David among Goliaths: open access publishing in scientific (semi)periphery

Bojan Macan, Ruđer Bošković Institute, Centre for scientific information, Bijenička cesta 54, 10000 Zagreb, Croatia; e-mail: bojan.macan@irb.hr; ORCID: 0000-0003-1138-4188

Lea Škorić*, University of Zagreb School of Medicine, Central Medical Library, Šalata 3, 10000 Zagreb, Croatia; e-mail: lea.skoric@mef.hr; ORCID: 0000-0002-1545-5791

Jelka Petrak, University of Zagreb School of Medicine, Central Medical Library, Šalata 3, 10000 Zagreb, Croatia; e-mail: jelka.petrak@mef.hr; ORCID: 0000-0003-1481-2598

Abstract

This case study analyses data on papers of Croatian authors published in 2017 from four Web of Science Core Collection citation indexes (SCI-EXP, SSCI, AHCI, and ESCI). The primary dataset (5,176 articles and reviews) was divided in two subsets, the OA subset (2,964 papers) and non-OA subset, (2,212 papers). We also used the primary dataset to create a subset of papers published in Croatian journals (1,588) as opposed to foreign ones. All were screened for full-text OA status, journal JCR quartile ranking, journal dominant discipline, and language of publication. OA papers prevailed with 74.4%. Most were available at publisher websites. The percentage of OA papers in Croatian journals was 99.8% and in foreign journals it was above the EU average (47%). The share of OA papers was highest in the humanities and social sciences, which also saw the highest share of papers in the Croatian language.

Key points:

- Three quarters of papers published by Croatian authors are available in open access.
- Most papers are in English, even in Croatian journals; Croatian prevailed only in the humanities.
- Croatian journals are mostly non-commercial and many receive government subsidies, provided they are OA; they generally do not charge for article processing.
- The prevalence of OA over non-OA papers is the highest in the humanities followed by the social sciences.
- Open access publishing might increase international visibility of journals from periphery or semi-periphery countries.
1. Introduction

The countries of the European Union (EU) produce one quarter of the world's scientific papers (Noorden & Butler, 2019), and, according to the European Commission (EC), 40.4% of them were freely accessible in 2017 (European Commission, 2020). The EC recommends (Commission Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information (No. 32018H0790), 2018) that all member states should adopt open access policies "to provide researchers and the public at large with access to peer-reviewed scientific publications, research data and other research outputs free of charge in an open and non-discriminatory manner as early as possible in the dissemination process, and enable the use and re-use of scientific research results". The policies should be particularly clear and detailed in national action plans about "dissemination of and open access to scientific publications resulting from publicly funded research." Between 2009 and 2018, Croatia had 50.8% OA papers and ranked third among the top European countries in this respect. Moreover, Croatia and the United Kingdom were the leading EU countries in the number of open-access papers (European Commission, 2020). Another large-scale study yielded similar results (Archambault et al., 2013).

According to its development indicators, Croatia belongs to the group of semi-peripheral countries (Sorinel, 2010). Domazet and Marinović Jerolimov (2014) describe European semi-periphery, especially its post-socialist countries, as "the countries with a high level of development, widespread educational and healthcare attainments, close to European cultural sphere". Croatia is still going through transitional changes in almost all of its economic, educational, social, and political life.

By joining the EU in 2013, Croatia has become its youngest member state. Its investment in research and development in 2017 was 0.85% of GDP, which is far below the EU average of 2.3% (The World Bank, 2019). However, the number of papers published by Croatian authors has steadily been growing. Between 2013 and 2017, this production was 31,853 documents in journals covered by the Web of Science Core Collection (WoSCC). Of these, 13,023 were available through open access. And while the increase in overall "output" in this 5-year period was 21%, the increase in the overall number of OA papers was as high as 58.7%. These increases were partly influenced by the changes in the WoSCC indexing policies and the introduction of Emerging Sources Citation Index (ESCI) (Testa, 2009; Web of Science: Emerging Sources Citation Index: Search and discover a trusted set of journals with comprehensive coverage by subject and region, 2020).

It has been a while since Guédon (2008) noted that OA journals would probably be crossing the divide between peripheral science, where they started, and "core" or "mainstream" science. He argued that international scientific competition had gradually led to a two-tier communication system, the international acting as the main quality arbiter and national reflecting national politics and priorities. He emphasized that Garfield's Science Citation Index had "a strategic role in the present power structure of world science", reflecting the dominance of mainstream journals. However, peripheral journals improved their position with the launch of the Emerging Sources Citation Index (ESCI), which ensured important research to be visible in WoSCC, "even if it is not yet internationally recognized" (What is the Emerging
Sources Citation Index?, 2020). For example, ESCI covers more than 50% of all Croatian journals indexed in WoSCC.

Kieć (2017) reported that authors from scientific periphery published in OA journals more often than authors from core countries, 46% vs 20% of papers, respectively, to be more exact. Archambault et al. (2013) found that publishing in gold OA journals was much more common in eastern Europe and proposed that it was because they preferred publishing in languages other than English. Croatia, however, does not fit that profile – most Croatian scientific journals, especially in the STEM fields, are published in English (Pulišelić & Petrak, 2006).

About 180 scientific and technical journals in Croatia are currently supported by public funds. Their publishers are mainly not-for-profit organisations (universities, research institutes, learned societies) under the patronage of the Croatian Ministry of Science & Education (MSE). Among the MSE’s most important criteria for subsidising a journal is that it is being indexed by at least one international bibliographic database and that it provides open access to complete journal’s content on the Hrčak repository portal (Republika Hrvatska, Ministarstvo znanosti i obrazovanja, 2018). Hrčak was launched in 2006 as an outlet for journals not having infrastructure for e-publishing and an interface for free access to all journal content (Stojanovski et al., 2009). As a platform Hrčak gave Croatian scientific journals better international visibility, which eventually increased the number of Croatian journals indexed in relevant international bibliographic databases. In 2007, Web of Science databases indexed only 16 Croatian scientific journals; in 2011, they indexed 61 (Macan et al., 2012), and their current number is 124 (Web of Science Master Journal List. Web of Science Master Journal List, 2020). In the meanwhile, again mainly because of Hrčak requirements, Croatian journals upgraded their editorial practices, which was the other main reason why so many are indexed in international bibliographic databases.

Piwowar et al. (2018) suggested further investigation to better understand what drives OA. The aim of this study was therefore to explore in depth the OA production of Croatian authors indexed by the WoSCC and published in 2017. We focused on OA papers, their disciplinary and quartile distributions, and the characteristics of Croatian OA journal publishing.

2. Methods

The primary sources of research data were the Clarivate Analytics WoSCC citation indexes: Science Citation Index Expanded (SCI-EXP), Social Sciences Citation Index (SSCI), Arts & Humanities Citation Index (AHCI), and ESCI. The initial search of WoSCC indexes was done in May 2019 using search query "Croatia OR Hrvatska" in the "Basic Search" address field, combined with 2017 in the Year Published field. The results were then filtered by document type to include only "article" and "review" papers. What remained was our primary dataset of 5,176 papers, which we used as the basis for all further analysis. Clarivate Analytics InCites was used for disciplinary distribution analysis as well as for quartile distribution (Q) of papers according to the journal’s impact factors (JIF). DOAJ and Hrčak were used to check the OA status of journals, especially the Croatian ones.

Our primary dataset was divided into three subsets according to the goals of our analysis:
I. **OA subset** comprised 2,964 papers. It combined two subsets:

- a subset of 2,574 papers obtained by refining the primary dataset by the WoS Open Access filter (all OA types according to the WoS typology were included) and
- a subset of 390 papers published in Croatian journals which were not marked in WOS as OA papers, but were identified manually by journal name search in Hrčak and DOAJ.

II. **Non-OA subset** included the remaining 2,212 papers.

- randomized sample: a random sample was created in order to check if there was any "hidden" OA paper within the non-OA subset. Randomisation was based on the expected OA proportion of 20%, 95% confidence interval, and type 1 error of 5%. The OA status of thus obtained 212 papers was checked manually using the article title as a Google Scholar search string. All retrieved papers were inspected to find respective OA full text.

III. **Croatian journals subset** consisted of 1,588 papers published by Croatian journals.

2.1. Data analysis

All papers were analyzed to determine OA to full text, disciplinary orientation and JCR quartile distribution of journals, journal’s country of origin, and language of publication (with focus on Croatia/Croatian).

Clarivate Analytics defines OA papers as "freely accessible peer reviewed versions of an article from either a publisher’s website or repository" (Clarivate Analytics, 2018). In WoSCC OA papers are grouped in six categories: DOAJ Gold, Other Gold, Bronze, Green Published, and Green Accepted. For the purpose of this study, we decided to use only two broad OA types: papers available at publisher websites (irrespective of their specific publishing model as gold, bronze, hybrid, delayed OA, open archives etc.) and those available on repositories, personal websites, and academic social networks (green OA). We did not verify the validity of WoS data.

Regarding the country of origin, the papers were divided in two groups: published in Croatian journals and published in foreign journals. There were three language groups: papers in English, papers in Croatian, and papers in other languages.

The records from WoSCC were transferred to the Clarivate Analytics InCites platform to determine discipline and JCR ranking. For disciplinary distribution of the papers we used the first level of OECD Category scheme, which corresponds to the Revised Field of Science and Technology (FOS) Classification of the Frascati Manual 2002 (OECD Publishing) (Directorate for science, technology and industry, Committee for scientific and technology policy, 2018). This classification consists of six major fields: natural sciences, engineering and technology, agricultural sciences, medical and health sciences, social sciences, and the humanities. This scheme is similar to the official Croatian classification of science (Nacionalno vijeće za znanost, 2009).
3. Results

3.1. Primary dataset

Our search of the four WoSCC citation indexes resulted in 5,176 papers with at least one Croatian affiliation. They were published in 117 Croatian and 2,018 foreign journals, most of them (87.2%) in the English language.

Three scientific fields dominated: natural sciences (37.7%), medical sciences (25.1%), and engineering & technology (21.2%) (Table 1).

3,768 (72.8%) of all papers in the primary dataset were published in JCR-ranked journals. Among them, 31.7% were published in Q1 journals, 25.7% in Q2, 18.8% in Q3, and 23.6% in Q4. Most papers in the Q1 category belong to natural sciences (63%). The distribution, however, varies greatly between Croatian and foreign journals (Table 1).

Papers published in journals covered by ESCI (23.7%) could not be distributed by the quartiles as IFs are not calculated for these journals. The same is true for the AHCI journals, which published 2.9% of papers. Journals covered by ESCI published 58.4% of all social sciences and 63.3% of all humanities papers (Table 1).

Table 1. Distribution of all papers and papers published in Croatian journals by scientific field, JCR quartile (where available), and ESCI

<table>
<thead>
<tr>
<th>OECD scientific field</th>
<th>All papers</th>
<th>Papers in Croatian journals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1</td>
<td>Q2</td>
</tr>
<tr>
<td>Natural sciences</td>
<td>753</td>
<td>548</td>
</tr>
<tr>
<td>Engineering and technology</td>
<td>305</td>
<td>195</td>
</tr>
<tr>
<td>Medical and health sciences</td>
<td>344</td>
<td>291</td>
</tr>
<tr>
<td>Agricultural sciences</td>
<td>83</td>
<td>51</td>
</tr>
<tr>
<td>Social sciences</td>
<td>51</td>
<td>85</td>
</tr>
<tr>
<td>Humanities</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Number of unique papers</td>
<td>1195</td>
<td>974</td>
</tr>
</tbody>
</table>

3.2. OA subset

The OA subset accounted for 57.3% (n=2,964) of all retrieved papers. Most full texts were available at publishers' web-sites (94.2%). Many of them also had their "green" versions available in repositories, academic social networks, and similar platforms (38.5%). 1,585 (53.5%) papers in this subset were published in Croatian journals, and all of them were available on publishers' web-sites.
The OA subset shows a disciplinary distribution similar to the primary dataset, with natural sciences having the greatest number of papers (30.2%) (Figure 1).

34.8% of papers were published in journals covered by ESCI, followed by the papers published in Q4 journals (20.6%) and Q1 journals (17.2%). Altogether, 60.9% of OA papers (n=1,806) were published in JCR IF journals (Figure 2).

**Figure 1.** Disciplinary distribution of papers within primary dataset and its subsets
3.3. Non-OA subset

The non-OA subset consisted of 2,212 papers. Only three were published in Croatian journals. Almost half of the papers (47.8%) in the non-OA subset belonged to the field of natural sciences and only 4.5% to the humanities (Figure 1).

Altogether, 88.7% (n=1,962) of papers in the non-OA subset were published in JCR IF journals (Figure 2). Most were published in Q1 journals (30.9%), followed by those in Q2 journals (25.8%), while only 8.8% were published in ESCI journals.

3.3.1. Non-OA randomized sample

Our analysis of the randomized sample (n=212) in the non-OA subset discovered that 85 papers were actually OA (40.1%). 20 had free full text access on the publisher’s web pages, while 65 were available in repositories (39 papers) or academic/social networks (26 papers). If we apply the 40.1% to the entire non-OA subset, as many as 887 papers in the non-OA subset could actually be OA. If we add these 887 papers to the number from the OA subset, the overall share of OA in the primary dataset rises to 74.4% (Figure 3).
Most of "green-only" papers were published in Q1 and Q2 journals (45 out of 65), which suggests the importance of repositories, social networks, and similar platforms for papers published in the most prestigious paywall journals with high JIF.

### 3.4. Croatian journals subset

Of all primary dataset papers 30.7% (n=1,588) were published in 117 Croatian journals. Only two Croatian journals in the primary dataset had a paywall (were non-OA), and they published three papers that met our search criteria. The remaining 115 journals publish in open access, most without charging for article processing. This is why 99.8% of all papers in Croatian journals were OA (n=1,585). They accounted for 53.5% of the papers in the OA subset and for 56.8% of those published on the publisher’s web pages.

Further subset analysis showed that 68 of the 117 journals were indexed by DOAJ, and that 91 used one of the Creative Commons licenses (most frequently CC BY), but most of them did not show this information in the article (e.g. in headers or footers). Twelve more journals did not refer specifically to the CC licence but allowed use of full texts for non-profit purposes in the journal’s policy section. 23 journals didn’t assign DOI to the papers published in 2017.

The OA status of 390 papers published in 34 Croatian journals was not determined in WoS. The main reason may be that 33 of those journals were not registered in DOAJ in 2017, and many of them had unclear copyright and OA policies.

More than half (52.8%) of papers were published in Croatian journals covered by ESCI, followed by 30.7% published in Q4 journals. Only seven papers were published in the three Croatian journals ranked Q1 (Table 1).
Disciplinary distribution showed that only 8.5% of all natural sciences papers and 15.6% of all medical & health sciences papers were published in Croatian journals. In contrast, Croatian journals published 57.1% of all social sciences and 73.5% of all humanities papers from the primary dataset.

Just over a third (36.5%) of papers in this subset were published in Croatian language (11.2% of the primary dataset), most of which in the humanities and social sciences. Croatian is a predominant language in humanities papers across all data sets: 82.2% in Croatian journals, 75.4% in the OA subset, and 60.4% in the primary dataset. In contrast, only three medical papers were published in Croatian (Figure 4).

![Figure 4. Distribution of papers in Croatian journals across scientific fields by the publication language](image)

4. Discussion and conclusions

Our findings show that 74.4% of papers with at least one Croatian affiliation are OA, which ranks Croatia shoulder to shoulder with the leading OA nations – Brasil, Netherlands, Switzerland, and United Kingdom – with more than 65% of OA in their overall WoS indexed publication output (Science-Metrix, 2018). For a country on the scientific (semi)periphery with relatively scarce financial and infrastructural research resources, open access to research output is an effective tool for increasing international visibility and research collaboration on the international scientific scene dominated by more productive and influential nations.

In his survey comparing OA publishing by authors from core and periphery countries Kieńć (2017) argues that most journals in periphery countries are owned by academic societies and universities and are run as non-profit initiatives, often on public money. This is also the case with the majority of Croatian journals which publish in OA. They operate on public money and can pursue OA policy without article processing.
charges. Namely, OA is mandated if the journal is to receive government support/subsidy. However, the government has not issued an OA policy mandating publishing publicly funded research in OA outlets. This has been left to individual research institutions to decide, and only a few have imposed an OA mandate so far (Macan, 2018). Many Croatian researchers signed the Croatian Open Access Declaration in 2012 "inviting the state administration, headed by the ministry responsible for science, as well as scientific and educational institutions, organisations, professional associations, and all the others involved in gathering and publishing scientific information to act decisively and in coordination in order to store all the Croatian scientific information in open access form" (Budin et al., 2012). Until now, Croatia has not produced a universal strategy that would encourage its researchers to publish all their research OA.

As Kieńc explains it, "...for authors from the global periphery OA is the least important factor in choosing a journal to publish work in, while journal impact factor and abstracting services are the most important and that these two factors favour ‘mainstream’ journals that are usually published in the core countries" (Kieńć, 2017). Similar was reported by a Taylor & Francis 2019 researcher survey (Taylor & Francis researcher survey 2019: Researchers’ views on the present and future of scholarly communication, 2019) – the availability of open options was less important to authors than the reputation or impact factor of the journal. This culture of "publish in high-profile journals or perish" (Wang et al., 2004) is prevalent in Croatia too. Our results show that 23.1% and 18.8% of all papers were published in Q1 and Q2 journals, respectively. However, the share of papers indexed by ESCI is the largest (23.7%), and this is where the prevalence of OA papers is the highest. The same goes for the Q4 journals. These results also correspond to the number of papers published in Croatian journals, which are mostly indexed by ESCI or belong to Q4.

Croatian journals account for 30.7% of all papers published by Croatian authors, and for 52.8% of all OA papers. They seem to be an important communication and visibility channel, if they are covered by major international bibliographic databases. Almost 60% (73 titles) are covered by ESCI, which means that they have not yet been included in major citation indexes such as SCI-EXP, SSCI, and/or AHCI (Web of Science Master Journal List. Web of Science Master Journal List, 2020) but have a fair chance to join the international community of those that are included.

The prevalence of OA over non-OA papers is the highest in the humanities (80.1%), followed by the social sciences (71%). Our results are opposite to Science-Metrix reports showing that, on the global scale, the humanities have the smallest percentage of OA publications (24%) followed by the economic and social sciences (44%) (Science-Metrix, 2018). More than half of the OA papers published by Croatian journals come from social sciences (28.3%) and the humanities (23%). This reflects the distribution of Croatian journals indexed in WoS by disciplines; because more than half of all WoSCC indexed Croatian journals are thematically oriented to social sciences (24) and to the arts & humanities (34). Kulczycki (2018) suggests that publication patterns in SSH in non-English speaking European countries not only depend on the discipline and size of the scientific community, but also on its cultural and historical heritage. The prevalence of SSH papers in Croatian journals indexed in WoSCC may reflect the importance of local social, historical and linguistic phenomena which have not yet been sufficiently explored. The SSH journals are considered the national identity factor (Matanović & Berbić Kolar, 2018), crucial for the survival of the
national culture and language (Zrinščak, 2011) and not only for the communication of the SSH research results. The prevalence of local topics in SSH papers is also supported by the number of papers published in the Croatian language: 82.2% of papers in the humanities and 43.6% in social sciences as opposed to only 1.5% in medical sciences and 8.7% in engineering and technology. Publishing in local journals in Croatian has also been encouraged by the national criteria for academic advancement in the fields of social sciences and humanities (Nacionalno vijeće za znanost, visoko obrazovanj ei tehnološki razvoj, 2017). Nevertheless, many Croatian journals have shifted to English to expand their audience and improve coverage by relevant bibliographic databases (Pulišelić & Petrak, 2006). In fact, 61.5% of papers from our Croatian journal subset have been published in English.

Why Croatian authors opt for publishing in OA journals is yet to be explored. Even though we did not study this particular issue, we can assume that papers published OA in foreign journals (which usually charge for article processing) are financed by international scientific projects, whereas with Croatian journals, this choice is mandated by the journal OA policy, which stems from public subsidies rather than author preference.

The analysis of our randomized sample indicates the importance of national or institutional repositories and academic/social networks that constitute the so-called "green" OA. Most of these "green-only" papers were published in Q1 and Q2 journals (n=45; 71.4%), which implies the importance of the green route for papers published in the most prestigious paywall journals with high JIF. Until recently, the awareness of the need for OA digital repositories among Croatian research and high education institutions was rather low, and their implementation and maintenance was a big technical challenge. Only researchers working on international competitive projects (such as the ones funded by the EC through the Horizon 2020 framework) were mandated to archive their papers into an OA repository or a similar infrastructure. The initiative for the development of a common national digital repository infrastructure resulted in the establishment of Digital Academic Archives and Repositories – DABAR (URL: https://dabar.srce.hr), a centrally implemented technological solution based on open-source software. It provides Croatian researchers free, reliable, and interoperable digital repository infrastructure (Macan & Petrak, 2019). Both infrastructures have been financed by public funds.

Croatian journals account for more than half of the Croatian OA output. Their publishers are mainly not-for-profit organizations. For most of them, journal publishing is a part of their public mission. Publishing in OA without restrictions nor for authors neither for readers it’s a good way to accomplish that mission. However, such OA programs require considerable investments. The Croatian example confirms that publicly funded support for scientific publishing on a non-profit basis (with hosting arrangements such as Hrčak or DABAR) greatly improves dissemination and access to the scientific research results as well as the chance of their international visibility.
5. References:


*Web of Science: Emerging Sources Citation Index: Search and discover a trusted set of journals with comprehensive coverage by subject and region*. (2020). Web of Science Group. Retrieved from https://clarivate.com/webofsciencegroup/solutions/webofscience-esci/


*What is the Emerging Sources Citation Index?* (2020). Editor Resources: Supporting Taylor & Francis Journal Editors. Retrieved from https://editorresources.taylorandfrancis.com/understanding-research-metrics/esci/