

Article

Grades of Openness: Open and Closed Articles in Norway

Susanne Mikki *, Øyvind L. Gjesdal and Tormod E. Strømme

University of Bergen Library, University of Bergen, 5007 Bergen, Norway; oyvind.gjesdal@uib.no (Ø.L.G.); tormod.stromme@uib.no (T.E.S.)

* Correspondence: susanne.mikki@uib.no; Tel.: +47-909-082-66

Received: 29 August 2018; Accepted: 19 November 2018; Published: 22 November 2018



Abstract: Based on the total scholarly article output of Norway, we investigated the coverage and degree of openness according to the following three bibliographic services: (1) Google Scholar, (2) oaDOI by Impact Story, and (3) 1findr by 1science. According to Google Scholar, we found that more than 70% of all Norwegian articles are openly available. However, the degrees of openness are profoundly lower according to oaDOI and 1findr at 31% and 52%, respectively. Varying degrees of openness are mainly caused by different interpretations of openness, with oaDOI being the most restrictive. Furthermore, open shares vary considerably by discipline, with the medicine and health sciences at the upper end and the humanities at the lower end. We also determined the citation frequencies using cited-by values in Google Scholar and applying year and subject normalization. We found a significant citation advantage for open articles. However, this was not the case for all types of openness. In fact, the category of open access journals was by far the lowest cited, indicating that young journals with a declared open access policy still lack recognition.

Keywords: bibliometrics; publication statistics; open access; citation impact

1. Introduction

There is a growing demand to make research freely available to everyone, and this has resulted in several developments over the last few years. First, governments, funders, and institutions are increasingly mandating open access (OA) to research publications. The Norwegian Government introduced a mandate on OA to scholarly articles in August 2017 [1] with the aim of making all Norwegian publicly funded research OA by 2024. Secondly, the Max Planck Digital Library initiative, OA2020 [2], which seeks to transform the publishing system by replacing the subscription model with an OA model, has been followed by a rise in OA big deals with publishers. Third, social sharing networks have increasingly made scholarly publications freely available, often in breach of copyright regulations [3]. Finally, the website Sci-Hub illegally hosts more than 70 million research articles, providing access to the majority of recently-published articles worldwide [4]. In total, these initiatives have resulted in a rapid increase in research publications that are free to read on the Internet.

This is confirmed by large-scale studies on the state of openness (none of these studies include Sci-Hub in their analysis). A study by ScienceMetrix that analyzed articles published between 1996 and 2013 found a global open availability level of over 45% [5]. That study was conducted using a custom automated system that would later become the proprietary database 1findr. Martín-Martín et al. [6] used Google Scholar to analyze levels of openness to articles published between 2009 and 2014 across all countries and fields of research and found that 54.6% were freely available. A similar study by Mikki [7] using Google Scholar, but limited to scholarly articles from Norway between 2011 and 2015, found that 68% were freely available. Piwowar et al. [8] used the Unpaywall database (formerly known as oaDOI) to analyze OA levels. The Unpaywall database only matches articles with DOIs and adheres

to a stricter definition of openness by only including articles that are available on journals' websites or in repositories. Based on three different samples of 100,000 articles, the study by Piwowar et al. [8] estimated that at least 28% of the scholarly literature is OA with levels up to 45% for 2015. By using the Unpaywall data through the Web of Science (WoS) interface, Bosman and Kramer [9] analyzed articles published between 2010 and 2017. They found increasing levels of OA of up to almost 30% in 2016. Lower levels than reported by Piwowar et al. [8] are due to a policy decision by WoS to exclude non-peer-reviewed article versions (typically in pre-print archives), which are provided by Unpaywall.

The most influential definition of OA comes from the 2002 Budapest Open Access Initiative (BOAI), which defines OA (to peer-reviewed literature) as "free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers" [10]. In a later recommendation, the BOAI separates between gratis and libre OA, where gratis OA removes price barriers, while libre OA also removes permission barriers [11]. The understanding of OA as "free to read" (gratis OA) is commonly held and is also valid for the freely available articles in this study. However, there are questions concerning legality and permanency. Many articles that are currently available might be removed due to copyright violations, and others that publishers have made freely available might be put behind paywalls. Based on these concerns, the term OA will only be used on subsets of the findings in this study. In general, we will use the terms freely available or openly available articles.

The citation impact of open and closed articles has been discussed repeatedly, and a significant advantage has been found for open articles [5,7,12,13]. Previous methodological issues in study design, such as citation normalization, OA time lags, and possible selection bias, the latter arguing that high-quality papers are preferably published openly [14], have since been taken into consideration and more carefully handled. Thanks to Unpaywall, we are now also able to distinguish between types of openness and to determine which types result in different degrees of impact. So far, gold articles seem to be the lowest cited (17% below average), which is right behind closed articles (10% below average) [8], and open journals still struggle with their reputations.

In this study, we investigated the three influential services (Google Scholar, oaDOI, and 1findr) and their ability to provide the community with full texts. We based our investigation on the total scholarly article output of all higher education institutions in Norway. The data from oaDOI and 1findr were collected during the winter 2017/2018, and we are aware that recent changes and improvements have been made to these services, which might lead to different results.

We aimed to answer the following questions:

1. How large is the share of open availability according to oaDOI?
2. What type of open availability is provided?
3. How does the citation impact vary by type of open availability?
4. How does this share compare to Google Scholar's free share?
5. How does this share compare to 1findr's free share?
6. How open are the different subject fields?

2. Applied Services, Data, and Method

We assumed that most of the scholarly peer-reviewed publication output of Norwegian higher education institutions is registered in Cristin, the Current Research Information System in Norway. The institutions and researchers have a strong incentive to register their publications in Cristin because governmental funding of higher education institutions in Norway is partly based on publication output. The funding is distributed among institutions based on the number and type of publications registered and reported through Cristin.

To be reported, publications must fulfil the requirements listed in the reporting instructions [15]. In general, this limits the reported publications to original research, including articles, review articles,

academic monographs, and anthologies. The institutions have the responsibility to ensure that all relevant publications are registered in Cristin and that only publications that fulfil the above-mentioned requirements are reported. Because national requirements on OA are posed on articles only, we limited our studies to these.

Based on metadata provided by Cristin, we queried Google Scholar, oaDOI, and 1findr. Preferably, we used the article DOI as a unique identifier for querying. For articles with missing or erroneous DOIs, we queried CrossRef to enrich our data set.

Applying the REST API provided by CrossRef, we used the ISSN in conjunction with other document parameters, such as the author and title (ex. <https://api.crossref.org/journals/0373-2630/works?query.author=Author&query.title=A%20Title>).

We soon learned that our CrossRef query, based on ISSN matches, was somewhat too restrictive and that a more flexible combination of metadata, or using other APIs such as those of CrossRef, to match many free-form citations to DOIs would be a better approach. Even though we missed some DOIs, we do not expect the results to have been notably affected.

We scored the results using the open source library SimMetrics, which compares titles by different similarity measures (Cosine and Jaccard) and returns values between 0 and 1 for no and complete matches. We then weighted the returned results differently for Publishing Year (0.5), First Author Name (0.7), and Title (1 for both Cosine and Jaccard), giving greatest weight to the Title, confer values in brackets. According to this weighting, the sum of all similarities was equal to 3.2 for exact matches. After thoroughly testing and manual checking, we decided that records with values larger than 2.2 should be considered exact matches and available metadata added to our set. Applying our similarity algorithm, we managed to enrich our dataset with 2663 DOIs (4%). When multiple results were returned, we selected the result based on max score.

As shown in Table 1, the share of all documents with a DOI was 82.3%. Of 87,439 articles, 71,953 had or could be allocated a DOI. We also found that the proportion of articles with a DOI increased continuously from 78.8% in 2011 to 89.2% in 2016.

Table 1. Number of articles indexed in Cristin and number of articles with DOIs.

| Year | Total Articles | Articles with a DOI | Percentage of Articles with a DOI |
|-----------|----------------|---------------------|-----------------------------------|
| 2011 | 12,821 | 10,102 | 78.8% |
| 2012 | 14,224 | 11,357 | 79.8% |
| 2013 | 13,916 | 11,272 | 81.0% |
| 2014 | 14,574 | 11,885 | 81.5% |
| 2015 | 15,370 | 12,584 | 81.9% |
| 2016 | 16,534 | 14,753 | 89.2% |
| All years | 87,439 | 71,953 | 82.3% |

2.1. Applied Services

Of the investigated services, Google Scholar is by far the largest container of scholarly bibliographic records [16]. According to our earlier study, it covers the entire article output of Norway [7]. For Google Scholar, no API is available for querying and extracting data. Instead, an in-house script [17] was developed for automatic web scraping, based either on the article DOI or the article title.

For articles published between 2011 and 2015, we used the data sampled earlier [7]. For articles published in 2016, we repeated the developed routines, adjusted the script according to changes made by Google Scholar, and manually solved the appearing CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart). Queries were carried out off campus in order to avoid full-text links to library subscriptions. Finally, we carefully compared input and output titles to eliminate false matches.

We only tracked Google Scholar's best full-text location and disregarded additional versions. This was because we reused data from our earlier study where information on article versions was not collected. In the present study, type of open availability was therefore only investigated through oaDOI and 1findr.

For the provided full-text links, we did not control for copyright compliance, involved usage licenses, or the actual provided content. We are aware of that this might have introduced errors, but for practical and time-saving reasons we had to skip this type of manual control mechanism. Still, Google Scholar is an important driver for open availability and serves here as a valuable comparative measure.

The focus of this study was on oaDOI and its outstanding ability to distinguish between types of openness. Since our investigation, the service has grown and now harvests over 50,000 publishers and includes about 20 million open documents [18]. The service is run by Impactstory, a nonprofit organization, using open source codes, thus making reuse and further development possible. At the time of our study, we queried the then older and less enhanced version, oaDOI, using the available API through the OpenRefine software. As the name oaDOI indicates, the service builds solely on DOIs, and publications missing DOIs are not included. Considering only articles with assigned DOIs, the coverage in oaDOI for our dataset was almost complete (99%).

The third investigated service was 1findr by 1science. According to their homepage, 1findr consists of 90 million scholarly articles, whereof 27 million are freely available [19]. At the time of the data collection for this study, 1findr was not free for academic use. However, for the purpose of this study the service kindly provided an API for querying. Similarly as for Google Scholar, we foremost used the DOI in our query, otherwise we used the title. For the title queries, we noticed that the 1findr API had a fuzzy operator (~as suffix), and we used this to increase our number of potential matches. We then compared the results from 1findr with the input data using SimMetrics to identify matching articles.

2.2. Categories of Open Availability

In the following discussion on open availability, we stick to the different classifications given by the investigated services. According to oaDOI, the following evidence for openness are given (Table 2).

Table 2. oaDOI categories of open availability.

| Category | Evidence in oaDOI |
|-----------------|--|
| Open Journal | oa journal (via issn in DOAJ) |
| Open Journal | oa journal (via journal title in DOAJ) |
| Open Journal | oa journal (via publisher name) |
| Open Repository | oa repository (via OAI-PMH doi match) |
| Open Repository | oa repository (via OAI-PMH title and first author match) |
| Open Repository | oa repository (via OAI-PMH title and last author match) |
| Open Repository | oa repository (via OAI-PMH title match) |
| Open Repository | oa repository (via pmcid lookup) |
| Open Toll | open (via crossref license) |
| Open Toll | open (via free pdf) |
| Open Toll | open (via page says license) |
| Open Toll | open (via page says Open Access) |
| Closed | All other articles, including those shared only on ASN * |

* ASN: Academic Social Networks.

For the aim of this study, the evidence was grouped by

1. Open Journal: Typically indexed by the DOAJ, the Directory of Open Access Journals
2. Open Repository: Open full-text in approved OA repositories.
3. Open Toll: Free via publisher sites, with or without an open license in a toll-access journal.
4. Closed: All remaining articles.

The oaDOI service specifies a particular category for articles that are open (via free pdf) without being licensed as such. Interestingly, this is the largest “open” category, typically consisting of delayed open content, hidden gold access, and last but not least content opened by the publishers themselves (probably to retain their market position [20]). The unclarified status of these articles makes this category questionable concerning the permanence of openness and reuse rights.

The categories of availability as defined in 1findr are shown in Table 3.

Table 3. 1findr categories of availability.

| Category | Description |
|----------|---|
| CL | Closed |
| GO | Gold: via DOAJ, hybrid, free to read |
| GR | Green: via repository, including academic social networks |
| UN | Unknown: open versions, may belong to GO or GR |

The wide categories of GR and UN are problematic because versions of manuscripts might be uploaded in infringement with copyright, and academic social networks (ASNs) might require registration and login before giving access. Nevertheless, we found 1findr useful to determine both coverage and availability.

2.3. Citations and Subject Fields

For the citation analysis, we used article data from 2011 to 2015. Citation counts were extracted from Google Scholar [7], and all articles were allocated a subject field (in total 85 fields) according to the journal they are published in. These subject fields are defined and approved by the national publishing committee [21]. In addition to publishing year, citation data are normalized by these subject fields.

3. Results and Discussion

In this section, we present and compare results by service. Furthermore, we relate our findings to similar studies in the existing scientific literature.

3.1. oaDOI

As mentioned in the previous section, we searched oaDOI by DOI whenever available. The articles that were missing a DOI were added to the Closed category (15,460 articles) (Figure 1). In fact, our previous study based on Google Scholar’s link provider showed that articles without a DOI are more likely to be closed than articles with DOIs [7].

Articles might be openly available through multiple categories, for example, Open Toll and Open Repositories. If nothing else was stated, Open Toll took precedence over Open Repository. In other words, categories were exclusive and reflected oaDOI’s “Best_location” field, presuming the original publisher site to be the primary and preferred location for the article. Because the self-archived content only represented the proportion of articles that were not provided openly elsewhere, this category was underrepresented and did not adequately reflect the entire volume of self-archived content available in repositories. When we changed the precedence of categories and chose self-archived content first, we found that this category increased from 4% to 18% of the total article set (Figure 2).

For our sample, the two largest open categories were open (via page says license), which are typically hybrid, and open (via free pdf), which are typically free to read without an open license (Figure 1).

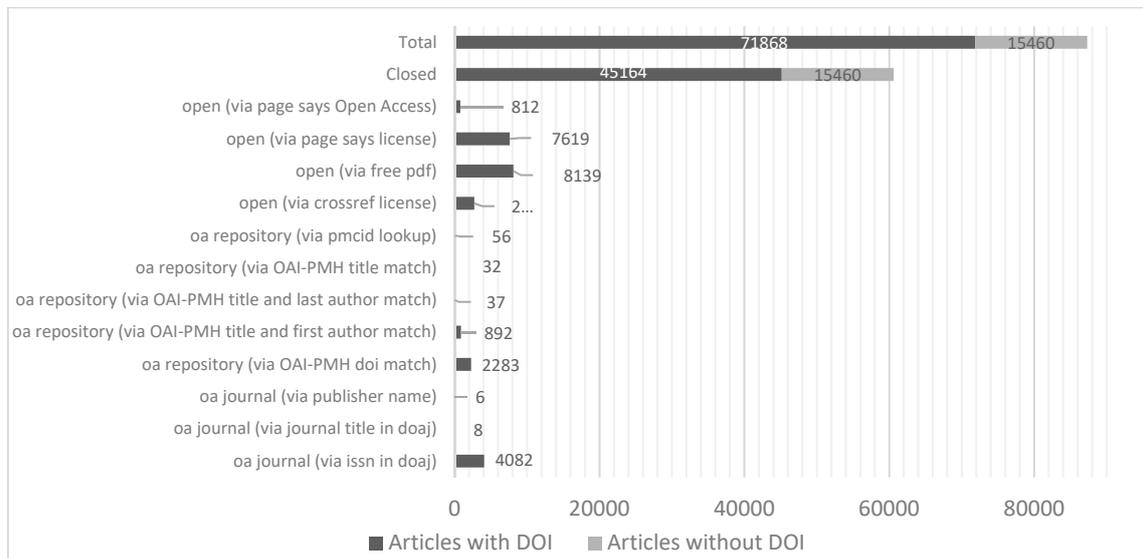


Figure 1. Open and closed articles according to oaDOI Best_location, 2011–2016.

According to our merged categories (see Table 2), we found that for all articles, including articles without a DOI, 31% were open, with 22% via Open Toll, 5% via Open Journal (typically DOAJ), and 4% via Open Repository.

Open availability gradually increased over the investigated period from 27% in 2011 to 34% in 2016 (Figure 2). However, these shares were considerably lower than those reported by Piwowar et al. [8], who found an open share of 44.7% for 2015 using a CrossRef sample. When we excluded articles without DOIs, the open share increased from 34% to 38%. This was still significantly lower than expected and might be explained by the fact that oaDOI, at the time of investigation, did not sufficiently harvest open repositories or otherwise failed to locate available full-text manuscripts.

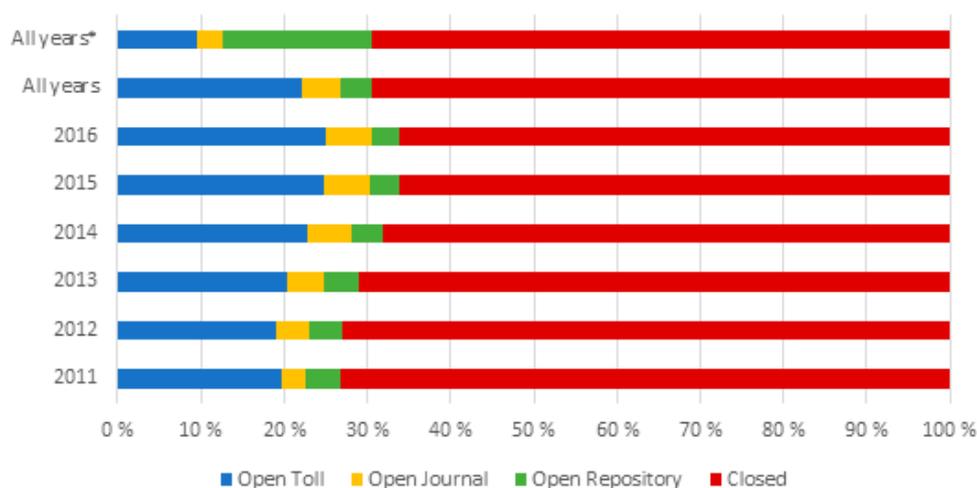


Figure 2. Open articles according to oaDOI’s Best_location, 2011–2016. The bar “All years*” reflects the total number of self-archived articles otherwise classified under Open Toll or Open Journal.

The category Open Journal made up 5% of all articles (Figure 2, all years). As argued above, this category might be overshadowed by other categories, in particular the category Open Toll.

When giving precedence to Open Journal, in case of multiple occurrences, this share increased to 16% in 2016, but still considerably lower than the findings of 20% OA for the University of Bergen [22]. oaDOI's underrepresentation of articles in that category, typically consisting of DOAJ journals, might be explained by many of these journals missing DOIs for articles represented in the Cristin data.

The bar labeled "All years *" in Figure 2 reflects the total number of self-archived articles, i.e., articles also available in Open Toll or Open Journal. The share of Open Repository then increased from 4% to 18%. Thanks to the efforts of libraries and thanks to strong OA requirements, a major part of the open articles, as located through oaDOI, indeed are available in repositories. However, the share is still low, and the potential of self-archiving is far from exhausted (see, for example, [23]).

3.2. oaDOI Compared to Google Scholar

Applying Google Scholar's link provider on our data, we found 56,359 articles (64%) freely available, twice as many compared to oaDOI (Figure 3). Figure 3 also shows 1090 articles that were closed in Google Scholar but open in oaDOI. This discrepancy might be because our oaDOI searches were performed one year later than searches in Google Scholar, and some articles might have become open during that time. It is also possible that Google Scholar lacks certain full-text sources that are covered by oaDOI, but due to Google's business model this information is secret and not verifiable.

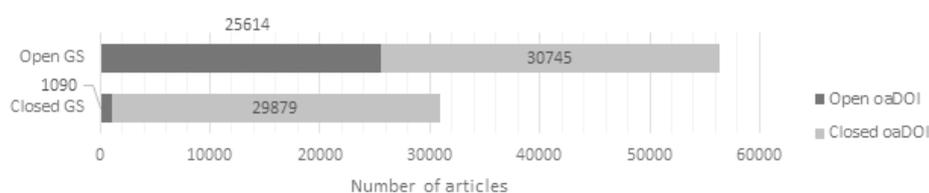


Figure 3. Open and closed articles according to Google Scholar (GS) and oaDOI, 2011–2016.

3.3. Citation Frequencies by Type of Openness

We determined the citation frequencies using Cited-by values according to Google Scholar and applying year and subject normalization, and we found a significant citation advantage for open articles (Figure 4a). However, this was not true for all types of openness according to oaDOI's classification. The category Open Journal was, in fact, by far the lowest cited. This disadvantage was even larger here than that reported by Piwowar et al. [8]. The categories Open Toll and Open Repository scored highest, with Open Repository at the very top. Articles belonging to Open Repository were most often cited both for self-archived-only and for all self-archived articles (Figure 4b,c).



Figure 4. Mean relative citation frequencies of articles indexed in (a) Google Scholar, (b) oaDOI Best_location (here Open Repository refers to self-archived-only), and (c) oaDOI giving precedence to Open Repository, i.e., all self-archived articles. The analysis was based on articles published between 2011 and 2015.

Our findings indicated that there indeed exists a so-called selection bias [24] where authors are more inclined to upload their most influential articles. Often these articles are a result of research that is supported by high-profile funding institutions with clear OA requirements. To fulfill these requirements, authors seem to self-archive their work but continue to publish in traditional subscription-based journals.

3.4. Openness by Subject

Within the Norwegian system, all journals are assigned one and only one subject for a total of 85 different subjects [21]. These 85 subjects are, in turn, exclusively assigned to four main subject categories. The open scholarly article output by these main subject categories is shown in Figure 5. The difference by service is obvious. While open shares for mathematics, natural sciences, and technology (MNT) and for medicine and health sciences were both 70% for Google Scholar, the respective values for oaDOI were 32% and 42% for all open categories together. Compared to medicine and health sciences, oaDOI obviously seems not to cover MNT subjects adequately.

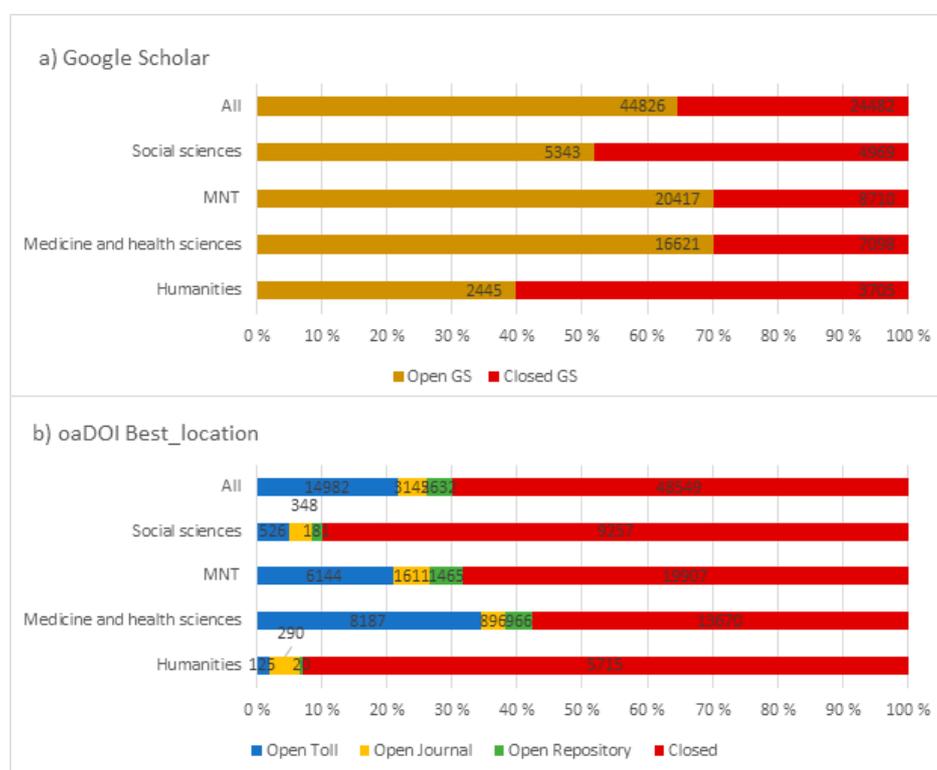


Figure 5. Open availability by subject (absolute numbers on bars) for (a) Google Scholar and (b) oaDOI Best_location for articles published between 2011 and 2015.

According to oaDOI in particular, Open Toll access was predominant for medicine and health sciences. However, these articles were foremost free to read-only, i.e., they are made open by the publishers without an open license. High open shares of read-only content were also reported by Martín-Martín et al. [6] for this discipline.

A closer look at what is open via Google Scholar only, i.e., closed according to oaDOI, reveals that the top providers for that closed content are researchgate.net and academia.edu (Table 4), confirming that ASNs are indeed important players in the field [6,7]. Because oaDOI excludes ASNs as full-text sources, these findings are not surprising. In addition, and more surprising, is that open repositories, in particular Norwegian repositories such as bibsys.no, uio.no, and diva-portal.org, are listed at the top of Table 4 as well. The same is true for arxiv.org and other recognized archives. From experience, we know that DOIs are not routinely added to these types of archives and, therefore, are not findable

by oaDOI. However, our Google Scholar searches were more sophisticated and identified additional articles, which again resulted in a higher recall.

Table 4. The numbers of full-text articles provided by Google Scholar and not provided by oaDOI (2011–2015). The first column lists the full-text providers according to Google Scholar’s best location.

| Provider | Humanities | Medicine and Health | MNT | Social Sciences | Total |
|---------------------|------------|---------------------|------|-----------------|-------|
| researchgate.net | 146 | 2766 | 4181 | 852 | 7945 |
| academia.edu | 284 | 694 | 1091 | 400 | 2469 |
| bibsys.no | 255 | 416 | 611 | 794 | 2076 |
| uio.no | 186 | 79 | 310 | 134 | 709 |
| psu.edu | 20 | 133 | 314 | 104 | 571 |
| diva-portal.org | 20 | 172 | 298 | 74 | 564 |
| acs.org | | 26 | 504 | | 530 |
| hio.no | 58 | 171 | 78 | 172 | 479 |
| arxiv.org | 1 | 2 | 447 | 1 | 451 |
| semanticscholar.org | 12 | 48 | 317 | 46 | 423 |
| uit.no | 104 | 107 | 134 | 72 | 417 |
| infona.pl | | 285 | 31 | 6 | 322 |
| uib.no | 42 | 64 | 153 | 45 | 304 |
| cambridge.org | 46 | 76 | 75 | 47 | 244 |
| ntnu.no | 2 | 5 | 166 | 32 | 205 |
| novus.no | 154 | | | 38 | 192 |
| bmj.com | | 189 | | | 189 |
| idunn.no | 33 | | | 120 | 153 |
| tandfonline.com | 8 | 75 | 20 | 18 | 121 |

3.5. 1findr

Finally, we briefly present some figures retrieved from 1findr. At the time of this study, we were able to retrieve 81% of all articles in Cristin (Table 5). Articles not retrieved, similarly to what we found for oaDOI, except for ASNs, belong to conference proceedings and to Norwegian journals (16,852 articles, 19%).

Table 5. Number of articles indexed by 1findr, 2011–2016.

| | Articles | % |
|-----------------------|----------|------|
| indexed by 1findr | 70,476 | 81% |
| not-indexed by 1findr | 16,852 | 19% |
| Total | 87,328 | 100% |

The overall openness score for 1findr was 52% (Figure 6). For type of open availability, as defined by the service (Table 3), the combined categories of “GO” and “GO;GR”, representing all gold OA articles, was the largest (21%). These findings are reasonable and are in agreement with our earlier findings on OA publishing according to DOAJ [22].

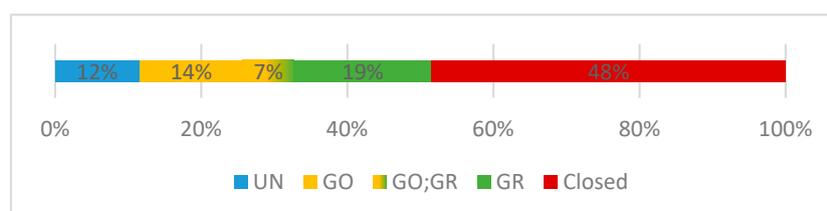


Figure 6. Open availability by 1findr for articles published between 2011 and 2016.

Because the categories given by 1findr and oaDOI are defined differently, they are not fully comparable. The largest difference is that 1findr considers articles via ASNs to be open GR articles.

4. Conclusions

In our study, we compared open availability by service and found that different services returned considerably different results. Google Scholar was by far the largest full-text provider. It linked 70% of the Norwegian article output to a free version, and corresponding shares for 1findr and oaDOI were 52% and 31%, respectively.

Different shares of openness are primarily caused by the services' different missions and what they choose to include in their results. For example, Google Scholar and 1findr include ASNs, which might contain material that is in violation of copyright, that has weak or unclear OA policies, or that does not guarantee persistent access. The dispute between publishers and ResearchGate proves the instability of these platforms as sources of full texts [25].

At the time of our investigation, the open share according to oaDOI was relatively low at 31%. Since then, oaDOI has changed its harvesting routines from mainly harvesting aggregated services to harvesting the original repositories directly, and oaDOI has reorganized its service under the brand Unpaywall. Our first tests seemed to have returned significantly higher recalls than reported in this study, which makes the service more relevant for further use and for build-upon services. We are also pleased to see that the service has added metadata on available document versions (submitted, accepted, or published). All of these measures are important and help the scholarly community to progress in building a true open-access infrastructure.

We only had a short look at 1findr. Apart from our concerns about including ASNs as full-text providers, 1findr returns reasonable open shares (52%). It is praiseworthy that the service recently opened their holdings for institutional use.

In future studies, it will be useful to consider OA policies implemented by funders because these mandates are likely to heavily influence the direction of OA. The European Commission has mandated OA to scientific articles in Horizon 2020 [26]. They list two ways to meet the requirements: (1) Self-archiving a peer-reviewed version of an article in a repository (with a maximum of a 6–12 month embargo) or (2) publishing an article as OA and then depositing the published version in a repository. These requirements are in line with the Norwegian Government's mandate and with mandates in other countries.

The requirement that all articles are to be made available through repositories ensures that articles are permanently OA. This will also make it easier to collect data on OA publishing and to monitor ongoing developments.

Although there is a rise in OA publishing and more publications are being made openly available, we have found that the available content is mainly free to read and not necessarily free to reuse and redistribute. Restriction on reuse hinders an effective research flow and harms the idea of open science. In the digital age and with the described services as important drivers towards OA, the time seems overdue to reshape the publishing scene.

Supplementary Materials: Data are openly available on BORA (<https://bora.uib.no/handle/1956/18308>). The code used to scrape data from Google Scholar is available on GitHub [17].

Author Contributions: Conceptualization: S.M., T.E.S., and Ø.L.G.; methodology: S.M. and Ø.L.G.; validation: S.M., T.E.S., and Ø.L.G.; formal analysis: S.M.; writing—original draft preparation: S.M.; writing—review and editing: T.E.S.; visualization: S.M.

Funding: This work is a result of our project Frie, vitenskapelige publikasjoner (Free, scholarly publications), funded by the National Library of Norway (Ref a68f29d9).

Acknowledgments: We thank Hemed Ali Al Ruwehy for some of the data preparation and Ingrid Cutler and Irene Eikefjord for reading and discussing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of the data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Ministry of Education and Research in Norway. National Goals and Guidelines for Open Access to Research Articles. Available online: <https://www.regjeringen.no/en/dokumenter/national-goals-and-guidelines-for-open-access-to-research-articles/id2567591/> (accessed on 22 July 2017).
2. Max Planck Digital Library. Open Access 2020. Available online: <https://oa2020.org/> (accessed on 20 August 2018).
3. Jamali, H.R. Copyright compliance and infringement in ResearchGate full-text journal articles. *Scientometrics* **2017**, *112*, 241–254. [CrossRef]
4. Greshake, B. Looking Into Pandora's Box: The Content Of Sci-Hub And Its Usage. *F1000Research* **2017**, *6*, 541. [CrossRef] [PubMed]
5. Archambault, É.; Caruso, J.; Nicol, A. State-of-Art Analysis of OA Strategies to Peer-Review Publications. 2014. Available online: http://www.science-metrix.com/sites/default/files/science-metrix/publications/d_2.1_sm_ec_dg-rtd_oa_policies_in_the_era_update_v05p_0.pdf (accessed on 30 May 2018).
6. Martín-Martín, A.; Costas, R.; van Leeuwen, T.; Delgado López-Cózar, E. Evidence of open access of scientific publications in Google Scholar: A large-scale analysis. *J. Informetr.* **2018**, *12*, 819–841. [CrossRef]
7. Mikki, S. Scholarly publications beyond pay-walls: increased citation advantage for open publishing. *Scientometrics* **2017**, *113*, 1529–1538. [CrossRef]
8. Piwowar, H.; Priem, J.; Larivière, V.; Alperin, J.P.; Matthias, L.; Norlander, B.; Farley, A.; West, J.; Haustein, S. The state of OA: A large-scale analysis of the prevalence and impact of Open Access articles. *PeerJ* **2018**, *6*, e4375. [CrossRef] [PubMed]
9. Bosman, J.M.; Kramer, B. Open access levels: A quantitative exploration using Web of Science and oaDOI data. *PeerJ Preprints* **2018**, *6*, e3520v1.
10. Chan, L.; Cuplinskas, D.; Eisen, M.; Friend, F.; Genova, Y.; Guédon, J.-C.; Hagemann, M.; Harnad, S.; Johnson, R.; Kupryte, R. Budapest Open Access Initiative. Available online: <http://www.budapestopenaccessinitiative.org/read> (accessed on 6 April 2018).
11. Suber, P. *Open Access*; MIT Press: Cambridge, MA, USA, 2012.
12. McCabe, M.J.; Snyder, C.M. Identifying the effect of Open Access on citations using a panel of scientific journals. *Econ. Inq.* **2014**, *52*, 1284–1300. [CrossRef]
13. Archambault, E.; Côté, G.; Struck, B.; Voorons, M. Research impact of paywalled versus open access papers. Science-Metrix and 1science 2016. Available online: <http://www.1science.com/oanumbr.html> (accessed on 7 November 2016).
14. Davis, P.M.; Walters, W.H. The impact of free access to the scientific literature: A review of recent research. *J. Med. Libr. Assoc.* **2011**, *99*, 208. [CrossRef] [PubMed]
15. Cristin. NVI-Reporting Instruction. Available online: <https://www.cristin.no/english/resources/reporting-instructions/#toc1> (accessed on 3 October 2018).
16. Orduna-Malea, E.; Ayllón, J.M.; Martín-Martín, A.; Delgado López-Cózar, E. Methods for estimating the size of Google Scholar. *Scientometrics* **2015**, *104*, 931–949. [CrossRef]
17. Gjesdal, Ø.L. Frie-Publikasjoner-Gs. Available online: <https://github.com/ubbdst/frie-publikasjoner-gs> (accessed on 2 October 2018).
18. Impactstory. Unpaywall. Available online: <https://unpaywall.org/> (accessed on 27 August 2018).
19. 1scienc. 1findr. Available online: <https://www.1science.com/1findr/> (accessed on 27 August 2018).
20. Posada, A.; Chen, G. Preliminary Findings: Rent Seeking by Elsevier. 2017. Available online: <http://knowledgegap.org/index.php/sub-projects/rent-seeking-and-financialization-of-the-academic-publishing-industry/preliminary-findings/> (accessed on 4 April 2018).
21. Det nasjonale Publiseringutvalget. NPI. Available online: <https://npi.nsd.no/fagfeltoversikt> (accessed on 13 April 2018).
22. Mikki, S.; Zygmuntowska, M. *Vitenskapelig publisering ved Universitetet i Bergen: Publiseringstatistikk 2005–2016*; Universitetsbiblioteket i Bergen: Bergen, Norway, 2017.
23. Laakso, M.J.S. Green open access policies of scholarly journal publishers: A study of what, when, and where self-archiving is allowed. *Scientometrics* **2014**, *99*, 475–494. [CrossRef]
24. Craig, I.D.; Plume, A.M.; McVeigh, M.E.; Pringle, J.; Amin, M. Do open access articles have greater citation impact?: a critical review of the literature. *J. Informetr.* **2007**, *1*, 239–248. [CrossRef]

25. Coalition for Responsible Sharing. Coalition Statement. Available online: <http://www.responsiblesharing.org/coalition-statement/> (accessed on 18 June 2018).
26. European Research Council. Guidelines on Implementation of Open Access to Scientific Publications and Research Data. (Horizon2020). Version 3.2. Available online: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf (accessed on 23 May 2017).



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).