

Open access challenge at the national level: comprehensive analysis of publication channels used by Finnish researchers in 2016-2017

Janne Pölönen¹, Raf Guns², Emanuel Kulczycki³, Mikael Laakso⁴ and Gunnar Sivertsen⁵

¹*janne.polonen@tsv.fi*

Federation of Finnish Learned Societies, Snellmaninkatu 13, 00170 Helsinki (Finland)

²*raf.guns@uantwerpen.be*

University of Antwerp, Faculty of Social Sciences, Centre for R&D Monitoring (ECOOM), Middelheimlaan 1, 2020 Antwerp (Belgium)

³*emek@amu.edu.pl*

Adam Mickiewicz University, Scholarly Communication Research Group, Szamarzewskiego 89c, 60-568 Poznań (Poland)

⁴*mikael.laakso@hanken.fi*

Hanken School of economics, Helsinki (Finland)

⁵*gunnar.sivertsen@nifu.no*

Nordic Institute for Studies in Innovation, Research and Education (NIFU), P.O. Box 2815, 0608 Tøyen, Oslo (Norway)

Abstract

The purpose of this paper is to provide a comprehensive picture of open access publishing in Finland. Data consists of the complete national peer-reviewed output of 48177 articles and books from 14 Finnish universities in 2016-2017 stored in the VIRTIA Publication Information Service. Each publication record contains an indication if it is openly available as Gold or Hybrid OA and/or if it is deposited in OA repository. Using this data, we investigate the share of openly available outputs across fields, as well as journal and book publishing, and analyse the open access status of all 10342 publication channels (journal/series and book publishers) used by Finnish researchers. We also examine the utility of international open access information sources, DOAJ and Bielefeld list for OA journals, and Sherpa/Romeo for self-archiving policies, in estimating the potential for open availability of peer-reviewed outputs, as well as the importance of the largest international commercial publishers in light of these comprehensive national data.

Introduction

In 2016, the European Union member states agreed to “open access to scientific publications as the default option by 2020 and to the best possible re-use of research data as a way to accelerate the transition towards an open science system” (Council of the European Union, 2016). The European Commission supports the transition with a strong open science agenda (European Commission, 2018). Most recently, a group of European research funders known as cOAlition S (which includes Finland’s largest research funder, the Academy of Finland) plans to make immediate open access and unrestricted use requirements for all published research funded by the signatories by 2020. This concern, in the first place, journal articles, while a longer transition period is admitted for peer-reviewed book publications.

Finland, like many European countries, is currently developing national strategies and incentives for advancing open access. In 2014-2017, the Ministry of Education and Culture funded a national project, the Open Science and Research Initiative, which set ambitious national targets for the share of openly available publications: 65% in 2017, 75% in 2018 and 100% in 2020 (Ilva, 2017b). According to European Open Science Monitor, the share of OA in Finland is 41.6%, so it ranks 19th out of 36 countries compared. Recently, the Finnish government has approved a new funding model for allocating core-funding annually to universities in 2021-2024. A publication indicator (Pölönen, 2018) will distribute 14 % of the funding, and the publication points based on publication type and channel are multiplied by 1.2

if the peer-reviewed output is openly available (independent of OA mechanism or embargo length). Meanwhile, the Ministry has invested in development of comprehensive national publication data that supports, in addition to the performance-based research funding system (PRFS), monitoring of open access publishing in Finland (Ilva, 2017a).

All countries face the challenge that the vast majority of peer-reviewed outlets used by researchers do not support Gold open access publishing. Many outlets allow individual papers to be made openly available on the publisher website, however, this hybrid OA model is considered unsustainable due to increasing costs and only partial open availability of outputs (Piwowar et al. 2018). It has also been observed that publishing in journals that allow self-archiving (Green OA) does not automatically mean that publications are actually deposited in open access repositories, highlighting a gap between potential and uptake (Laakso, 2014; Björk et al., 2014). Green journals may impose embargoes for the peer-reviewed post-print and publisher version, making them not compliant for example with the Plan S requirements.

Directory of Open Access Journals (DOAJ) and Sherpa/Romeo are the most frequently used information sources to identify Gold and Green OA channels. Nevertheless, even these sources may not provide full coverage of Gold and Green channels. Bielefeld university, for example, provides an ISSN-Matching of Gold OA Journals based – in addition to DOAJ – also to the Directory of Open Access Scholarly Resources (ROAD), PubMed Central (PMC) and Open APC (OAPC) (Wohlgemuth et al. 2016). It is, however, an open question to what extent these existing sources cover the whole variety of publication channels used at the national level, or help estimating the level of open availability of peer-reviewed outputs.

Given that the five largest international commercial publishers account for more than half of the journal output indexed in Web of Science (WoS) (Larivière et al. 2015), most attention at both international and national level is focused on pressuring and/or negotiating with these publishers for open availability of publications. But as Larivière et al. (2015) point out, WoS purports to cover only the most cited international subset of scholarly journals. Especially in the social sciences and humanities (SSH), WoS coverage is seriously wanting due to the importance of national language and book publishing (Kulczycki et al., 2018). In many SSH disciplines, the majority of journal articles are published in national or regional outlets not indexed in WoS (van Leeuwen & Sivertsen 2014; Sivertsen 2016). In addition, up to one half of peer-reviewed outputs in Humanities, and around one-third in the social sciences, are book publications (Engels et al. 2018).

The challenge of implementing and providing open access at a national level has various aspects of which we highlight three. Firstly, analysing what share of national output is published as OA and in how many and what kind of channels. This cannot be easily calculated on the basis of international databases such as WoS or Scopus, or Google Scholar (Martín-Martín, 2018). The implication is that only countries in which current research information systems with full coverage of the SSH publications (Sīle et al, 2018) have been developed can provide an accurate picture of publication patterns and OA publishing in all fields and across publication types. Such an analysis is important as a basis for tailor-made science policy instruments. Secondly, implementing OA at the national level requires infrastructure, tools and resources for open publishing (Sivertsen, 2018). It is an important prerequisite in the ongoing process of flipping journals to the OA model. One possible solution is to use the Open Journal Systems developed by the Public Knowledge Project, to provide a translation of the system into the national language, to provide some training materials, and to ensure resources and create incentives for flipping national journals to the OA model (Ilva, 2018). Another option is to build a national OA platform from the very beginning, as in Croatia (Stojanovski et al, 2009) or in Québec (Larivière and Macaluso, 2011). Thirdly, the challenge of open access at the national level is to provide all mentioned analyses and materials, infrastructure, and platforms also for the peer-reviewed book publications. Scholarly monographs, book chapters and edited volumes play a

key role in the social sciences, humanities, and law domains (Montgomery et al, 2018). Thus, not only journal articles but also books should be fully integrated into the OA scholarship.

In this paper we investigate the extent of such challenges by means of a comprehensive analysis of open access publishing in Finland based on complete national publication data. The national information sources remain under-exploited in analysis of open access publishing, and have focused predominantly on journal publishing (Ilva, 2017b; Kronman, 2017; Mikki, 2017; Mikki et al., 2018). Our main research questions are:

1. What is the share of openly available peer-reviewed journal and book publications across fields of science in Finland?
2. How many journals/series and book publishers do Finnish researchers use for publishing peer-reviewed outputs across fields of science, and how large is their share that provides for full, partial or no open availability of Finnish outputs?
3. How large is the share of journals/series that have been identified in VIRTAs as OA channels, and what share of these outlets are indexed in DOAJ and Bielefeld list?
4. How large is the share of journals/series that are indexed in Sherpa/Romeo, and does the self-archiving policy influence the share of Finnish outputs in those journals that are openly available?
5. How large is the share of book publishers that are identified in VIRTAs as OA channels or have permitted self-archiving?
6. To what extent do the largest international commercial publishers dominate the publishing of Finnish researchers, and are there differences between fields?

Data and Methods

The data consists of unique peer-reviewed outputs published in 2016-2017 that the 14 Finnish universities have reported to the Ministry of Education and Culture and that are stored in the VIRTAs publication information service (Sīle et al., 2017; Sīle et al., 2018; Pōlōnen, 2018). In VIRTAs, co-publications of Finnish universities appear as duplicates, however, duplicates are automatically identified on the basis of publication information and indicated in the data. In this study, we use deduplicated publication counts. For each publication, the reporting university has indicated the publication type, OECD field of science, peer review status and open availability. This study includes peer-reviewed articles in journals, books and proceedings, as well as monographs and edited works from all fields of science. For the year 2017 the data collection is not yet entirely complete.

The years 2016 and 2017 have been selected because universities have indicated the open availability of peer-reviewed outputs according to renewed definitions (Ilva, 2017a). Firstly, it is indicated for each output if it is openly available in either Gold or Hybrid OA publication channel. Secondly, it is indicated if the publication is openly available in an OA repository. Information on embargoes or OA licenses, however, is not available in the data. Consequently, it is possible to establish if a peer-reviewed publication is openly available in an OA or Hybrid channel, deposited in a repository, or both. The open availability of a publication can be verified using the URL provided in its metadata. The validation of openly available publications takes place at the universities, and involves both researchers and data collection personnel from the university libraries.

In VIRTAs, the publication channel – journal/series or book publisher – of each peer-reviewed output has been identified by matching the publication's bibliographic metadata to the Publication Forum authority list of publication channels. The authority list covers all journals/series and book publishers actually used by researchers affiliated with the 14 Finnish universities. Journals/series include mostly journals but also some book series with ISSN code, as well as some conference proceedings without ISSN. Book publishers mostly have a registered ISBN. For journals/series with ISSN, the Publication Forum channel register contains

the name of the publisher retrieved from the International ISSN Centre. We have complemented the ISSN Centre data with publisher information in the Scopus journal list. It is also indicated if the channel is included in the Directory of Open Access Journals (DOAJ), the Bielefeld list of open access journals, and what the self-archiving policy is according to Sherpa/Romeo. In 2016-2017, the 14 Finnish universities published a total of 48177 unique peer-reviewed outputs in 10342 publication channels, of which 91.9 % are journals/series and 8.1 % are book publishers (Table 1). 16.5 % of outputs are published with book publishers, while 83.5 % are published in journals/series. Only 62 % of all peer-reviewed outputs are published in journals indexed in Scopus and 52 % in WoS journals (Figure 1). There are, however, large differences between fields in the share of outputs in journals/series, as well as in Scopus and WoS coverage.

Table 1: Number of journals/series and book publishers and their share of outputs by main fields of science

Field of Science	Publication channels			Outputs		
	Journals/ Series	Book publishers		In Journals/ Series	In Book publishers	
	N	%	%	N	%	%
Natural sciences	3750	95.3 %	4.7 %	15230	89.7 %	10.3 %
Engineering	1888	91.1 %	8.9 %	6647	81.2 %	18.8 %
Medicine and health	2541	98.4 %	1.6 %	10189	98.5 %	1.5 %
Agriculture and forestry	404	93.3 %	6.7 %	900	95.1 %	4.9 %
Social sciences	3307	89.0 %	11.0 %	10608	72.4 %	27.6 %
Arts & humanities	1782	78.0 %	22.0 %	5920	64.7 %	35.3 %
All fields	10342	91.9 %	8.1 %	48177	83.5 %	16.5 %

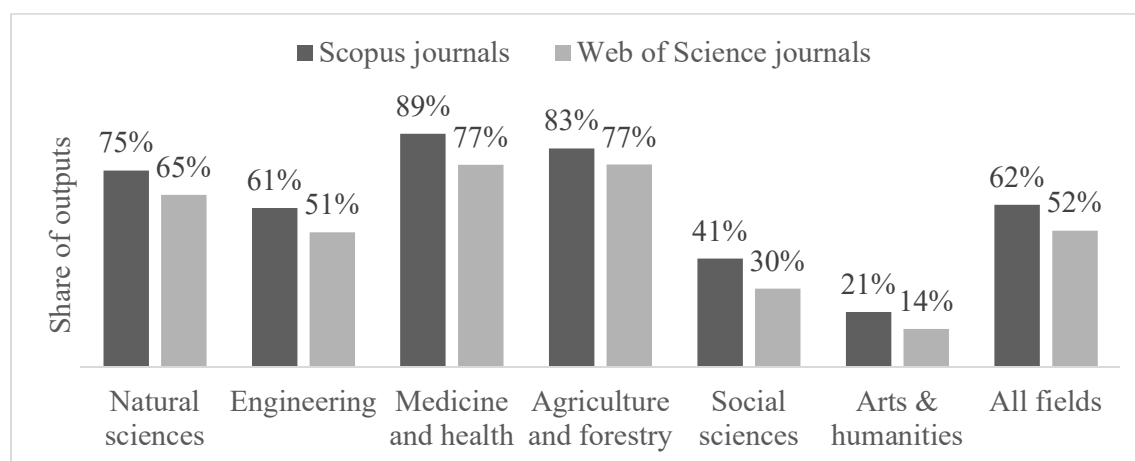


Figure 1: Scopus and WoS coverage of outputs by field of science

Results

Identification of open access status of publication channels based on VIRTA

In VIRTA, there is some evidence of open availability of outputs for one-half of the 10342 publication channels that Finnish researchers have used in 2016-2017 (Table 2). But there is considerable variation in the share of the Finnish outputs that are openly available in different channels. In roughly one-fourth of the channels (24.7 %) all Finnish outputs are openly

available, and in one-fourth (25.5 %) of the channels the open availability is only partial. Half (49.8 %) of the publication channels do not seem to have any publications reported as being openly available. This pattern is observed, more or less, in all the main fields, although the share of channels providing no form of open availability is somewhat larger in SSH. This is because open availability is more restricted in the case of book publishers than journal/series.

Table 2: Number of journals/series and book publishers and their share of outputs by main fields of science

Field of science and channel type	Publication Channels (N)	Share of openly available outputs in channel					
		100 %	<100% >=75%	<75% >=50%	<50% >=25%	<25% >0%	0 %
Natural sciences	3750	20.5 %	2.8 %	9.3 %	12.4 %	12.1 %	42.9 %
Engineering	1888	16.4 %	2.8 %	7.5 %	11.5 %	15.7 %	46.2 %
Medicine and health	2541	24.1 %	1.6 %	9.1 %	12.0 %	11.2 %	42.0 %
Agriculture and forestry	404	21.8 %	3.0 %	6.2 %	10.6 %	19.1 %	39.4 %
Social sciences	3307	24.6 %	2.7 %	9.6 %	10.0 %	7.9 %	45.2 %
Arts & humanities	1782	22.6 %	3.3 %	7.8 %	8.1 %	7.2 %	51.0 %
All fields	10342	24.7 %	1.8 %	7.9 %	8.6 %	7.2 %	49.8 %
- Journal/series	9500	25.6 %	1.8 %	8.1 %	9.1 %	7.1 %	48.3 %
- Book publisher	842	14.8 %	1.0 %	6.2 %	3.7 %	8.2 %	66.2 %

Of 9500 journals/series the Finnish researchers used as publication channels, 2074 have at least one peer-reviewed output stored in VIRTAs that has been indicated as being openly available in a Gold OA channel (21.8 % of journals/series). In the case of 281 journals/series, outputs are marked as being openly available in both Gold and Hybrid OA channel (3 %), so there is some ambiguity about the OA status of the channel. Outputs from 1137 journals/series have been indicated as being openly available in a Hybrid OA channel (12 %). There are further 1416 journals/series, from which outputs are indicated in VIRTAs as being openly available in an OA repository (14.9 %) but not in a Gold or Hybrid channel. For 4592 journals/series used by Finnish researchers we have no indication of any form of open access in VIRTAs (48.3 %). The share of journals/series identified as Gold OA channels is smaller for the largest commercial publishers than for the other publishers (Table 3).

Table 3. Type of open access of journals/series by publisher as identified in VIRTAs

Publisher	Publication channels (N)	Gold OA channel	Gold or Hybrid channel	Hybrid OA channel	Only self-archiving	No indication of open access
		%	%	%	%	%
Elsevier	1373	7.2 %	3.3 %	20.2 %	22.1 %	47.2 %
Springer Nature	605	10.4 %	3.0 %	23.0 %	13.2 %	50.4 %
Wiley-Blackwell	595	8.2 %	2.2 %	18.2 %	16.0 %	55.5 %
Taylor & Francis	553	7.6 %	1.8 %	11.9 %	19.9 %	58.8 %
Sage	273	9.9 %	0.7 %	7.7 %	27.1 %	54.6 %
ACS	46	6.5 %	6.5 %	34.8 %	23.9 %	28.3 %
Other	6055	29.6 %	3.1 %	8.4 %	12.3 %	46.6 %
All journals/series	9500	21.8 %	3.0 %	12.0 %	14.9 %	48.3 %

For the book publishers there is no comprehensive source on OA-status or self-archiving policy, such as DOAJ and Sherpa/Romeo for journals. The VIRTAs data indicates, however, that 186 different publishers have at least one output registered as being openly available in a Gold OA channel (22.1 % of the publishers). Outputs from 6 book publishers are indicated as being openly available in both Gold and Hybrid OA channels (0.7 %), so there is ambiguity about the OA status, and 4 book publishers have been identified as Hybrid channels (0.5 %). There are further 89 book publishers, of which outputs have been indicated as being self-archived in an open access repository (10.6 %) but they are not openly available in the publisher website. For 557 book publishers used by Finnish researchers there is no indication of open availability of any outputs (66.2 %). The share of Gold OA channels is about the same for both journals/series and book publishers, however, the availability of Hybrid OA and self-archiving options appear much more limited in the latter case.

There is a considerable difference in share of openly available outputs according to open access status of the channel, as well as according to publications channel type (Figure 2). The share of outputs indicated as being openly available in VIRTAs is largest in the identified Gold OA channels (79 %), followed by Hybrid OA channels (31 %) and smallest in journals/series with only self-archived outputs (26 %). The same is observed in case of book publishers, however, the overall share of openly available outputs is much smaller.

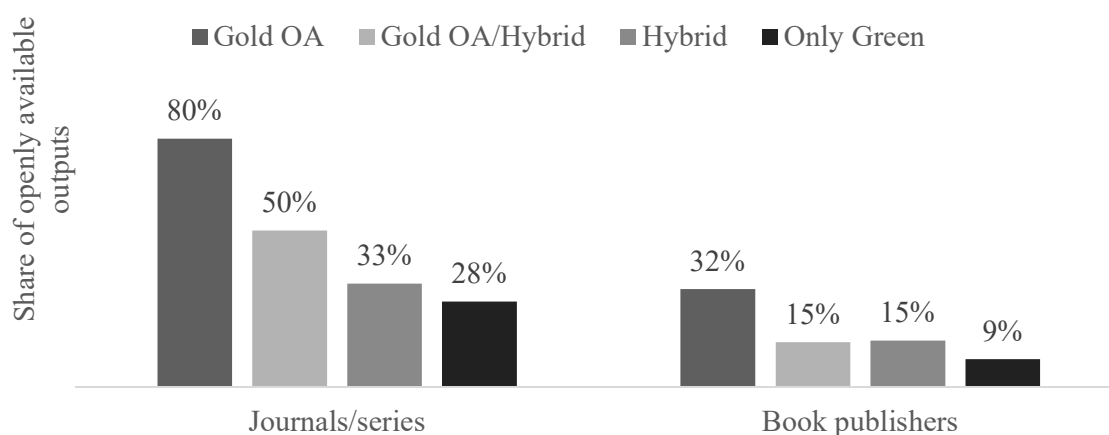


Figure 2. Open Access Status of publication channels as identified in VIRTAs and share of openly available outputs.

Comparison of VIRTAs based open access status of journals/series with DOAJ and Bielefeld list
 Of all 9500 journals/series used by Finnish researchers, 1237 are Gold OA journals indexed in DOAJ with or without a green tick (13 %) (Table 4). Furthermore, 372 journals/series are included in the Bielefeld list of open access journals but are not indexed in DOAJ (3.9 %). Comparison with VIRTAs data suggests that inclusion of journal/series in DOAJ and the Bielefeld list is a good predictor of open access, as 96 % outputs from channels in DOAJ and 78 % from channels in Bielefeld are actually indicated in VIRTAs as being openly available. For journals/series outside DOAJ and the Bielefeld list, the share of openly available outputs is considerably smaller (25 %), yet as large as 54 % in case of those journals/series indicated as Gold OA channels.

Together, DOAJ and the Bielefeld list cover over 60 % of all journals/series identified as Gold OA channels based on the VIRTAs data (including Gold/Hybrid OA journals). Combining all information sources it is possible to identify a total of 2553 potential Gold OA journals, of which 48 % based on DOAJ, 15 % based on the Bielefeld list, and an additional 37 % based on VIRTAs (Figure 2). It is noteworthy that 37 % of all potential Gold OA channels are not included

in either DOAJ or the Bielefeld list (it has not been possible for us to manually verify their Gold OA status).

Table 4. Comparison of VIRT A based Open Access Status with DOAJ and Bielefeld list, and share of openly available outputs.

	Publication channels	Outputs	Openly available outputs	Openly available outputs
	N	N	N	%
DOAJ	1237	6013	5765	95.9 %
+Bielefeld	372	1249	973	77.9 %
Not in DOAJ or Bielefeld list	7891	32977	8289	25.1 %
All publication channels	9500	40239	15027	37.3 %

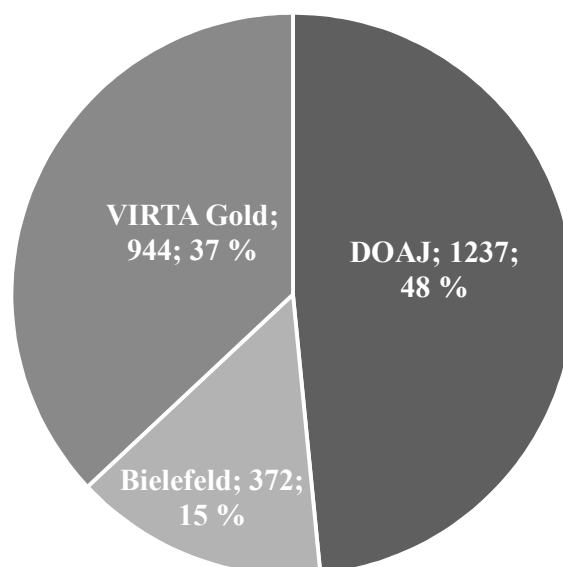


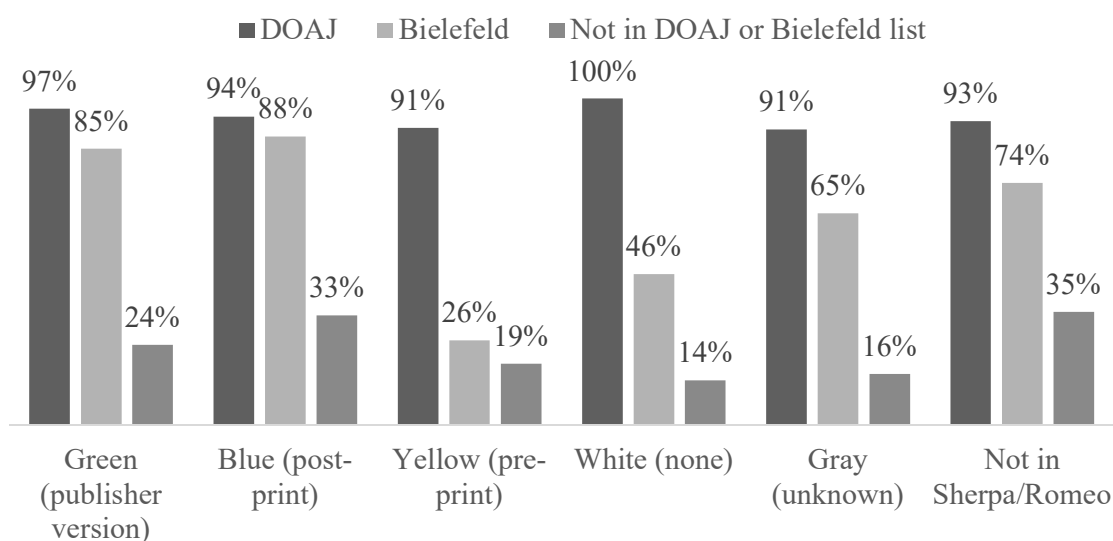
Figure 3. Share of potential Gold OA journals identified based on DOAJ, Bielefeld list and VIRT A

Comparison of Sherpa/Romeo self-archiving policies of journals/series with DOAJ and Bielefeld list and open availability of outputs in VIRT A

Sherpa/Romeo codes indicating the self-archiving policies cover 7537 journals/series (79 % of all journals/series) used by Finnish researchers (Table 5). Sherpa/Romeo includes almost all DOAJ journals (95 %), and a considerable share of Bielefeld listed journals (43 %). Overall, however, the inclusion of journals in Sherpa/Romeo is not a very good predictor of open availability of outputs, the share of which in VIRT A is practically the same as in the case of journals not included in the Sherpa/Romeo service (Figure 3). The share of openly available outputs is much larger for channels included in DOAJ or the Bielefeld list, than for the other channels included in Sherpa/Romeo. Availability of the Gold route clearly has resulted in more complete open availability of outputs than the Green route. The differences in self-archiving policy do not make a great difference, especially if we look at journals/series not in DOAJ or the Bielefeld list.

Table 6. Sherpa/Romeo codes and share of openly available outputs.

Sherpa/Romeo self-archiving policy	Publication channels	
	N	%
Green (publisher version)	5034	53.0 %
Blue (post-print)	361	3.8 %
Yellow (pre-print)	1346	14.2 %
White (none)	267	2.8 %
Gray (unknown)	529	5.6 %
Not in Sherpa/Romeo	1963	20.7 %
All publication channels	9500	100 %

**Figure 3. Sherpa/Romeo codes and share of openly available outputs in DOAJ and Bielefeld listed journals**

The importance of the largest international commercial publishers and open availability of the outputs across fields

Publication channels owned by Elsevier account for 20.1 % of the 14 Finnish universities' journal outputs in all fields of science counted together (Table 6). Next come Springer Nature (12.8 %), Wiley-Blackwell (9.2 %) and Taylor & Francis (6.8 %). Sage and the American Chemical Society (ACS), which are often also considered among the "big" commercial publishers, account for 2.3 % and 1.9 % respectively. Taken together, these publishers account for 53.1 % of the journal output. This is consonant with studies based on Web of Science data, even though national VIRTAs data includes many journals/series not indexed in WoS. If we take into account also peer-reviewed conference articles and book publications, these publishers' joint share of Finnish output diminishes to less than half (44.3 %). VIRTAs data also suggests that the commercial publishers included in this study are most dominant in Medicine and Agriculture, and least dominant in the social sciences and especially humanities. Thus, our study corroborates the findings of Larivière et al. (2015) concerning the humanities being the field least dominated by the big publishers. In our analysis, however, social sciences is among the least, not the most, dominated fields (this holds true even if we limit analysis to journal articles).

Of all Finnish 2016-2017 peer-reviewed outputs one-third is openly available (33.6 %) and two-thirds are not openly available (66.4 %) (Table 7). The share of openly available outputs is somewhat smaller in case of the large commercial publishers (except Springer Nature) than other publishers. The share of openly available outputs is also larger among journal articles than conference and book publications. Overall, the differences between fields are not great. Nevertheless, natural sciences (39 %) and medicine (37 %) have the largest, while SSH (30 %) and especially engineering (26 %) have smallest share of openly available outputs (Figure 4).

Table 6. The six largest commercial publishers' share of outputs by field of science and publication type

Field and publication type	Outputs N	Elsevier	Springer Nature	Wiley-Blackwell	Taylor & Francis	Sage	ACS	Other
		%	%	%	%	%	%	%
Natural sciences	15230	17.5 %	17.8 %	8.1 %	2.3 %	0.4 %	3.0 %	50.9 %
Engineering	6647	22.5 %	8.9 %	4.3 %	2.9 %	0.9 %	2.4 %	58.1 %
Medicine and health	10189	19.9 %	17.7 %	13.1 %	6.3 %	2.4 %	0.6 %	40.1 %
Agriculture and forestry	900	27.6 %	12.2 %	10.1 %	5.4 %	0.4 %	0.7 %	43.6 %
Social sciences	10608	8.6 %	8.7 %	3.9 %	14.0 %	4.2 %	0.0 %	60.5 %
Arts & humanities	5920	2.1 %	4.5 %	1.4 %	8.7 %	1.2 %	0.0 %	82.1 %
All fields	48177	14.9 %	12.8 %	6.9 %	6.6 %	1.8 %	1.4 %	55.7 %
- Journal article	34507	20.1 %	12.8 %	9.2 %	6.8 %	2.3 %	1.9 %	46.9 %
- Conference article	6283	2.6 %	9.9 %	0.3 %	0.9 %	0.0 %	0.0 %	86.3 %
- Book publication	7387	1.3 %	15.0 %	1.9 %	10.4 %	0.6 %	0.0 %	70.7 %

Table 7. Type of open availability of outputs by publisher and publication type

Publisher	Outputs N	Only publisher service	Publisher service and self-archived	Only self-archived	Not open access
		%	%	%	%
Elsevier	7188	5.2 %	8.8 %	11.7 %	74.3 %
Springer Nature	6164	8.4 %	25.5 %	7.0 %	59.2 %
Wiley-Blackwell	3328	5.0 %	9.4 %	8.4 %	77.2 %
Taylor & Francis	3163	3.4 %	6.7 %	11.1 %	78.8 %
Sage	855	3.9 %	6.8 %	14.6 %	74.7 %
ACS	651	1.8 %	5.5 %	9.1 %	83.6 %
Other	26828	13.2 %	16.2 %	8.2 %	62.4 %
All publishers	48177	9.8 %	14.9 %	8.9 %	66.4 %
- Journal article	34507	9.9 %	18.7 %	9.6 %	61.8 %
- Conference article	6283	11.8 %	7.2 %	9.6 %	71.4 %
- Book publication	7387	7.9 %	3.4 %	5.1 %	83.5 %

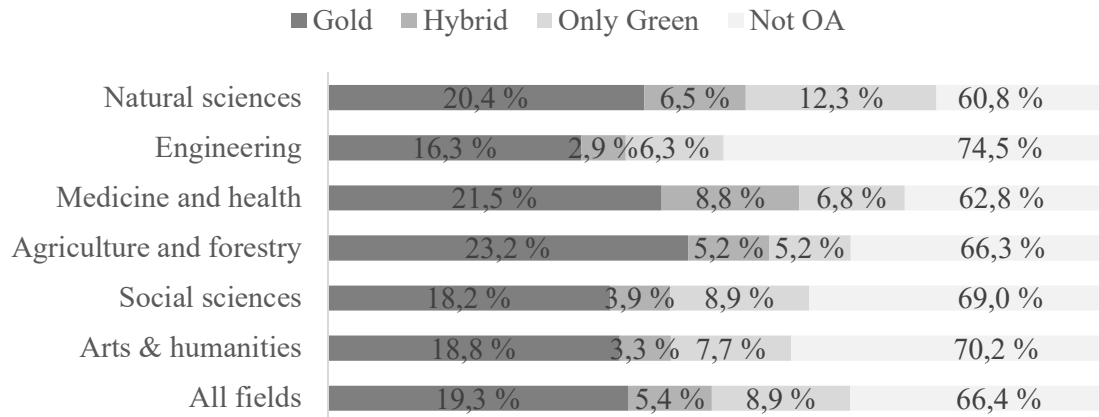


Figure 4. Type of open availability of outputs by field of science

Discussion and conclusions

The international data sources (Web of Science and Scopus), which are most often used for monitoring open access publishing, privilege journal outputs and STEM fields. The national publication data stored in the VIRTAs publication information service from the 14 Finnish universities, including 48 177 peer-reviewed outputs from 2016-2017, provides a more complete picture of open access by also including book publications as well as all SSH journal publications. Scopus journals cover only 62 %, and WoS journals 52 %, of all these outputs. Taking all fields and publications types into account, the share of openly available outputs is 34 %. For 25 % of the outputs open availability is provided in a Gold or Hybrid channel, while 9 % are openly available only in repositories. The differences between fields in the share of openly available outputs range from 39 % in the natural sciences to 26 % in engineering.

The Finnish researchers used 10 342 different publication channels as outlets, including 9500 journals/series and 842 book publishers. In 25 % of the channels all Finnish outputs are openly available. In 25 % of the channels, however, the open availability is only partial, and in case of 50 % of the channels no openly available outputs have been reported in VIRTAs. It is important to remember that we rely here on universities' self-reported OA status of publications. These results mean that, for Finland to achieve the target of open availability of all peer-reviewed outputs in the near future, around 5000 currently used channels should either be replaced with alternative open access channels or should flip to the required gold or green open access publishing models. Around 2500 channels already provide for open availability of some outputs, hence closing the gap between potential and uptake is the key.

The majority of journals/series used by the Finnish researchers (79 %) have a self-archiving policy registered in Sherpa/Romeo. Analysis of the share of Finnish outputs published in these journals shows that a relatively small share is openly available, irrespective of the self-archiving policy indicated with colour code, unless the outlet also provides open availability via Gold OA (DOAJ-indexed or Bielefeld listed journals). The share of openly available outputs is only slightly larger in the case of Hybrid OA channels than in channels permitting only self-archiving. Our results confirm that there indeed is considerable potential for advancing open availability via Green route (Laakso 2014; Björk et al. 2014). It remains to be seen if open access incentives, such as the extra-weight for openly available publications in the Finnish universities' core funding-model, help to increase the uptake.

As expected, publishing in DOAJ-indexed journals is a good predictor of open availability of outputs. However, only 13 % of the journals/series used by the Finnish researchers are indexed in DOAJ. These account for 15 % of all peer-reviewed journal outputs, and 38 % of all openly

available journal outputs (including book publications, the shares are 12 % and 35 % respectively). A total of 944 journals/series identified in VIRTAs as OA channels are not covered in DOAJ or Bielefeld list. It has not been possible for us to investigate if these journals/series might meet the DOAJ criteria. Nevertheless, our findings point at considerable gap in the information sources on OA channels. Combining all OA information, it was possible to identify 2553 potential Gold OA journals, of which DOAJ covers 48 % and the Bielefeld list additional 15 %. Our findings suggest that relying on external information sources, such as DOAJ, in the identification of open access publications may not result in complete picture of Gold OA publishing.

We also investigated the importance of large international publishers. Elsevier, Springer Nature, Wiley-Blackwell, Taylor & Francis, Sage and American Chemical Society account for 53 % of the Finnish peer-reviewed journal output, and 44 % of all outputs including conferences and book publications. In all, their dominance appears less pronounced than in analyses using Web of Science data, especially in case of humanities as well as social sciences (Larivière et al., 2015). This means that negotiations with the largest international publishers can provide only partial solution to advancement of open access, which – especially in the SSH – depends on open access publishing models adopted by large variety of relatively small journal and book publishers operating in national context (Ilva, 2018; Late et al., 2018).

The VIRTAs publication data provides valuable information on the open availability of peer-reviewed book publications compared to journal articles (conference articles as a group is a mixture of both these publication types). The share of articles in books, monographs and edited works that are openly available is smaller (17 %) than that of journal articles (39 %). Nevertheless, 186 different book publishers (22 % of all publishers used by the Finnish researchers) are identified in VIRTAs as Gold OA channels providing for open availability via publisher website at least to some of the outputs. Hybrid and self-archiving options appear, however, more restricted in case of book publishers. Our findings highlight the need for international register of academic/scholarly book publishers that would contain information – like DOAJ – on their peer-review practices, as well as open access status and self-archiving policies.

In all, we conclude that national publication data can provide valuable information on the open availability of peer-reviewed outputs. To enhance comprehensive and comparable monitoring of open access we recommend development of well-structured and comprehensive national and international publication information sources.

References

- Björk, B. C., Laakso, M., Welling, P., & Paetau, P. (2014). Anatomy of green open access. *Journal of the American Society for Information Science and Technology*. doi: 10.1002/asi.22963
- Council of the European Union (2016). *Outcome of the council meeting, 3470th Council meeting, Competitiveness (Internal Market, Industry, Research and Space) Brussels, 26 and 27 May 2016*. Retrieved January 17, 2018 from: <https://www.consilium.europa.eu/media/22779/st09357en16.pdf>.
- Engels, T., Starčič, A., Kulczycki, E., Pölönen, J. & Sivertsen, G. (2018). Are book publications disappearing from scholarly communication in the social sciences and humanities? *Aslib Journal of Information Management*, 70:6 (2018).
- European Commission (2018). *OSPP-REC: Open Science Policy Platform Recommendations*. https://ec.europa.eu/research/openscience/pdf/integrated_advice_opspp_recommendations.pdf.
- Ilva, J. (2017a). Towards reliable data – counting the Finnish Open Access publications. *Procedia Computer Science* 106, 299–304.
- Ilva, J. (2017b). Suomalaisten yliopistojen avoimet julkaisut vuonna 2016 OKM:n julkaisutiedonkeruun tietojen valossa, *Informatiivitutkimus* 3–4 (36): 51-69.
- Ilva, J. (2018). Looking for commitment: Finnish open access journals, infrastructure and funding. *Insights*, 31, 25.

- Kronman, U. (2017). Open Access i SwePub 2010 – 2016. Retrieved January 17, 2018 from: https://web.archive.org/web/20171216091206/http://openaccess.blogg.kb.se/files/2017/12/Open_Access_i_SwePub_2010-2016_v1.pdf
- Kulczycki, E., Engels, T., Pölonen, J., Bruun, K., Duskova, M., Guns, R., Nowotniak, R., Petr, M., Sivertsen, G., Starčić, A., & Zuccala, A. (2018). Publication patterns in the social sciences and humanities: evidence from eight European countries, *Scientometrics*, 26.3.2018.
- Laakso, M. (2014). Green open access policies of scholarly journal publishers: a study of what, when, and where self-archiving is allowed, *Scientometrics*, 99, 475–494.
- Larivière V, Haustein S, Mongeon P (2015). The Oligopoly of Academic Publishers in the Digital Era. *PLoS One*, 10(6): e0127502.
- Larivière, Vincent, and Benoit Macaluso. Improving the coverage of social science and humanities researchers' output: The case of the Érudit journal platform. *Journal of the American Society for Information Science and Technology*, 62.12 (2011): 2437-2442.
- Late, E., Korkeamäki, L., Pölonen, J. & Syrjämäki, S. (2018). The role of learned societies in scholarly publishing in Finland. In M. Schirone et al. (eds), *Book of Abstracts, Nordic Workshop on Bibliometrics and Research Policy, Borås 7-9 November 2018*.
- Martín-Martín, A., Costas, R., van Leeuwen, T. & Delgado López-Cózar, E. (2018). Evidence of open access of scientific publications in Google Scholar: A large-scale analysis, *Journal of Informetrics*, 12, 819–841.
- Mikki, S. (2017). Scholarly Publications beyond Pay-walls. Increased Citation Advantage for Open Publishing. *Scientometrics* 113: 1529.
- Mikki, S.; Gjesdal, Ø.L.; Strømme, T.E. (2018). Grades of Openness: Open and Closed Articles in Norway. *Publications*, 6, 46.
- Montgomery, L. and Neylon, C. and Ozaygen, A. and Saunders, N. and Pinter, F. (2018). *The Visibility of Open Access Monographs in a European Context: Full Report, The Visibility of Open Access Monographs in a European Context: Full Report*. Curtin University of Technology, Humanities Research and Graduate Studies.
- Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, Farley A, West J, Haustein S. (2017). The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ Preprints* 5:e3119v1
- Pölonen, J. (2018). Applications of, and Experiences with, the Norwegian Model in Finland, *Journal of Data and Information Science*, 3(4), 31–44.
- Sîle, L., Guns, R., Sivertsen, G., & Engels, T. C. E. (2017). *European Databases and Repositories for Social Sciences and Humanities Research Output*. Antwerp: ECOOM & ENRESSH. Retrieved January 17, 2018 from: <https://doi.org/10.6084/m9.figshare.5172322.v2>.
- Sîle, L., Pölonen, J., Sivertsen, G., Guns, R., Engels, T., Arefiev, P., Dušková, M., Faurbæk, L., Hollg, A., Kulczycki, E., Nelhans, G., Petr, M., Pisk, M., Soós, S., Stojanovskil, J., Stone, A., Šušolo, J., & Teitelbaum, R. (2018). Comprehensiveness of national bibliographic databases for social sciences and humanities: findings from a European survey, *Research Evaluation*, 27(4), 310-322. doi: 10.1093/reseval/rvy016.
- Sivertsen, G. (2016). Patterns of internationalization and criteria for research assessment in the social sciences and humanities. *Scientometrics*, 107(2), 357–368.
- Sivertsen, G. (2018). Balanced multilingualism in science, *BiD: Textos Universitaris de Biblioteconomia i Documentació*, (40). Doi: <http://dx.doi.org/10.1344/BiD2018.40.25>.
- Stojanovski, J., Jelka, P. & Bojan, M. (2009). The Croatian national open access journal platform. *Learned Publishing* 22.4: 263-273.
- van Leeuwen, T. & Sivertsen, G. (2014). Scholarly publication patterns in the social sciences and humanities and their relationship with research assessment. In *Science, Technology and Innovation Indicators*, Utrecht.
- Wohlgemuth, M., Rimmert, C. & Winterhager, M. (2016). *ISSN-Matching of Gold OA Journals (ISSN-GOLD-OA)*. Bielefeld University. doi:10.4119/unibi/2906347.