

Supporting Open Access nationwide

Jadranka Stojanovski

jadranka.stojanovski@irb.hr

University of Zadar, LIS Dept, F. Tuđmana 24i, HR-23000 Zadar (Croatia)
Ruđer Bošković Institute, Library, Bijenička c. 54, HR-10000 Zagreb (Croatia)

Abstract

To support Croatian scholarly publishing environment, and inspired by global open access movement, the portal of Croatian scientific journals HRČAK (<http://hrcak.srce.hr>) was introduced in 2006 offering an open access publishing platform for Croatian journals. Today, HRČAK gathers about 290 scholarly and professional Croatian journals. This paper is focused on the currency and visibility of the journals included in HRČAK, giving accurate statistical data about HRČAK repository, its growth and development. Collaboration with Croatian publishers, namely those are mainly academic and research institutions or professional societies, on the continuous work of raising the quality of Croatian scientific journals is presented in this paper. HRČAK journals are available for harvesting using OAI-PMH protocol and papers are distributed through many different repositories, archives, databases and search engines. The future plans include work on full-text documents, inclusion of the additional types of publications and formats, harvesting process improvements, additional functionalities and standardization.

Introduction

In the era of ACTA¹, PIPA², SOPA³, RWA⁴ and other threats to the open access movement it is important to recognize specific practices which are present in the scholarly publishing environment, and to be aware of the present changes and future trends. Today's research is interdisciplinary and collaborative, requiring fast and efficient information and data exchange, as well as full and open access to published research. But predominant part of scholarly publications is still in hands of publishers with profit oriented business model. Although today the average researcher at the university or research institute has access to a much broader range of scholarly literature than even before in the print era, open access (OA) movement promoting access to articles without any restrictions posed by subscriptions is more important than ever (Björk et al., 2010). OA journals are more visible and readable and that's why comparison of OA and NOA articles showed that OA articles have consistently more citations, varying from 36%-172% by discipline and year (Hajjem, Harnad, & Gingras, 2005).

The most important step in bringing research from small and developing scientific communities to the mainstream science was the inclusion of the journal in international bibliographic databases and making the journal available in free full text on the Internet (Marušić, Sambunjak, & Marušić, 2006, page 372). It is also a widely accepted opinion that publicly funded research data should be openly available. This statement is supported in the official science & technology policy issued by the Croatian Ministry of Science, Education and Sports (MSES) where is specified that results of research "...financed by public resources must be accessible to the public in the form of open publications or databases" (Stojanovski, Petrak, & Macan, 2009), and taken as the base for the promotion of the OA movement in Croatia.

¹ Anti-Counterfeiting Trade Agreement

² Protect IP Act

³ Stop Online Piracy Act

⁴ Research Works Act

The transformation of Croatian scientific journals from print to digital started in late nineties and at very slow pace. In 2002 there were only 21 Croatian scientific journal with online version and full-articles available (Konjević, 2003). Publishers were not ready for evolutionary changes in publication process for many reasons: lack of awareness about existing trends, poor ICT equipment, low level of computer literacy, lack of technical support, etc. To support Croatian scholarly publishing environment, and inspired by global open access movement, the portal of Croatian scientific journals HRČAK (<http://hrcak.srce.hr>) was introduced in 2006 offering OA publishing platform for Croatian journals. The abbreviation HRČAK (hamster in English language) is coming from the phrase “HRvatski ČAsopisi” (Croatian Journals in English language). HRČAK was conceptualized and designed by a small group of librarians, researchers and information specialists, and created in cooperation with programmers at University Computing Centre SRCE, with a MSES support. Primary intention of HRČAK team was to offer free building tool for digital versions of the Croatian journals, and from the very beginning open access was promoted as the most appropriate model for scholarly publishing in Croatia.

There are two widely accepted routes to OA: digital repositories or archives (“green”) and OA journals (“gold”), often considered as two parallel tracks. “Gold” journals are available to their readers free of charge, and “green” uses online repositories to store digital duplicates of published research output and make them freely and openly available across the Internet (Hubbard, 2011). “Gold” applies only to publication through publishers, while “green” relates to self-archiving. Although this classification caused often confusion (Murray-Rust, 2011) it is widely adopted in OA community. However, it has recently become apparent that there is potential for repositories and journals to interact with each other on an ongoing basis and between them to form a coherent OA scholarly communication system (Pinfield, 2009). In Croatia, the majority of journals followed “golden” open access road providing digital journals, including peer-reviewed articles freely available, and HRČAK version of an online journal is the only digital version of the journal. But some journals have their online version independent from the HRČAK platform, and editorials are submitting articles in HRČAK as additional copy, and this part can be considered as “green”, although it is not submitted by author.

The main goal of HRČAK was to provide a simple and free interface for creating digital online version of Croatian journals, technical support for publishers, maintenance and long-term data protection and management. With time visibility of the Croatian scientific journals was improved, and research in Croatia in general become one of the priorities. Visibility can be measured by different criteria: number of visits, usage of Open Access Initiative – Protocol for Metadata Harvesting (OAI-PMH) for sharing data, data traffic, replication of bibliographic content by online databases, number of citations, etc. Research question in this paper is to explore in what extent can analysis of different data sets reflect visibility of HRČAK portal, including journals and their content?

Data and methods

In this paper the visibility of the journals included in HRČAK is analysed and discussed, giving accurate statistical data about HRČAK repository, its growth and development. The descriptive data analysis and interpretation was used in this study implying a simple quantitative summary of data sets that have been collected from different sources. The data about journals, their size, number of full-text articles stored, were collected from HRČAK database. The data about traffic, visits, and harvesters were extracted from a Web server log files using Webalizer, program that looks at Web server logs and produces detailed reports accessible online. The data about harvesters using OAI-PMH protocol were analysed

separately from the usage statistics representing visits to regular Web pages at <http://hrcak.srce.hr>, from which usage of the editorials and staff was excluded. Scopus (Elsevier), Web of Science - WoS (Thomson Reuters) and Journal Citation Reports – JCR (Thomson Reuters) databases were consulted to get the data about indexed journals. All data were gathered and analysed in the period March-May 2012.

Results and discussion

There are about 220 scientific and professional journals in Croatia publicly funded and mostly published by academic and research institutions or professional societies, half of them coming from social sciences and humanities. According 2007 study the median grant for journals by MSES was 46.000 HRK, e.g. approximately 6.200 EUR (Macan & Stojanovski, 2008). To get a grant from MSES publishers are evaluated by several criteria: publishers should be not-for-profit organisations, the journal should be issued regularly, the articles should pass peer review process, journal content should be focused more on original scientific papers, certain number of journal papers should have corresponding author from foreign countries, the journal should be in OA, the journal should be indexed in relevant international bibliographic databases, publishers should use journal management and publishing system, the journal should be edited according *Guidelines for editing and formatting journals*⁵, the journals should have cultural, national and local importance for scientific community, etc.⁶

The number of journals included in HRČAK increased steadily from 2006 till 2012, and in May 2012 HRČAK gathered 292 scholarly and professional Croatian journals. Distribution of the journals by discipline represented the majority of journals coming from humanities and social sciences (33 percent and 28 percent accordingly), which correspond to the number of journals supported by MSES in these disciplines (Stojanovski et al., 2009). Social sciences and humanities journals are closely oriented to the Croatian language, culture and history. Other disciplines are represented by 7 to 11 percent (Figure 1) also following the proportion of the total number of journals published in these disciplines. Croatian researchers from sciences, biomedicine and health, technical sciences and biotechnical sciences publish also in international journals, and that's why there is no need for lot of domestic journals in those scientific fields. Also advancement and evaluation criteria motivate particularly science researchers to publish in more prominent international journals.

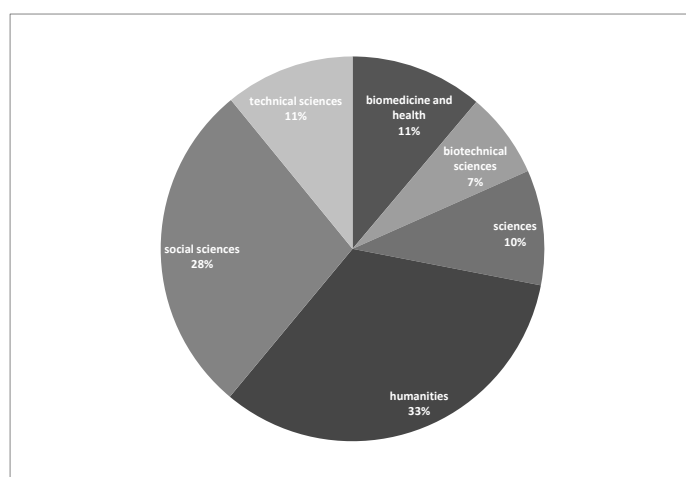


Figure 1. Distribution of journals included in HRČAK portal by discipline

⁵ <http://public.mzos.hr/Default.aspx?art=9110&sec=3202>

⁶ Guidelines about financial support of scholarly publishing for 2010 issued by the Croatian Ministry of Science, Education and Sport.

HRČAK growth was impressive going much beyond team expectations. The majority of the publishers accepted immediately HRČAK portal after introduction of OA approach, also visible data about journal/issue/paper visits and downloads by HRČAK played significant role in the further self-promotion. Very soon after the first release, without any systematic promotional activities, HRČAK portal become the most popular place for accessing the scholarly content. There are many reasons for the HRČAK success and some of them are (not ordered by importance):

- readers are more likely to access scholarly content through single portal instead of trying to find and select the content of interest from many different websites;
- free infrastructure and support from HRČAK team were encouraging for many Croatian publishers unable to publish digital version of their journals by own resources;
- credibility and competences of HRČAK team and University Computing Centre SRCE built a trustworthy relationship between Croatian publishers and HRČAK team from the very beginning;
- MSES, which supports majority of the journals, shared the main HRČAK goals such as online environment, open access, increased visibility, user-centred approach, and advancements in editorial work;
- HRČAK platform is used for collaboration and cooperation, establishing a dialog and empowering small publishers' position;
- regular statistical reports with data about repository growth, number of visits, data traffic, etc.

In May 2011 for 292 included journals there were 74.404 full-text articles stored in HRČAK repository. The size of the repository is 89.2 GB: 85 GB for file system and 4.2 GB for database. Some publishers provided the content of their journals since volume 1, using HRČAK as the place to archive their complete collections.

Data traffic in 2011 according Webalizer was 5,94 TB in total, with average traffic of 465 GB per month (Table 1), and traffic growth is following content growth. HRČAK is in top ten world repositories, according to Ranking of Web of World Repositories⁷. Data about traffic done by OAI harvesters are excluded from the statistics.

Table 1. Traffic in the period 2006-2011

<i>year</i>	<i>traffic (GB)</i>
2006	116.48
2007	1259.89
2008	2307.23
2009	3733.34
2010	5589.74
2011	5953.98

In the time period 2006-2011 top number of visits was in 2010 (Figure 2). During 2011 5.421.200 visits were recorded, with average of 451.767 visits per month, where visit is defined as any number of requests from the same IP address processed during a period of 30 minutes. Server processed in total 51.011.881 requests (hits) during 2011, and the most busiest months were October and November (Figure 3).

⁷ <http://repositories.webometrics.info/>

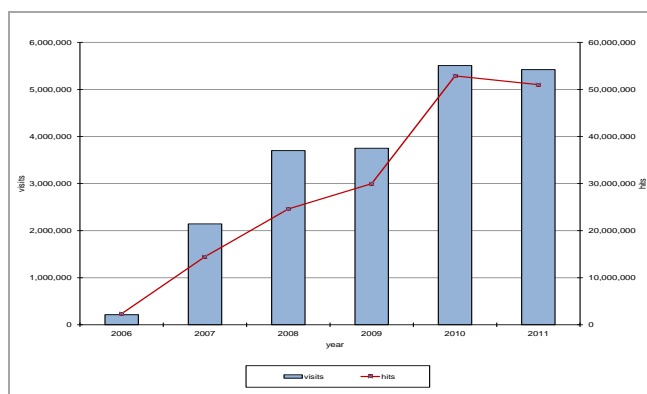


Figure 2. Visits and hits by year 2006-2011

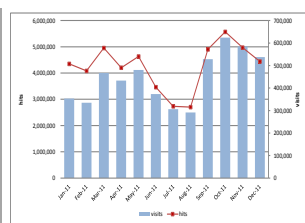


Figure 3. Visits and hits by month 2011

HRČAK journals are available for harvesting using OAI-PMH protocol⁸, RepositoryIdentifier: hrcak.hr; SampleIdentifier: oai:hrcak.hr:25) and papers are distributed through many different repositories, archives, databases and search engines. In 2011 the most active harvesters with resolved IP addresses were: Google Scholar, IFREMER, Yahoo! Research, Sakura Internet (Japan), Long Island University, OCLC, Bielefeld University, Scientific Commons Project, QSensei, Purdue University Libraries, Università degli Studi di Firenze, Physikalisch-Technische Bundesanstalt, University of Southampton, School of Electronics and Computer Science and Staatsbibliothek zu Berlin.

Harvesters visited HRČAK on average 2508 times per month in 2011, and 3543 times per month during the first four months in 2012, which is significant increase. The majority of harvesters are located inside US commercial domain, Japan, France, US, Croatia, South Africa, Swiss, Bosnia and Herzegovina and Germany. A lot of IP addresses remain unresolved, and even for known domains it was difficult to reveal in which repository, database, archive or catalogue the harvested data were included. Important information about a crawler, shared during “crawler introduction”, can improve the data exchange and should become a standard.

Besides disseminating the data about Croatian scientific and professional journals and published papers through OAI-PMH, HRČAK team is involved in the process of the inclusion of Croatian journals in most prominent free and free databases like Web of Science (WoS), Scopus, Google Scholar, EBSCO databases, DOAJ, Google Scholar, OAIster, BASE etc. Although HRČAK is OA portal exchanging the data with OA databases, archives, repositories and other digital collections primarily, inclusion into popular, even commercial products is improving visibility of published papers significantly. For articles indexed in citation databases the number of citation can be recorded which can be used for evaluation and advancement purposes. In 2007 there were only 16 Croatian journals included in WoS (Macan & Stojanovski, 2008). To satisfy inclusion criteria publishers usually need to improve some elements of the publishing process and to follow professional standards for editing more strictly. HRČAK as collaboration tool supports publishers during this process in different ways. HRČAK team provides Croatian publishers with contacts, links to selection/inclusion criteria and help with their applications. These efforts together with WoS expansion of the content coverage resulted in more than 60 Croatian journals (54 HRČAK) indexed in Web of Science in 2011. Scopus database, launched in 2004, systematically included Croatian journals from the beginning and currently indexing 105 Croatian journals (90 HRČAK). From 292 journals included in HRČAK 31.5 percent (92 titles) are indexed in WoS or Scopus

⁸ <http://hrcak.srce.hr/oai>

databases: 48 journals are indexed in both databases, 6 journals are indexed only in WoS, and 42 journals are indexed only in Scopus. HRČAK journals included in WoS are distributed evenly across disciplines: natural sciences (11), biomedicine (8), technical sciences (11), biotechnology (5), social sciences (9), and humanities (10). According JCR there are 35 Croatian journals with impact factor (IF), and only 4 of them are non-HRČAK journals. Best Croatian journals according IF are included in HRČAK, and first on the ranked list is *Croatian Medical Journal* (IF 1.455). Two non-HRČAK journals with relatively high IF are *Operators and Matrices* (IF 0.542) and *Mathematical Inequalities and Applications* (IF 0.524). Both journals have their web sites and articles can be accessed by pay per view/download (advanced payment of 10 EUR per article is charged) or with login for journal subscribers. Other two non-HRČAK journals, *Transactions of FAMENA* (IF 0.143) and *Neurologia Croatica* (IF 0.125), have web pages with general information about journal and published articles, but full-text articles are available only by print subscription. Other Croatian journals included in WoS are still not listed in JCR, since minimum of two year of presence in WoS is needed for IF calculation. In the cooperation with database publishers and aggregators HRČAK team is working together with their technical teams on data harvesting using OAI-PMH protocol instead of traditional ways of data gathering. Croatian publishers are using HRČAK mailing list to exchange their experiences which was recognized as a good motivation tool. The popular search engines and databases that uses OAI-PMH on a regular basis are Google Scholar (GS) where all HRČAK journals and papers are indexed and searchable, OAIster (OCLC) and BASE (Bielefeld).

To support the whole publishing process open source journal management system OJS - Open Journals System⁹ was introduced in 2008 to the Croatian publishers, and the data synchronisation with HRČAK repository has been established. For publishers without own servers there is central version of OJS ver. 2.2.2.0. installed at SRCE¹⁰, and also local OJS installations are supported. At the present about twenty publishers are using OJS for the automation of their publishing process, from paper submission and peer review to the selection of papers and online journal publishing. Although OJS has been designed to help enhance high quality standard for journal publishing, reduce cost, and provide efficient editorial processes (Willinski J. according Kwaku Smit Esseh, 2006) for majority of Croatian publishers it encompasses redefinition of the existing editorial process. This could be the reason for relatively small number of publisher who accepted OJS as a tool for journal management. To educate, share opinions and to meet editors' requirements, workshops and round tables for discussions about current topics are regularly organized.

Conclusions

The analysis in this paper revealed growth in all aspects of HRČAK portal: number of included journals, content, usage statistics, and exchange of data, proving that OA movement find fertile ground inside small Croatian publishing environment. One of the reasons for HRČAK big success is coherence with MSES criteria for journals grants, including promotion of OA, usage of journal management system and improved visibility of scholarly publishing in Croatia through exchange of data and journal indexing in popular databases. Editorials and authors can also track data about number of visits and downloads available on the journal, issue and article level, measuring their popularity.

There is still a lot of space for improvements, especially regarding interoperability and data exchange with other national and international information services. The future plans include

⁹ <http://pkp.sfu.ca/ojs>

¹⁰ <http://hrcak.srce.hr/ojs>

work on existing full-text documents which are at the present only partially searchable, inclusion of the additional types of publications, formats and research data gathering, harvesting process improvements, additional search functionalities, articles in press, website redesign and more standardization on different levels. Metadata description can be also improved. Although scientific journals in Croatia are recently suffering from serious budget cuts, this could be considered also as new challenge leading to further modernisation and improvements, and leaving less space for unnecessary expenditures related with traditional publishing processes. The collaboration between publishers could be still improved, and more educational workshops and meetings should be organized. There is also a strong will to broaden HRČAK coverage with conference proceedings and to include OA journals from other neighbouring countries.

Literature:

- Björk, B.-C., Welling, P., Laakso, M., Majlender, P., Hedlund, T., & Gudnason, G. (2010). Open access to the scientific journal literature: situation 2009. *PloS one*, 5(6), e11273. doi:10.1371/journal.pone.0011273
- Hajjem, C., Harnad, S., & Gingras, Y. (2005). Ten-Year Cross-Disciplinary Comparison of the Growth of Open Access and How it Increases Research Citation Impact. *IEEE Data Engineering Bulletin*, 28(4), 39-47.
- Hubbard, B. (2011). Green, Blue, Yellow, White & Gold. Retrieved from <http://www.sherpa.ac.uk/documents/sherpaplusdocs/Nottingham-colour-guide.pdf>
- Konjević, S. (2003). Hrvatski znanstveni časopisi na Internetu. *Vijesnik bibliotekara Hrvatske*, 4, 111-118. Retrieved from <http://eprints.rclis.org/bitstream/10760/5598/1/Konjevic.pdf>
- Kwaku Smit Esseh, S. (2006). *The open access (Open Journal Systems) paradigm and the production of scholarly journals in developing countries*. Retrieved from summit.sfu.ca/system/files/iritems1/2371/etd2183.pdf
- Macan, B., & Stojanovski, J. (2008). Analiza novčane potpore Ministarstva znanosti, obrazovanja i športa hrvatskim znanstvenim časopisima. *Kemija u industriji*, 57(3), 115-122. Retrieved from <http://hrcak.srce.hr/file/32964>
- Marušić, M., Sambunjak, D., & Marušić, A. (2006). Life of small medical journal--how bibliographical indexing and international visibility affected editorial work in Croatian Medical Journal. *Croatian medical journal*, 47(3), 372-5. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2080425&tool=pmcentrez&rendertype=abstract>
- Murray-Rust, P. (2011). Green and gold open access? Libre and gratis. Reasons why readers and re-users matter. Retrieved April 12, 2012, from <http://blogs.ch.cam.ac.uk/pmr/2011/07/19/green-and-gold-open-access-libre-and-gratis-reasons-why-readers-and-re-users-matter/>

Pinfield, S. (2009). Journals and repositories: an evolving relationship? *Learned Publishing*, 22(3), 165-175. doi:10.1087/2009302

Stojanovski, J., Petrak, J., & Macan, B. (2009). The Croatian national open access journal platform. *Learned Publishing*, 22(4), 263-273. doi:10.1087/20090402

ACKNOWLEDGEMENTS. I wish to thank colleagues at University Computing Centre SRCE and their leader Miroslav Milinović for providing the data about traffic and usage statistics.