 respondents, 228 answered the survey as researchers, 35 as translators and 56 as publishers. These numbers represent the total number of respondents in each perspective, including participants who answered to more than one. A number of 79 people did not finish the survey, and therefore were considered as invalid responses. Other 2 respondents were excluded for giving unrelated answers to the questions, thus totalizing 81 invalid answers.

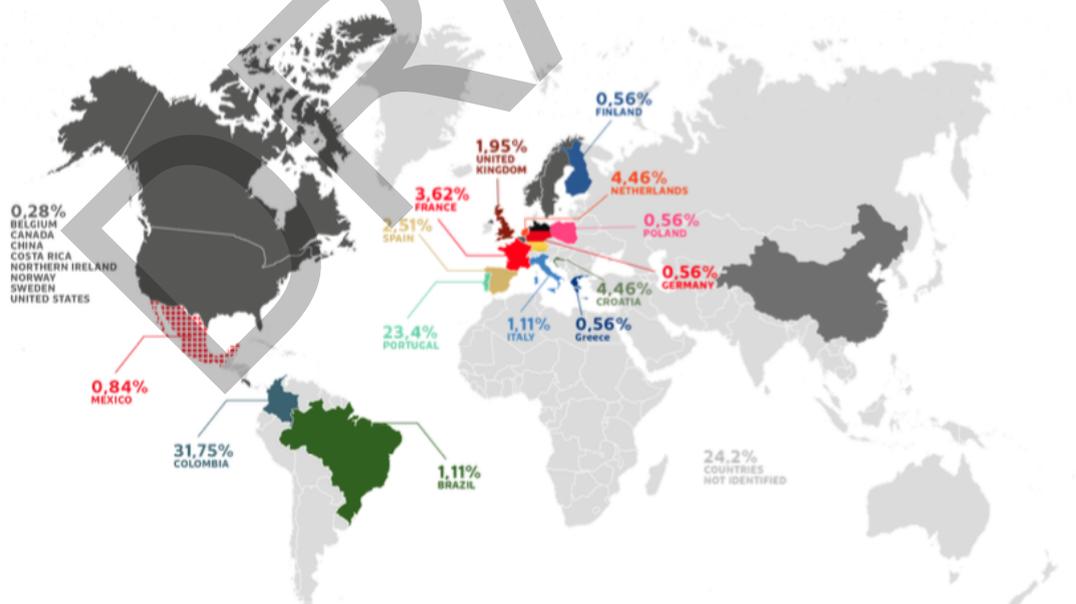


From the total of responses, 190 people responded exclusively as researchers, 10 exclusively as translators and 32 exclusively as publishers. The number of combinations between two types of answers were 21 people responding as both researcher and translator, 23 as researcher and publisher, and 8 as translator and publisher. Moreover, 6 respondents answered the questions posed in the three perspectives.

### Characterization of the sample

The respondents were distributed mainly by Europe and South America, with Portugal leading the number of answers, with 23.4% of the total (N=359). Croatia (4.46%), the Netherlands (4.46%) and France (3.62%) were the other most representative countries in Europe. However, a large number of respondents were from Colombia (31.75%), which resulted from a strong snowballing strategy led by researchers in that country. Brazil (1.1%) and Mexico (0.84%) were also represented. The overview of the countries represented can be seen in Figure 1 below.

Figure 1 – Distribution of the survey respondents per country



Source: Own creation

The results of the data analysis start by reinforcing the prevalence of English as the most common working language in scholarly publications (32.2% of the valid answers [N=601]). Portuguese (16.7%) and French (12.8%) are the

second most representative among the nine given language options (i.e. German, Greek, Italian, Croatian, Dutch, Polish, French, English, and Portuguese – the languages used by OPERAS core group members). However, 25.9% (n=155) of respondents selected the option “Other”, which can be explained by the absence of Spanish among the options offered in the question. Because of the great participation of Latin American respondents, mainly from Colombia, most of the answers in the “Other” specified Spanish. Other languages that were indicated were Latin, Ancient Greek, Chinese and Korean, but with no more than two answers each. The overall results regarding the working languages of respondents can be seen in Figure 2.

Figure 2 – Working languages

## What are your working languages?



Source: Own creation

## Results

### *The researchers' perspective*

#### *a) General overview regarding multilingualism*

In the researchers' profile, participants who answered about their position (N=202) divided into a share of 43.0% (n=87) of teachers, 22.8% (n=46) of senior researchers and 14.4% (n=29) of junior researchers. Besides, 7.9% (n=16) were PhD Students and “Other” non-specified profiles represented



11.9% (n=24) of the valid answers. The share of respondents that selected the option “Other” included, for instance, directors of library services, librarians and archivists, retired professors, independent and unemployed researchers and a writer. Given the nature of the OPERAS-P project, the survey was mainly focused on professionals acting in the area of Social Sciences and Humanities; nonetheless, 8.1% (n=24) of respondents who have identified their research areas (N=294) reported being from other scientific fields, i.e. Sport Sciences, Pharmacology, Medicine, Engineering, Biological Sciences and Geology. With regard to the overall sample of participants (N=294 valid answers), the areas with more representation included professionals from Education (n=33, 11.2%), History (n=20, 6.8%), Business Administration (n=17, 5.8%), and Literature (n=17, 5.8%).

As to the type of scholarly publications, the most selected among the answers (N=433) for the last 5 years were: i) journal articles (n=134, 30.9%), followed by book chapters (n=104, 24.0%) and papers in conference proceedings (n=82, 19.0%). Edited or co-edited books were selected by 15.7% (n=68) of the answers and monographs represented a total of 10.4% (n=45). Considering the scholarly production as a whole, the questionnaire turned to the languages in which the responding researchers usually publish their work. In this respect, English was highlighted in 37.0% (n=97) of the answers (N=262), followed by Portuguese (19.0%, n=50) and French (7.3%, n=19). However, 28.2% (n=74) selected the option “Others”, in which Spanish was specified, except for one case (1.35%, n=1 from the "Other" sample) indicating Russian. Completing the full sample (N=262) from this answer, Croatian (3.8%, n=10) and German (3.1%, n=8) were also selected, along with Italian (0.8%, n=2), Greek and Polish (both with 0.4%, n=1 each).

When asked about which languages their research results had been translated into, 33.9% (n=65) of the respondents (from a total of 192 valid answers in the question) reported none, indicating that translation of their works is not a practice for slightly more than a third of the participants. However, among the other 66.1% of the answers, 32.8% (n=63) had their works translated into English, 6.3% (n=12) into Portuguese and 5.2% (n=10) into French. Around 13% (n=25) of the respondents had their research results translated into languages not listed in the options – Spanish being the most recurrent, followed by Catalan, Russian, Slovenian and Czech. Finally, from the provided list of languages, German and Italian (both with 3.6%, n=7 each) and Croatian (1.6%, n=3) were also nominated.

Getting into the specific questions related to the views and reflections on multilingualism in scholarly communication, the questionnaire then included



open questions, starting with a focus on the advantages and disadvantages of publishing in the respondents' mother tongue and in English. The questions about the advantages and disadvantages had an average of 145 answers each; after excluding invalid answers (non-related to the questions) and categorizing the commentaries - once each answer could have more than one argument -, the numbers were slightly changed, as it will be indicated in the following paragraphs. The first one asked about advantages of publishing in the mother tongue (N=158 categorized answers). Here, the perception of a possible bigger and better outreach to readers - both academic and non-academic - was the most cited advantage of publishing in the mother tongue (29.7% of answers, n=47).

The easiness to write in the mother tongue (17.8%, n=28) and the possibility of more fluent, better and richer expression of ideas (11.4%, n=18) were the other types of most frequent answers. Besides, being capable of a more accurate expression in the mother tongue (8.2%, n=18), and being able to write faster (7.6%, n=12) were also indicated. Issues related to peer recognition were cited 5.7% (n=9) of the times; ease to publish and the identity/language promotion/preservation were both cited on 2.5% (n=4) of the times each. Other answers received less than 2% of indications, totalizing 14.6% (n=23). Among them, there were comments such as positioning and relevance, impact, access to researchers, linguistic diversity, not having extra costs, writing in a field where the language itself was the topic; "no" advantage and having English as a mother tongue were also noted.

It is noteworthy that the perception about being more read in their native language was the most cited answer. In some regions (specially in the context of Iberoamerica), the publication in Spanish and Portuguese can have a broader reach, according to answers such as respondent no. 276: "Hispanic community is able to understand my scientific articles"; and from respondent no. 123 that reported that his/her work is "more accessible to Portuguese students". Another aspect of answers included in the category of better outreach to a wider public/audience was the intention to be read by a non-academic audience, as is the case with respondent no. 88: "to reach a non academic public".

On the other hand, the disadvantages of publishing in the mother tongue (N= 136 categorized answers) included mainly the limited potential for international outreach (29.4%, n=40) and the perceived limitation in the number of readers in comparison with publishing in English (28.7%, n=39). Commentaries about low capacity of dissemination and diffusion of the works represented 10.3% (n=14) of the answers; a similar percentage appeared for the non-acceptance or limitation of options for publishing in journals (9.6%, n=13). Some specific



issues were mentioned among the commentaries about the limitation of diffusion, such as the difficulty to find colleagues to read their works when written in their mother tongue, as in respondent no. 235: "I can't transfer and exchange papers to the colleagues who do not know Croatian". The undervaluation of the mother tongues were cited as well, in 4.4% (n=6) of the times, as well as difficulties with academic jargon in the mother tongue and other translation problems (2.9%, n=4). A 4.4% (n=6) of the valid commentaries said that there were no disadvantages in writing in their mother tongue. There were also some considerations of disadvantages on performance evaluation (4.4%, n=6) and on the impact of the work (5.9%, n=8).

In what regards the publication in English, the main advantages mentioned by the participants (N=153) included the possibility of more - and international - dissemination and impact of research (27.45%, n=42), the potential of reaching more readers (26.8%, n=41), the peer recognition and the access to more prestigious journals (13%, n=20 for both categories, 6.5% each). The fact of perceiving English as the dominant language was cited 5.3% of times (n=8). The respondents also included comments about better performance evaluation (2.6%, n=4), clarity and standardization (n=2 each, 2.6% both). Other single comments (n=1, 0.65% each) included considerations about having direct contact with international researchers, greater competitiveness and the possibility of developing terminology; these single comments, on the other hand, included as well answers that could be interpreted as disadvantages, such as a greater difficulty to write and the need for checking available tools of translation. One respondent (0.65%) indicated that English was his/her mother language.

In respect to the perceived disadvantages of publishing in English (N=153), the bigger difficulty to write (20.3%, n=31) was the main point. Besides, less accessibility of readers was also significant (11.1% of answers, n=17), many times in relation to the regional context, such as in answer no. 388: "There are fewer readers of languages other than Spanish" and no. 368: "A high percentage of the student population - (include high school and undergraduate) [in] Latin America does not read English". The loss of nuances of meaning in translation were cited 10.5% of the times (n=16). "No disadvantages" were pointed 9.2% of times (n=14), followed by the idea of a loss in diversity and in the socio-cultural context (7.8%, n=12). The high costs of revisions and translations were pointed out on 7.2% of the answers (n=11). A bigger time consumption for the production of texts was cited 6.5% of times (n=10), followed closely by perceptions of a bigger difficulty for publishing, given the reviewers demands for authors whose native language is not English (5.9%, n=9). A smaller impact, especially in the regional context, was cited 5.3% of



times (n=8). Other categories of answers included: not being their native language, more hard to take control of the quality, a negative opinion about publishing in English or about the language itself, "all" non-specified disadvantages, the lack of adequate reviewers, English being their native language, more competition and direct relation of the research topic with other language. These other categories totalized 16,2% of the answers (n=25).

The survey also questioned, in an open answer, whether the respondents had ever faced any sort of prejudice or criticism for using their mother tongue (other than English) in their scientific publications. As to this, more than half of the respondents (N= 129 of valid answers) answered negatively (53.5%, n=69), while 27,1% (n=35) answered that they have suffered some kind of prejudice, whether frequently or eventually. Other 19.4% (n=25) of the responses didn't give a definite answer, making, instead, commentaries about the subject, such as in respondent no. 368: "Not exactly prejudice or critics, but journals with high impact receive only manuscripts in English, including some journals in Colombia and Brazil". Another example, from respondent no. 310, shows a consideration of English as the dominant language, but does not show if he or she has ever suffered some kind of prejudice: "Publishing in English is more prestigious. My university gives more economic incentives when publishing in English".

Among the set of commentaries of prejudice, the respondents considered that they: i) had their works either undervalued when written in their native tongue, ii) faced a limited understanding and usage of the data, or even iii) felt the pressure for using English as a standard. It is noteworthy that, among the sample that reported not having faced prejudice, 2 respondents (1,5%) said that they didn't because they never wrote in their mother tongue (respondents no. 33 and 136). Other 6 people (4.6%, among which there were respondents that answered both "yes" or didn't give a definite answer) pointed out some subjective prejudice, such as: "Sí, se asume que aquellos investigadores que publican en inglés no solo tienen más 'prestigio' sino también credibilidad ante la comunidad de investigadores. Esto es completamente errado" (no. 388). Feelings of prejudice and criticism were also reported in more informal situations, such as conversations with fellow researchers. Career evaluation was also mentioned (2.3% of times, n=3), in the sense that publications were deemed having less value when they were not in English. Undervaluing of research and researchers whose language is not English was cited on 14% of the answers (n=18), such as in: "I have not had any inconvenience so far, although there is a tendency to stereotype that, for example, in Latin America it is not investigated with the same rigor as in other places" (respondent no. 277).

Finally, the encouraging/pressure to publish in English was cited on 6,2% (n=8) answers. Not all answers in this question included commentaries.

Among the people who responded negatively to this question – meaning that no prejudice or critics were ever felt –, the prestige of publications in their own mother language was once reported (0.7%): “No, as my mother tongue is Portuguese, with a large audience and prestige in Brazil” (respondent no. 88) - a scenario that is linked to some answers in the above mentioned questions in what respects the potential and relevance of regional contexts. Overall, the negative answers were rarely commented, differently from the positive answers, which had a set of different considerations, as seen above.

*b) About a Multilingualism platform*

Following the questions about publications and languages, the survey directly addressed a question regarding how the researchers felt about the possibility of having translations made in collaborative networks into the languages of the group members. The question asked whether such translations would be able to promote more inclusion, and a sense of value and of being significant for every member. For 81.4% (n=105) of the participants (N=129), the answer to this question was positive; for 6.2% (n=8), it was negative, and 12.4% (n=16) gave other kinds of response, either saying that they don't know, considering some aspects of the idea without giving a definite answer, providing a reflection on the theme or even some advice on what they think it would be the best proposal. There has been a total of 47 commented answers (which provided more content than just answering "yes", "no" or "I don't know" and adding reflections to the question), which could have more than one category of commentary inside of them, thus totalizing 56 commentaries. The most common comments focused on multilingualism as a way of producing new knowledge (17.9%, n=10), followed by comments that saw the proposed collaborative networks as promoters of more interculturality and interdisciplinarity (12.5%, n=7). Other frequent answers included different proposals for fostering collaboration in the networks (10.7%, n=6), reiteration of the perception that the collaborative works could wider the audience of readers (10.7%, n=6), with arguments taking into account the amount of time and resources invested for making translations (7.1%, n=4), and the defense of a bigger effort among researchers to learn more languages and to respect all languages as equally important (7.1%, n=4).

With less representation, but having three comments each (5.4% each category, 16.2% in total) there were: the affirmation that having English as the common language was fine; that the success of the collaborative networks



depended on how they would be implemented; and the necessity of paying more attention to revision and quality control during the translations. Other 17.8% (n=10) of the answers were concentrated on the following categories (each one with one or two comments each): advocating that the translation of abstracts was enough for the promotion of more inclusive groups; the perception that translations are often flawed and that the proposal of a multilingual collaborative network would have a low efficiency; affirmations of not believing in the possibility of true multilingualism in European networks; and appeals to considerations about the social impact of such networks.

Following this first approach to the possibility of a network-based collaboration for translation, respondents were asked to think about a platform that would bring together researchers, translators and publishers, in which participants could assume more than one of these roles. Having this in mind, they were asked whether they saw advantages in such a platform. In this matter, 83.3% (n=130) of the participants with valid answers (N=156) agreed, 3.8% (n=6) disagreed and 12.9% (n=20) did not answer. When asked whether they would be willing to participate by translating or contributing to translate their colleague's work into languages in which they are proficient, 65.4% (n=102 of 156 valid answers) answered positively, 14.1% (n=22) negatively and 20.5% (n=32) did not answer. The former were then asked to indicate in which languages they were proficient and the results (N=220) were English (cited 29.5% of times, n=65), Spanish (22.7%, n=50), Portuguese (17.7%, n=39) and French (15.9%, n=35). German, Italian and Croatian were selected by 8 participants each (3.6% each, 10.8% the three languages). Ancient Greek, Latin, Polish, Romanian, Latvian, Russian and Chinese were selected by one participant each (3.2%, n=7).

The remaining questions answered by researchers were designed to measure the degree to which they preferred to publish their works with appropriate CC open licenses – 74.4% (116 of N=156) stated that they favor open license publications, 8.3% (n=13) answered that they do not and 17.3% (n=27) did not answer –, as well as other issues regarding the possibilities of collaborative translations. One of the questions focused on whether respondents would prefer to translate general or more specialized topics. The majority of the respondents pointed that they would prefer to translate specialized topics (64.7%, n=110 of N=170), while 35.3% (n=60) would choose to work with general topics. Among the specific topics, a wide range of themes was identified, including not only the Social Sciences and Humanities but also other Sciences, such as Pharmacy, Medicine, Sports Sciences, Energy, Environment, Tropical Ecology, Business and Entrepreneurship.



Participants were then asked about whether they would consider adopting a recommended writing style (disambiguation, simplified sentences, lexical consistency) in order to avoid typical machine translation errors, if the platform incorporated machine translation tools. Around 68.6% (n=107) of the participants (N=156) would consider using the recommended style, 9.6% (n=15) would not and 21.8% (n=34) did not answer.

Considering the possibility that researchers would have to choose their own publications to be translated into other languages, participants were asked what criteria they would consider in the selection process. Once it was an open-ended question, it gave way to a wide range of answers (total of codified answers N=151/original N=128).

Questions respecting a good relationship with the translator, translation quality and the need of translators that are specialized on the research topics represented 20% of the answers (n=30). The relevance of research results (10.6%, n=16), the will to reach a wider public - on academia or outside of it - (8.6%, n=13) and the opportunities regarding the research topics, such as the impact and global relevance (also 8.6%, n=13) were the next most cited arguments.

Quality of the work and the potential interest of the topic/research among research networks were both cited 5.3% of times (n=8) each, thus totalizing a 10.6% share of the comments. The choice for more recent and original works was cited on 4.6% (n=7) of the times, as well as questions regarding rigour, methodology and clarity (4.6%, n=7), totalizing other 9.2% of answers in this question. Less frequent answers involved: the reputation of the journal or the publisher with which the work would be translated and published (3.3%, n=5), the choice to translate in languages related to the topic or the social relevance of the language to be used (3.3%, n=5), the universal or wide scope of the topic (2.6%, n=4), the will to promote some language or culture (2.6%, n=4) and some particular interest in a given topic (2.6%, n=4). Other reminiscent answers talked about costs, the aim of translating all works, the place where the work was first published, the availability of open access, the choice for languages of topic related experts, personal relations, committees of experts, the use of the same criteria used for having a work accepted for publication, the ability to add value to the publication and to collaborate in a respectful way. All of these last comments were cited either 1 (0.7%) or two times (1.3%), representing a share of 8% (n=12). Finally, unrelated or invalid answers, or reports of not having understood the questions were pointed out 15 times (10%).

Still considering the situation in which researchers had to choose their own works to be translated into other languages, they were asked to indicate which foreign languages they would consider for translating their works into. The most chosen language (N=226 language indications) would be English, with 26.5% (n=60) of the answers. English was followed by French (15.5%, n=35), Spanish (12.8%, n=29), Chinese (15.5%, n=22), German (8.4%, n=19), Portuguese (6.2%, n=14) and Italian (5.3%, n=12). A number of 12 respondents (5.3%) didn't provide an answer about the other languages they would choose for translating their own works. Other 7 people (3.1%) commented that they would choose any language, "other" languages in general or any languages with interested readers. Russian and Arabic were also cited, with 1.8% (n=4) and 1.3% (n=3) of the choices, respectively. Aramaic, Croatian, "Asian languages", Korean, Scandinavian languages, Japanese, Latvian, Indian and Dutch were each chosen once (0.4% each), totalizing 4% of choices (n=9).

Because of the open nature of the question, part of the respondents also made comments about their language choices (N=90). In this sense, the most common comment (36.7%, n=33) regarded the relative universality or wide coverage of the languages chosen - in this case, mainly English, but also Spanish, Chinese, Portuguese, and then German, Italian and French were also related to these comments. Another part (17.8%, n=16) of the respondents mentioned that their choice would depend on the topic and area of expertise. There was the same number (17.8%, n=16) of answers talking about a wider academic reach related to their choices, while the perception of increased readability was pointed out on 8.9% of the answers (n=8). Four respondents (4.4%) took into consideration their personal taste for the language and/or culture chosen, while 3.3% (n=3) would choose languages in which they were used to read or in which they felt fluent enough to control the quality of translation. In the opposite view, one respondent (1.1%) said that he/she would choose to translate into languages in which he/she is not proficient. The fight against monolingualism in some countries was directly pointed out in 3 (3.3%) answers, while a personal will to interact with other research networks and to cooperate was cited twice (2.2%). Other single comments included topics such as to ensure continuity of scientific languages - related mainly with the so-called "peripheral" languages (1.1%) - and the idea that the more languages, the better (1.1%). One person said that he/she doesn't need translations (1.1%), and a last one said that he/she would choose a language because it is very different from his/her mother tongue (1.1%).

The two last questions addressing researchers asked if they considered that the number of translations in different languages could improve visibility,



accessibility and readability of their work – which could result in an increase in the number of visits, downloads, mentions and sharing. In this case, 84.0% (n=131) of the participants (N=156) agreed with the claim, 3.8% (n=6) disagreed and 12.2% (n=19) did not answer. The last question focused on altmetrics, in which participants were asked if the number of languages into which a work is translated could be seen as alternative metrics to measure its usability and impact: 69.9% (n=109 of N=156) agreed with this premise, 14.7% (n=23) disagreed and 15.4% (n=24) did not answer.

### *The translators' perspective*

In the section of the survey dedicated to collecting the translator's perspective regarding a collaborative multilingual platform for bibliodiversity, a total of 35 answers were gathered. Within this group, as seen in the first section of this report, 6 (17.1%) people responded as researcher, translator and publisher, 21 (60%) as researcher and translator, and 8 (22.9%) as both translator and publisher. It is important to remember, once again, that the questionnaire allowed for respondents not to answer all questions; thus, each question could have a variable number of respondents, as seen in the researchers section.

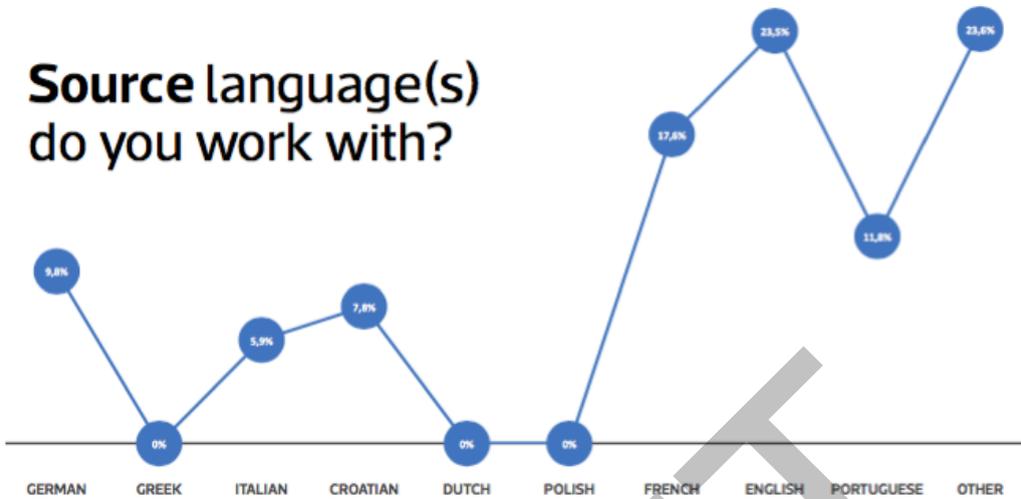
Considering the identification of the translators, more than a half identified themselves as professional translators (55.2%, n=16 from 29 answers). Other 24.1% (n=7) identified themselves as experts in their scientific areas; 13.8% (n=4) were non-professional translators and 6.9% (n=2) selected "other" non-listed options. The range of source and target languages with which the translators worked is detailed in Figures 3 and 4 below. In Figure 3, it is possible to see that English (24%, n=12), French (18%, n=9) and Portuguese (12%, n=6) were the most prominent source languages. Details from the provided open answers in the option "Other", 10% (n=5) included Spanish, 6% (n=3) for Latin, 4% (n=2) Ancient Greek and 2% (n=1) Chinese.

Figure 4 shows that English (25%, n=9) is the main target language, followed by French (16.7%, n=6), Portuguese and Croatian (both with 13.9%, n=5 each), German (11.1%, n=4) and Italian (2.8%, n=1). Detailed from the "other" option, there was Spanish (11.1%, n=4), Ancient Greek (2.8%, n=1) and Latin (2.8%, n=1).

Figure 3 – Source languages for translators



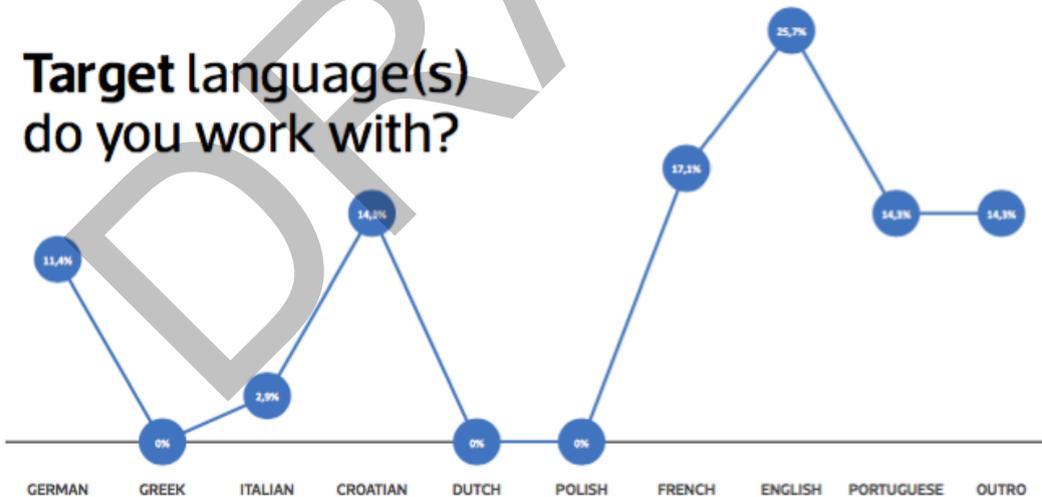
### Source language(s) do you work with?



Source: Authors elaboration

Figure 4 – Target languages for translators

### Target language(s) do you work with?



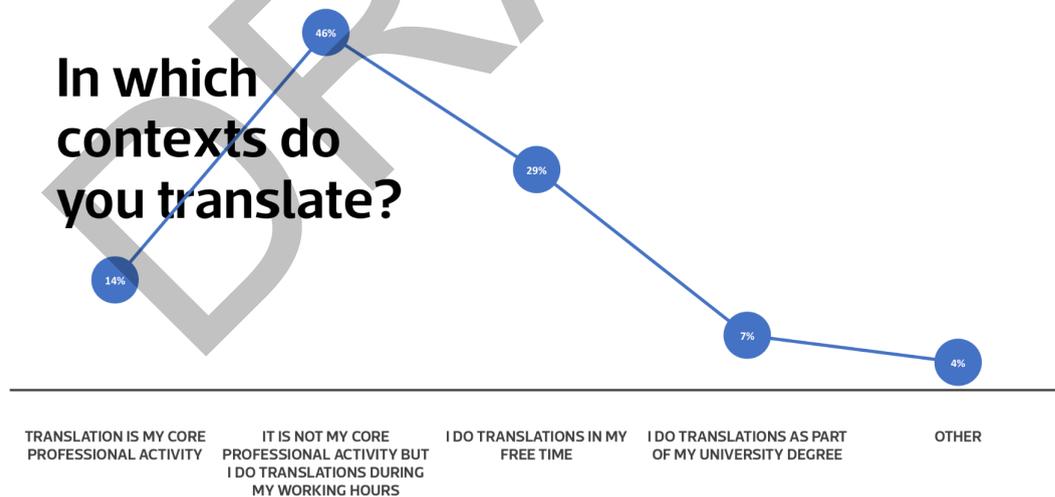
Source: Authors elaboration

In what concerns the translators' work experience, 81% (n=17) of the respondents (N=21) reported having more than 5 years of experience, while 14% (n=3) had 1-5 years and 5% (n=1) no previous experience in the field. As

for working with Computer Assisted Translation (CAT) tools, 52% (n=12) of the participants (N=23) reported using a specific CAT tool, 39% (n=9) do not use them and 9% (n=2) did not answer. Among the tools used, 38% (n=8) reported using desktop tools (such as SDL Trados, memoQ, OmegaT etc); 29% (n=6) online tools (e.g., Matecat, Wordfast Anywhere etc.); and 24% (n=5) automatic translation with post-editing. Other 10% (n=2) did not answer. Furthermore, 67% (n=14) of respondents indicated that they did not have specific training to use this type of tools but were willing to, while 33% (n=7) have already received training to use CAT tools.

Almost half of the respondents (46%, n=13) reported that translation is not their core professional activity, but that they do translations during their working hours; 29% (n=8) translate in their free time and only 14% (n=4) mention translation as their core professional activity. For 7% (n=2 from 28 responses), the translation activity is part of their university degree and 4% (n=1) chose "Other" option. A graphical representation of these findings are presented in Figure 5.

Figure 5 – Translation activity



Source: Authors elaboration

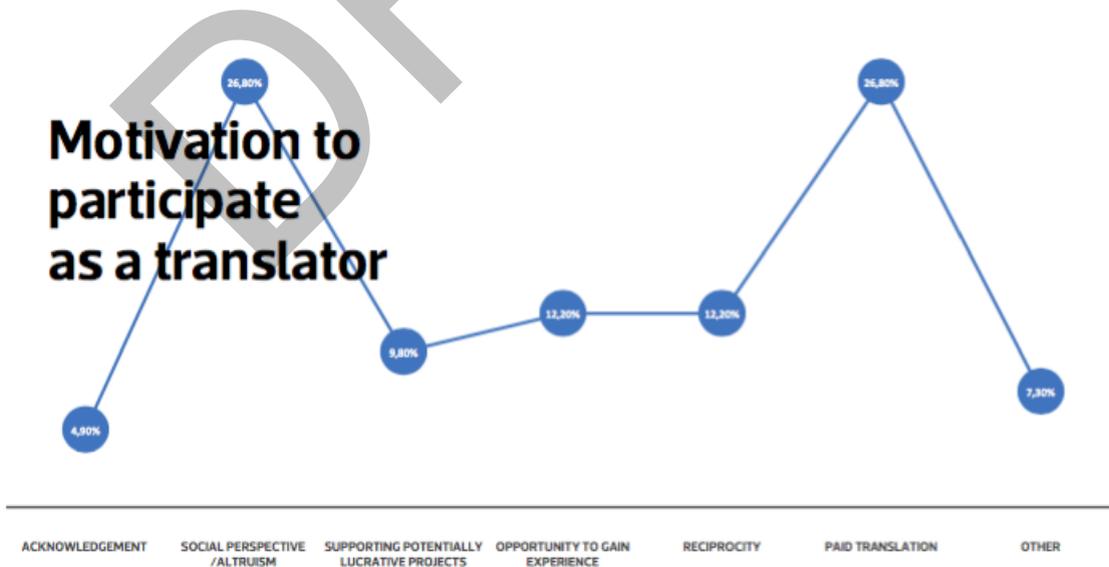
Moving forward to the motivation for participating in a collaborative multilingual translation platform, such as the one proposed in the study, the results of the



data analysis show a balance between people who wish to participate from a social perspective and altruism<sup>1</sup> (26.8%, n=11) and people who consider participating to get paid for their translations (26.8%, n=11). The opportunity to gain experience and the reciprocity – being able to receive the same kind of collaboration in return – were both present in 12.2% (n=5) of the answers, totalizing 24.4% of answers. Supporting potentially lucrative projects resulting from voluntary participation was also mentioned (9.8%, n=4), as well as participating in order to receive some kind of acknowledgement (e.g. certification) (4.9%, n=2). Other non-listed, more dispersed options, were also retrieved (7.3%, n=3), as presented in Figure 6.

The contact/interaction level between translator and author (to ask/answer questions) was another topic addressed in the questionnaire. The respondents' answers distributed as follows: i) a relation of responsiveness (38.5%, n=10), ii) the permanent availability model (34.6%, n=9), iii), the contact only after receiving the work (15.4%, n=4), iv), and the contact prior to the translation (11.5%, n=3). As to the process of revision, the respondents' preference was quite balanced: 51.7% (n=15) of the respondents (N=29) indicate that the revision should be monolingual (made by an expert in the field or a technical proofreader) and 48.3% (n=14) considered that the revision should be bilingual, made by a translator or linguistic proofreader.

Figure 6 – reported motivations to participate in the platform as a translator



<sup>1</sup> This option also detailed that this perspective would mean dissemination of relevant content, like a researcher's work, theories, new paradigms etc.



Source: Authors elaboration

For 73.9% (n=17) of the participants (N=23), the choice of translating a work should be made by the translator, while 26.1% (n=6) consider that it is the publisher's choice. As to the extent and type of works to be translated, 73.1% (n=19) prefer to translate complete articles, books or monographs, while 26.9% (n=7) prefer to translate abstracts. A majority of the participants prefer to translate works in their own field of expertise (72.7%, n=16), while 27.3% (n=6) are willing to translate in any given field.

The frequency of the work was also taken into account. In regard to the translation of abstracts, the majority (57.1%, n=12 from 21 answers) would translate several abstracts per year; 28.6% (n=6) would do one translation per month and 14.3% (n=3) would prefer to do just one translation per year. Considering the translation of full articles, books or monographs, the majority (57.1%, n=12 from 21 answers) would choose to do several translations a year, 23.8% (n=5) would do one per year, 14.3% (n=3) would choose to do less than one a year, and 4.8% (n=1) would do one per month.

The time spent in the translation of abstracts (N=21) would be up to 3 days for 81% (n=17) of the respondents and up to one week for the other 19% (n=4). For a full article (average 8000 words), 81.8% (n=18 from 22 answers) admitted that it would usually take them up to one month to translate and 18.2% (n=4) up to three months. In terms of feedback regarding the translation work, most respondents expected it to be either from the author or the proofreader (44.1%, n=15 for each case), with less respondents expecting it from the target audience (11.8%, n=4 from 34 responses).

The ethical issue of confidentiality was also addressed in the survey. In this respect, for 82.6% (n=19) of the respondents (N=23), there is a need for a dedicated Code of Conduct for Translation. No more than 4.3% (only one respondent) disagreed with this need and 13.0% (n=3) did not answer about the issue.

In the final question, participants were asked to indicate when a translator should refuse a work. To answer this question, they were given a set of options and could select all that applied, which generated a number of 64 answers. They distributed as follows: i) when the translator is not up to the task (25%, n=16), ii), when the translator cannot meet a reasonable deadline (23,4%, n=15), iii) when the translator does not master the scientific or academic field (23.4%, n=15), or iv) when the translator does not master the academic jargon of the research area (26.6%, n=17).



### *The publishers' perspective*

This section aimed at establishing a comprehensive characterization of the publishers participating in the survey and getting to know their preferences and willingness to take part in the activities concerning the idealized platform. As already mentioned, 57 publishers answered the questionnaire. From the valid responses in the identification question (N=45), most publishers were research organisations (31.1%), followed by public organisations (24.4%) and commercial publishers (13.3%). Learned societies represented 8.9% of the respondents, and 22.2% classified themselves as "Other", mostly specifying as university presses (6.7%, n=3). The ones that were attached to public organisations (N=35) were mostly integrated into educational establishments (65.7%, n=23).

The financial resources of the participating publishers come mainly from public funding and own revenues (31.4% of the answers, n=16 each; 62.8% both), followed by funding from research projects (21.6%, n=11) and national funding (3.9%, n=2). Respondents also mentioned that most publishing decisions are made by the editorial boards (45.8%, n=22), while 22.9% (n=11) reported that the decisions encompass all the three options presented in the questionnaire (publishers, editorial boards and series/collection directors). Moreover, 12.5% (n=6) indicated that the decision is made by the collection/series director and 4.2% (n=2) selected the option "other". They also state that most of the peer-reviews correspond to double-blind (77.8%) or blind (22.2%) reviews. The majority of the participants (57.6% from 33 answers) published, on average per year between 2017 and 2019, 1-20 books, while 15.2% published between 20-50, and other 15.2% between 50-70 books. There are 9.1% who published between 70 and 100 books per year and 3.0% selected the option "other".

Regarding the types of books published, participants could select from the options provided all that applied, which generated 85 answers. The results yielded by the data analysis are the following: scholarly books (24.7%, n=21), essays (21.2%, n=18), technical books (18.8%, n=16), textbooks (17.6%, n=15), poetry (7.1%, n=6), fiction (4.7%, n=4), and theatre (1.2%, n=1). The option "others" was selected four times (4.7%).

With regard to the publication of journals in 2019, 62.5% (n=10) of 16 valid answers reported to have published between 1-5 journals, 25% (n=4) between 5-10, 6.3% (n=1) between 10-50, and 6.3% (n=1) published more than 50 journals. The areas covered by the journals ranged from all the options among Social Sciences and Humanities, but others, namely on Medicine, Mathematics Education, Veterinary and Advertising, were also indicated. A variety of areas

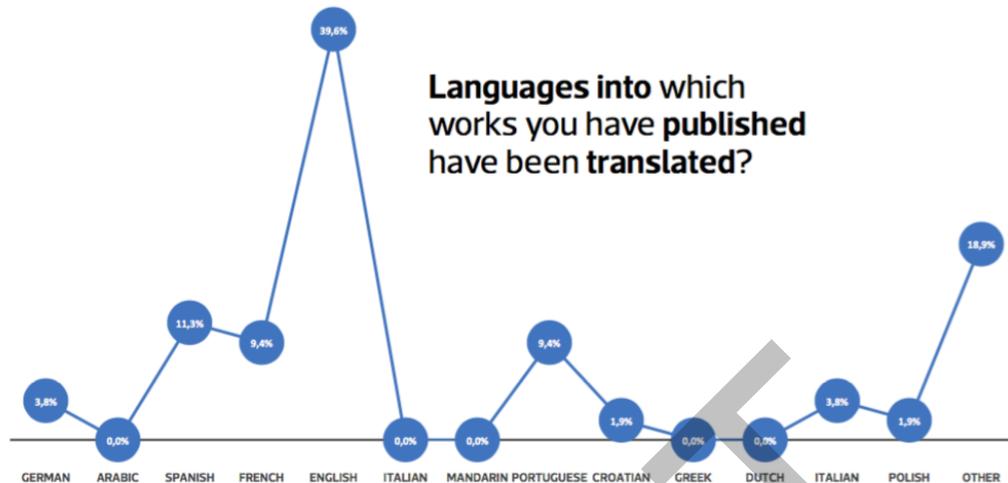


was also identified with regard to the books edited by the participants, with titles published in the different branches of Social Sciences and Humanities. With regard to the books, the areas with a larger number of titles published between 2017-2019 were Education (N=19, 8.6% of the 233 selected answers), History (N=15, appearing 6.4% of the times), Economics and Finance (N=14, 6%).

Most publishers (74.3% from 35 answers) informed that they allow for open access publication globally, while 17.1% do not and 8.6% did not answer the question. Considering the period from 2017 to 2019, the average of books published in open access annually was between 1-20 for 60.6% (n=20) of the respondents (N=33), between 20-50 for 24.2% (n=8), 50-70 for 6.1% (n=2), and 70-100 for 3% (n=1). About 6,1% (n=2) selected the option "other". Regarding the number of translations published annually in the same period of time, a vast majority (81.1% from 33 answers) reported translating between 1-20 works a year, 6.1% published 20-50 and 12.1% did not specify.

The languages most published by the respondents were English (mentioned on 30.7%, n=23 from the 75 answers), Spanish (n=22, 29.3% of the mentions) and Portuguese (n=12, 16%), but they also reported publishing in French (5.3% of times, n=4), German (N=3), Croatian and Italian (2.7%, n=2 each, 5.3% both), and Polish (1.3%, n=1). The languages into which the works published by the participants (N=52 answers) were translated included mainly English, Spanish, French and Portuguese, but also Italian, Polish and Croatian (see Figure 7). A significant majority of the respondents (71.4% from 35 answers) agreed that they would be willing to provide their translated works to develop multilingual corpora for training specialized machine translation engines. Nevertheless, 14.3% disagreed, and another 14.3% did not answer the question.

Figure 7 – Languages into which the published works have been translated



Source: Authors elaboration

Some other questions were also asked in order to better understand the translators' perceptions concerning, for example, the services they provided to authors, average number of books and journals printed, as well as the publishing management system tools used more recurrently. Even though it is not the focus of this report, it is pertinent to refer that the collected data helped to understand the institutional situation of the participants, which can in turn help to have a better understanding of the possibilities available for innovative models for fostering multilingualism and bibliodiversity in scholarly publications.

### Global contribution for future planning

The results presented in this survey showed that there is a strong openness among researchers, translators and publishers in viewing the amplification of multilingualism as an advantage both for fostering international collaborative works and for promoting interculturality, inclusion and equity. Among researchers, one of the strongest claims for until now dismissing the possibilities of translation would be the consumption of time, the price and the flaws in the process, issues that could be faced and solved through a collaborative system that could use expertise in specific areas to support and facilitate this kind of work. Also, the exchange of experiences and specificities among researchers from different languages, but sharing the same areas of study, could be a relevant contribution for enriching the international collaboration and impact of works. Considering that, as we have seen above, more than 35% of the researchers never translate their works, a platform or networking environment to put together interested researchers, translators and

publishers could improve the impact and the internationalization of works in a wide variety of areas, in the field of Social Sciences and Humanities.

It was possible to perceive that a significant part of the translators willing to participate would do this kind of work as a form of altruism and as a way of enhancing their expertise (particularly when it is combined with reciprocity), what is another indicative of a feasible positive impact from a network to be created with the purpose of improving bibliodiversity through multilingualism. Therefore, the work of colleagues from different origins and languages could be better valued. Also, the range of languages published by the editorial houses could be enriched in the process, widening their impact and opening opportunities for them to publish works from a wider scope of authors worldwide, a scenario that would in turn open up possibilities for readers – whether academic or from the general public – and a wider audience of researchers. This would improve, more generally, the scholarly communication landscape on an international scale, helping as well what tends usually to be considered ‘national’ (the usage of local languages) to become more clearly ‘international’ (by putting them on the radar of wider networks and on the scope of collaborative interest groups).

## Phase 3

**Phase 3: Contributions to developing the conceptual design of a platform pilot of a shared translation service at the scholarly communication level (involving publishers, translators, researchers)” [corresponding to goals 4 and 5]** Initial draft by Delfim Leão with the collaboration of the OPERAS Special Interest Group on Multilingualism, with contributions by Agnieszka Szulińska, Dulip Withanage, Elea Giménez Toledo, Jadranka Stojanovski, Jean-François Nomine and Susanna Fiorini.

A scenario that has become increasingly clear during the development of the different phases of this report is that multilingualism is perceived as a strong manifestation of bibliodiversity, particularly important in the area of Social Sciences and Humanities. This does not preclude the usage of English as a communication language, as long as the advantages of using a *lingua franca* does not imply the risk of turning it into a *lingua unica* in terms of scientific and scholarly communication. Instead of that, it is strongly advocated the implementation of innovative solutions that have the ability of enhancing balanced multilingualism in scholarly communication, in information-sharing, and in collaborative knowledge construction.



At the OPERAS level, this is being put into practice by the TRIPLE project (<https://www.operas-eu.org/projects/triple/>), by the creation of the OPERAS National Nodes (<https://www.operas-eu.org/projects/operas-ger/>), and more programmatically by the work of the Special Interest Group on Multilingualism (<https://www.operas-eu.org/special-interest-groups/multilingualism/>), whose members have directly participated in the making of his report.

In the case of TRIPLE, its main goal is to develop a full multilingual and multicultural solution for the appropriation of SSH resources. The discovery tool is based on a multilingual standardization vocabulary of concepts and authoritative labels translated and validated in an initial set of 9 languages (Croatian, English, German, Greek, French, Italian, Polish, Portuguese and Spanish), which will be enlarged into other languages in the future.

At a complimentary level, the implementation of National Nodes plays a pivotal role in establishing a connection point to outside the OPERAS community, helping to identify needs, as well as to provide services and training. Among them is the necessity of finding a balanced solution for the keeping of national languages as fully scientific valid modes of expression, while promoting them in the international level. This strategy is fully aligned with the development and implementation of TRIPLE solutions and may benefit directly from a service dedicated to supporting translations and stimulating cross-publications.

The Special Interest Group (SIG) for Multilingualism is directly involved in studying and promoting the development of the translation platform as one of the future OPERAS services. In fact, it is distinctive of each OPERAS SIGs to aggregate a group of expertise with a broad coverage of international representativeness, who work collaboratively and share information about other initiatives in order to prepare projects on their topic. Although the specifications of the translation platform are to be developed in a dedicated project, and therefore won't be approached here in detail, its conceptual framework is based in three key-pillars, which act as guiding development criteria:

- it shall be a community based platform: leveraging on the federative nature of OPERAS consortium, it will boost the collaborative work of researchers, translators and publishers, by creating conditions for cooperation and providing information that will enable each scholarly work to identify the appropriate publisher profile, the adequate scientific milieu and the right partnership, in order to disseminate specialized or local scientific production in a wider environment;

- the platform shall be conceived as a social infrastructure: by federating technical knowledge and scholarly expertise, it will stimulate the sharing of tools, methodologies and practices, so that a broad user community can test and scale what is being developed separately by individual partners;
- it must contribute to make national production become internationally relevant: the literature review showed that the notion of international publishing is closely linked to the idea of publishing in English in large international publishing houses; however, by putting a broad universe of small publishers and their authors in contact with each other, it will be possible to find an alternative way to internationalise scholarly production, to enhance specific catalogues and relaunch multilingualism as an expression of bibliodiversity, of inclusion and of scientific maturity.

## References

- ALLEA – All European Academies. (2017). *The European Code of Conduct for Research Integrity*. Berlin: ALLEA – All European Academies. Available at [https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics\\_code-of-conduct\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics_code-of-conduct_en.pdf)
- Balula, A., & Leão, D. (2019). Is multilingualism seen as added-value in bibliodiversity?: A literature review focussed on business and research contexts. *ELPUB 2019 23d International Conference on Electronic Publishing*. doi: 10.4000/proceedings.elpub.2019.17
- Balula, A., & Leão, D. (2021). Multilingualism within Scholarly Communication in SSH – a literature review. *JLIS.it*. 12(2), 88-98. doi: [10.4403/jlis.it-12672](https://doi.org/10.4403/jlis.it-12672)
- del Rio Riande, G., Tóth-Czifra, E., Wuttke, U., Moranville, Y. (2020). OpenMethods: A Compass for a More Open Digital Humanities. *Preprints*, 2020030016. doi: 10.20944/preprints202003.0016.v1
- Giglia, E. (2019). OPERAS: bringing the long tail of Social Sciences and Humanities into Open Science. *Italian Journal of Library and Information Science*. 10(1), 140–156. doi: 10.4403/jlis.it-12523.
- Guns, R., Eykens, J., & Engels, T. C. E. (2019). To What Extent Do Successive Cohorts Adopt Different Publication Patterns? Peer Review, Language Use,



and Publication Types in the Social Sciences and Humanities. *Frontiers in Research Metrics and Analytics*, 3. doi: 10.3389/frma.2018.00038

- Ibrahim, S., Fathalla, S., Yazdi, H. S., Lehmann, J., & Jabeen, H. (2019). From Monolingual to Multilingual Ontologies: The Role of Cross-Lingual Ontology Enrichment. in Acosta M., Cudré-Mauroux P., Maleshkova M., Pellegrini T., Sack H., Sure-Vetter Y. (eds). *Semantic Systems. The Power of AI and Knowledge Graphs. SEMANTICS 2019. Lecture Notes in Computer Science*, vol 11702, 215–230. Springer. doi: 10.1007/978-3-030-33220-4\_16
- Jokić, M., Mervar, A., & Mateljan, S. (2019). Comparative analysis of book citations in social science journals by Central and Eastern European authors. *Scientometrics*, 120(3), 1005–1029. doi: 10.1007/s11192-019-03176-y
- Koutsomitropoulos, D. A. (2019). Semantic annotation and harvesting of federated scholarly data using ontologies. *Digital Library Perspectives*, 35(3/4), 157–171. doi: 10.1108/dlp-12-2018-0038
- Kulczycki, E., Guns, R., Pölönen, J., Engels, T. C. E., Rozkosz, E. A., Zuccala, A. A., Bruun, K., Eskola, O., Starčić, A., Petr, M., & Sivertsen, G. (2020). Multilingual Publishing in the Social Sciences and Humanities: A Seven-Country European Study. *Journal of the Association for Information Science and Technology*, 1-15. doi: 10.1002/asi.24336
- Luzón, M. J. (2019). “Meet our group!”: Addressing multiple audiences on the websites of Spanish research groups. *International Journal of English Studies*, 19(2), 37-59. doi: 10.6018/ijes.382561
- Mounier, P. (2018). ‘Publication favela’ or bibliodiversity? Open access publishing viewed from a European perspective. *Learned Publishing*, 31, 299–305. doi: 10.1002/leap.1194
- Sivertsen, G. (2018). Balanced Multilingualism in Science. *BiD: textos universitaris de biblioteconomia i documentació*, 40. doi: 10.1344/BiD2018.40.25
- Sivertsen, G., Guns, R., Kulczycki, E., & Pölönen, J. (2019). The use of Gold Open Access in four European countries: An analysis at the level of articles. In *17th International Conference on Scientometrics & Informetrics: ISSI 2019: With a Special STI Indicators Conference Track: Proceedings* (vol. 2, pp. 1600–1605). Rome: Edizioni Efesto.

