



# Archival description using UNIMARC: Presidential Library experience

## Opis arhivskega gradiva z uporabo formata UNIMARC na primeru Predsedniške knjižnice v Sankt Peterburgu

**Olga Zhlobinskaya<sup>1</sup>**

**ABSTRACT:** The paper presents basic approaches to managing archival resources in the Presidential library (Saint-Petersburg, Russia). All resources in the collection are described using the RUSMARC format (national implementation of UNIMARC), which provides sufficient instruments to keep the hierarchical structure of archival description and to give detailed characteristics of the materials. The workflow of creating RUSMARC records for archives is presented in detail, including description from the general to the specific, using only level-appropriate data to avoid redundancy, creating name access points, indexing with Library Bibliographic Classification class numbers and subject headings according to the policy developed by the Presidential Library. Enhanced users' search interface is also presented; it includes both traditional library and archival resources with specific options for each type of resource. The conclusion is made that co-existence of both types of resources in the library catalogue using RUSMARC (UNIMARC) made the catalogue more efficient and user-friendly.

**KEYWORDS:** Presidential Library, archival resources, cataloguing, bibliographic description, RUSMARC, UNIMARC, access points, search interface

**IZVLEČEK:** Članek predstavlja osnovne pristope k upravljanju arhivskih virov v Predsedniški knjižnici (St. Peterburg, Rusija). Vsi viri v knjižnični zbirki so opisani z uporabo formata RUSMARC (nacionalna različica UNIMARC-a), ki zagotavlja dovolj možnosti za ohranitev hierarhične strukture pri opisu arhivskega gradiva in omogoča navedbo podrobnih značilnosti gradiva. Podrobno je predstavljen potek dela pri kreiranju zapisov RUSMARC za arhivske vire, vključno z opisom od splošnega k specifičnemu ob uporabi zgolj podatkov, ki ustrezajo posameznemu nivoju, da bi se s tem izognili nepotrebnim podatkom, ter s kreiranjem točk dostopa za imena, indeksiranjem na osnovi klasifikacijskih števil Knjižnične bibliografske klasifikacije (Library Bibliographic Classification) in predmetnih oznak v skladu s politiko, ki jo je oblikovala Predsedniška knjižnica. Predstavljen je tudi izboljšani uporabniški vmesnik za iskanje, ki vključuje tako običajne knjižnične vire kot tudi arhivske vire z možnostmi, specifičnimi za vsakega od obeh skupin virov. Zaključek je, da je zaradi vključenosti obeh skupin virov v knjižničnem katalogu na osnovi uporabe formata RUSMARC (UNIMARC) katalog učinkovitejši in uporabniku prijaznejši.

**KLJUČNE BESEDE:** Predsedniška knjižnica, arhivski viri, katalogizacija, bibliografski opis, RUSMARC, UNIMARC, točke dostopa, iskalni vmesniki

## 1 Introduction

The Presidential Library in Saint-Petersburg, one of Russia's three national libraries, was established in 2009 as a nationwide repository of digital materials either born-digital or digitized copies of the most important documents on the history, theory and practice of Russian statehood and Russian language.

The first library collection was based on digitized materials of the Russian State Historical Archive. The main sources of acquisition are state and regional archives (including the Russian State Historical Archive (RGIA), State Archive of Russian Federation (GARF), Leningrad Regional State Archive in Vyborg, Archive of Foreign Policy of the Russian Empire, etc.), national libraries and major research libraries of the country. The Presidential Library works in cooperation with regional libraries and research institutions in Russia and abroad, the Directorate for Navigation and Oceanography of the Russian Federation Defence Ministry, the Russian Geographical Society, other non-profit organizations and private collections. At the end of 2019, our collection included about 830 000 items, and more than half of the collection are archival materials.

All the materials in the collection are described in the RUSMARC format, which was developed as a national implementation of UNIMARC considering the specifics of national practice and some differences existing between ISBD and Russian Cataloguing Rules (Skvortsov and Zhlobinskaya, 2018).

Since most of the Presidential Library collection consists of archival materials, one of the most important tasks to be solved from the very beginning of the library was the description of such materials. How to describe them? And how to manage archival records – keep a separate catalogue or include the archives into the integrated library catalogue along with all other types of materials? It is evident that archival and library materials are quite similar by nature in that they document human history. But due to different traditions, standards and description practices, archival materials are often not listed in library catalogues. On the other hand, we supposed that the inclusion of archival catalogue records into the integrated online catalogue could enable users to locate archival resources more easily.

## 2 Designing description for archival resources in RUSMARC

In 2010, the Presidential Library established a working group to develop the principles and approaches to integrated presentation and access to library, archival and museum resources in compliance with modern international standards. The Working Group included representatives from major libraries and archives of the country (Presidential Library, National Library of Russia, Russian State Library, Russian State Historical Archive, State Archive of Russian Federation), National RUSMARC Service, National Library Information Centre "LIBNET", library and archival software vendors (DIT-M, AS Company, Alt-Soft Company). The main objectives of the Working Group were:

- to analyse national and world practice of archival description as well as international standards and machine-readable formats in the field;
- to define data elements for different levels of archival description (fonds – series – file – item);
- to develop specific requirements for archival description in the library;

- and finally, if necessary, to develop appropriate updates to the RUSMARC format for archives.

The Working Group studied major international standards used by archives. Three of them – *General International Standard Archival Description (ISAD (G))* (Zhabko ed., 2011b), *ISAAR (CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families* (Zhabko ed., 2011a), and *Encoded Archival Description (EAD) Tag Library (version 2002)* (Zhabko ed., 2011c) were translated by the Presidential Library and published in Russian for local users. On the national level we did not have (and still do not have) a national standard for archives, just recommendations and general regulations issued by the Federal Archival Agency, and thus, the practice used in the country might vary from one archive to another.

The major standards used in libraries are ISBD (in Russia – Russian Cataloguing Rules and national standard for bibliographic description GOST R 7.0.100-2018) and MARC formats. MARC formats were not supposed to be used for archives and in general they do not allow to create detailed archival description but nevertheless they still can be used for specific purposes.

Most researchers note that unlike a bibliographic record, which usually focuses on an individual manifestation of a published work, an archival description is highly hierarchical and typically it involves a complex group of interrelated unique materials whose shared provenance and hierarchical arrangement must be shown very clearly. A MARC record, which gives a horizontal view of the object is not the best instrument to describe archives. But in fact, UNIMARC users are lucky because UNIMARC and UNIMARC-based formats have the mechanism for building comprehensive multilevel descriptions and establishing links between different levels.

A comparative study of RUSMARC (UNIMARC) and ISAD(G) showed that UNIMARC lacked some data elements which were essential for archives (Zhlobinskaya et al., 2013). Based on the results of the Working Group activities, in 2012 National RUSMARC Service and the Presidential Library submitted proposals to the IFLA Permanent UNIMARC Committee (PUC) to make changes and additions in UNIMARC, which we supposed to be necessary to present archives. After discussion, most of these proposals were approved by PUC, and later similar updates were made in RUSMARC. Also, a draft of the *Guidelines for Archives* (further Guidelines), in UNIMARC was developed, which is in the phase of finalising.

### **3 General principles of archival description in UNIMARC**

The draft of the Guidelines gives the principles of using UNIMARC in presenting archival materials, lists data elements that could be commonly included in the description of archives and indicates where each data element would be placed in a UNIMARC record. Section 4 gives examples of full records for archives. Appendix A contains the Table of correspondence of EAD elements/attributes and UNIMARC, which includes the most significant EAD elements mapped to UNIMARC fields. General principles documented in these Guidelines are type of control, multilevel description and dates.

#### **3.1 Type of control**

First of all, when archival descriptions are incorporated into library catalogues, it is necessary to specify that this particular record describes archival material rather than traditional library

resources. This information would make it possible to implement different options of search, facet navigation and displaying search results to users. Materials might be defined as archival by the manner in which they are described, and here the most important is not the type of materials (text, photo, etc.) but the way in which materials are arranged, physically organised and handled by the custodian. Type of control is specified in the record label (character position 9), and code "a" (archival control) is defined as "a method of describing and handling materials wherein the focus is on the contextual relationships between the items and on their provenance, rather than on bibliographic detail".

### **3.2 Multilevel description**

Archival description is normally presented in a highly hierarchical structure proceeding from general description to more specific descriptions of its components, the description of each part being linked to its higher unit.

In UNIMARC the whole hierarchy may be built using a mechanism for linking entries (group of fields 46X), and archival description can be constructed with any desired level of detail – it may be collection level, if it is sufficient, or it may be set of records for all desired levels. The hierarchical level is indicated in the record label, character position 8. Additionally, the position of each described unit within the hierarchy is identified explicitly – with field 251 *Organization and Arrangement of Materials*. Normally, standard terminology is used here – fonds, series, file, item.

### **3.3 Dates**

In ISAD(G) dates are considered to be one of the essential elements when describing archives; any description of an archival collection or a part of the collection should always include the date(s) when the documents in the collection (or part of the collection) were created. If any date is not known exactly, an estimated date is given, and this is indicated in the description.

Most rules for archival description recognize dates of the collection as a part of the title, and in a UNIMARC record, the dates may be entered into field 200 *Title and Statement of Responsibility*, subfields \$j (*Inclusive dates*) and \$k (*Bulk dates*). When the dates of creation are not considered as a part of the title, they may be written in the field 210 \$d *Date of publication, distribution, etc.*

Dates are also given in coded form in field 100 *General Processing Data*, character positions 8–16.

These principles were taken as a foundation for archival description in the Presidential Library.

## **4 Workflow of describing archive resources**

The workflow of describing archives is presented in the edition *Metodicheskie rekomendatsii po katalogizatsii i indeksirovaniyu elektronnykh kopiy dokumentov* (Guidelines for Cataloguing and Indexing digital copies of documents) published by the Presidential Library in 2014 (Zav'yalova et al., 2014). Description in the Presidential Library proceeds from the general to the specific, which corresponds to the basic principle of the Guidelines and, in general, to the basic theoretical principle of archival description, i.e. the principle of "respect des fonds".

The process begins with the creation of a collection level for fonds. Then records are created for series, file(s), item(s) – depending on the level of details desired. Normally it would be the fonds-series-file chain. But when necessary, an item-level description may be created, or even a component part description. All records are linked with fields 46-. Though typically a UNIMARC record may include relationships from the higher level to the lower ones, in RUSMARC it is generally implemented to define upward relationships – from the lower hierarchical level to the higher level (e.g. article to journal relationship, or, in the case of archives – from file to series, or from series to fonds). General structure of multilevel archival description is shown in Figure 1.

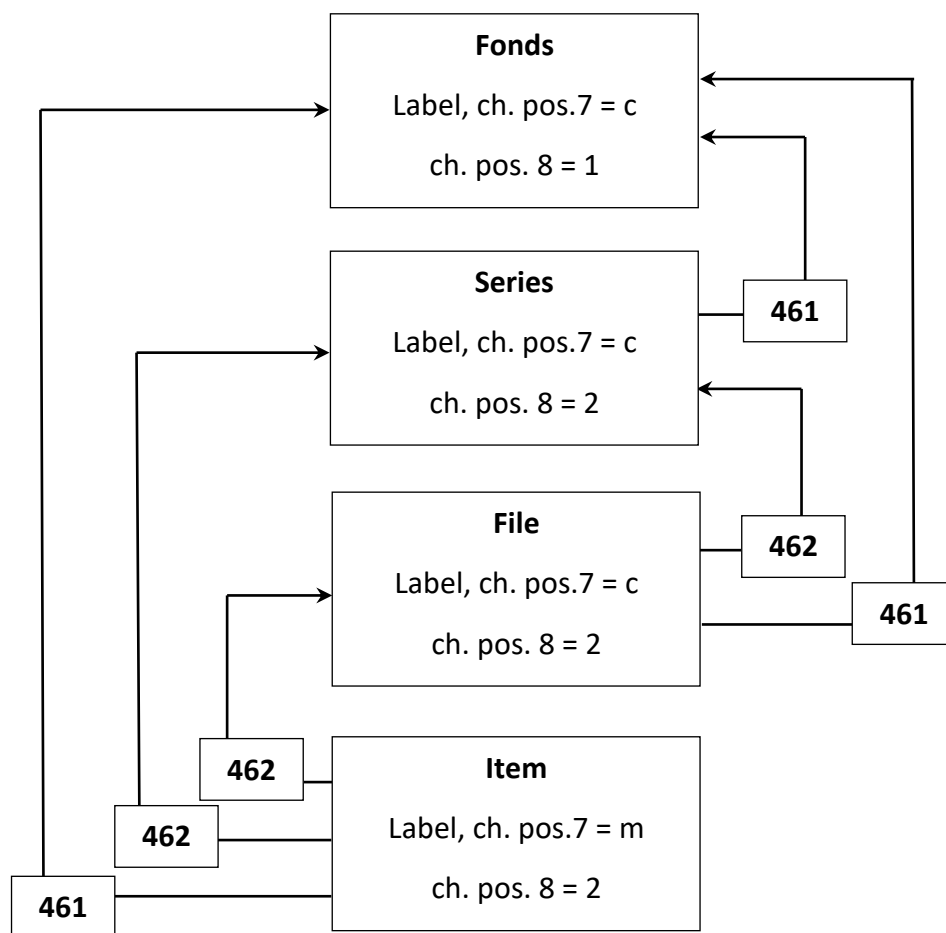


Figure 1: Multilevel archival description in RUSMARC (UNIMARC)

The basic set of data elements specified in the record label and hierarchical links used for different archival levels are given in Table 1.

Table 1: Record label and links in archival description

Data element	Fonds	Series	File	Item
<b>Label, ch. pos. 6 (Type of record)</b>	Basic type of material represented in the record, regardless of the physical format (language materials, notated music, cartographic material, etc. or multimedia)			
<b>Label, ch. pos. 7 (Bibliographic level)</b>	c (collection)	c (collection)	c (collection)	m monograph
<b>Label, ch. pos. 8 (Hierarchical level)</b>	1 (highest level record)	2 (record below highest level)	2 (record below highest level)	2 (record below highest level)
<b>Label, ch. pos. 9 (Type of control)</b>	a (archival control)			
<b>Label, ch. pos. 18 (Descriptive cataloguing form)</b>	x (ISBD provisions are not applicable to the type of resource)			
<b>Links (hierarchical)</b>	No	461 link to FONDS	461 link to FONDS 462 link to SERIES	461 link to FONDS 462 link to SERIES 462 link to FILE

The type of material is specified in Label, ch. pos. 6. For archival fonds normally it would be code "m" (multimedia); at the level of file and item the code value is determined depending of the main content of the unit, regardless of physical media. Thus, RGIA Fond 759 op. 91 d. 100 "Photographs and print portraits of historical persons of the 19th cent. – beginning of the 20th cent." consists of 95 portraits, so code value "k = two-dimensional graphics" is used.

Field 100 contains inclusive dates of the collection. Additionally dates may be given in field 122 when (a) the indication of dates in field 100 is insufficient and it is necessary to specify precise dates of the collection, or (b) inclusive dates in field 200 are given as separate spans of dates and/or dates separated with commas, if there is a significant gap in the chronological sequence of the materials of the collection). See examples:

EX 1:

*Inclusive dates: January, 1 1950 – March, 31 1964*

*122 2#\$ad19500101\$ad19640331*

EX 2:

*Inclusive dates: 2003, 2006, 2010*

*122 1#\$ad2003\$ad2006\$ad2010*

## EX 3:

*Inclusive dates: 1917–1991, 2001*

*122 2#\$ad1917\$ad1991*

*122 0#ad2001*

In general, an archival description may include any material-specific coded data fields depending on the type of material. E.g. for archival units comprising textual materials, field 101 *Language of Resource* is used to specify the language(s) used in the unit. If the archival item is a drawing or a photograph – coded data might be given in field 116 *Coded Data Field: Graphics* to make the description more precise and detailed.

To avoid redundancy of information, information that is common to all lower-level units must be present at the highest appropriate level. On the lower levels, only information appropriate for the level being described must be provided. As a result, elements like material-specific fields for photographs, maps, etc. would be used at item level, and in some cases at file level.

In block 2-- the following fields of descriptive information are used:

200\$a – title of the archival unit

200\$j – inclusive dates

200\$k – bulk dates (when needed)

210 – not used

215\$a – number of items in the unit

251\$c – hierarchical position of the unit in the whole hierarchy of description ("фонд", "опись", "дело", "документ")

In block 3-- any note may be entered as applicable, e.g. note on former owners (field 317), any significant actions with the unit, including arrangement, appraisal, preservation, etc. (field 318), source of acquisition (field 345), conditions governing access and/or reproduction (field 371), textual description of the unit contents (field 330), etc.

Fields in block 4-- provide links to higher level records. Other fields may also be used as appropriate, e.g. field 464 for component parts of the archival item being described.

All access points (personal names, corporate names, family names, geographical names, subject headings) at any level of description are controlled with the authority files that are common for library and archival materials; it makes search results more complete and precise.

The name of the creator (person, corporate body or family) is given in field 7-0 (and 7-1, when there is more than one creator). Names of corporate bodies, persons, families having some other form of responsibility for the unit are given in other fields of block 7--.

Every record at any hierarchical level is indexed with subject headings and class numbers (Library Bibliographic Classification) using the policy of indexing archival materials developed by the Presidential Library.

Example of the record:

Label	06822cmc2a2200000#x#450#
001	6ca00cb3-0a74-4abc-96e7-999e3d90f730
100	##\$a20150128l18401843k##lOrusy0189####ca
101	0#\$ager
102	##\$aDE
120	##\$ab
121	##\$aa   aau#a#
123	1#\$aa
124	##\$aa\$bd
200	1#\$aКарта Персии и Месопотамии, составлена К. Циммерманом\$j1840-1843 г.
215	##\$a1 лист\$скарт.\$d62x48 см
251	##\$cДело
300	##\$aЗаголовок на карте на немецком языке: West-Persien und Mesopotamien. C. Ritters Erdkunde, Buch 3. West-Asien. Iranische Welt, bearbeit von Carl Zimmermann, second Lieutenant im 21 Infanterie Regiment, herausgegeben durch C. Ritter und F. A. O Etzel. Berlin, Verlag v. G. Reimer, 1840-1843
300	##\$aБумага. Тушь, карандаш, акварель
461	#0\$1001a74412a3-3cc5-4d21-8f0a-dfe2fbd9e72d\$12001#\$aДепартамент народного просвещения Министерства народного просвещения\$1251##\$cФонд\$1852##\$j733
462	#0\$10019cc3627e-19b4-421e-a74f-17fa1dd1f54b\$12001#\$aГеографические карты со специальными обозначениями, относящимися к ведомству народного просвещения \$vД. 73\$1251##\$cОпись\$1852##\$j206
607	##\$3RU\NLR\AUTH\66476693\$aМесопотамия (Западная Азия)\$jКарты\$2nlr_sh
607	##\$3RU\NLR\AUTH\6660241\$aИран\$jКарты\$2nlr_sh
617	##\$aИран\$2prlib_sh
617	##\$aИрак\$2prlib_sh
617	##\$aСирия\$2prlib_sh
608	##\$3DITM\AF\0000000386\$aГеографические карты\$2prlib_sh
686	##\$2rubbk\$a26.89(533.5)я64\$vLBC/M
686	##\$a26.17(533.5)\$vLBC/M\$2rubbk
686	##\$a26.89(5Ирн)я64\$vLBC/M\$2rubbk
686	##\$a26.17(5Ирн)\$vLBC/M\$2rubbk
801	#0\$aRU\$bРГИА\$c20150128
801	#1\$aRU\$bГлавНИВЦ\$c20150128\$2rusmarc
852	##\$aРГИА\$j73



Fond level records always include subject heading using the model "<creator name> -- Archives" (field 600, 601 or 602). E.g. record for GARF F.R6991. Совет по делам религий при Совете министров (Council for religious affairs under the Council of ministers of USSR) has the following subject heading: 601 01\$аСССР\$bСовет по делам религий\$jАрхивы (\$аUSSR\$bCouncil for religious affairs\$jArchives).

At series level, the subject heading is always the topic, which is normally present in the title of the series. E.g. Opis' 7 of the above fond GARF R6991 is entitled "Личные дела высших служителей культа" (Personnel files of ministers of religion). The subject heading assigned to it is: 606 1#\$аСлужители культа\$уСССР\$jЛичные дела (\$аClergy\$уUSSR\$jPersonnel files).

At file level, the main subject is defined depending on the content of items that are included in the file. Indexing at item level is type-specific. Here, the subject headings of any category may be used (name, topic, etc.).

For example: The record above describes a file from the fonds of Department of Education (RGIA. F. 733, op. 206, d. 73). The file includes one document – a map of Persia and Mesopotamia.

The record above includes fields 120, 121, 123 and 124 (material-specific data for maps), and it also is indexed with geographical subject headings (fields 607 and 617), form the subject heading (608) and the LBC class numbers.

The approach described above helped enhance the users' search interface. After conducting the search in the catalogue, the user may click a specific tab to choose a specific category of materials – library or archival material (Figure 2).

The screenshot displays the search interface of the Presidential Library. The search query is "ближний восток" (Middle East). The results are categorized into "Library materials" (631 found, 471 002 total) and "Архивные материалы" (187 found, 123 385 total). The interface includes a search bar, a "Найти" button, and a sidebar with sorting options (Sorting order, Level, Inclusive dates, Collection). The main content area shows search results with titles and descriptions, such as "Записка неуставленного лица о политике Англии на Ближнем Востоке и о роли славянской общественности в борьбе с этой политикой, с сопроводительным письмом на имя Александра Илларионовича Васильчикова".

Figure 2: Search interface (Presidential Library, [www.prlib.ru](http://www.prlib.ru))

Both library and archive materials have specific options to define the order of displaying retrieved records or to refine the search results. Archival records may be sorted by relevance, title, creator, inclusive dates, date of publication on the portal. Also, the user may choose to browse records at a specific archival level – fonds, series, file or item, may specify dates they are interested in, or browse materials included in a specific digital collection. Sorting and navigation options are given in the right-hand panel of the screen.

When viewing the selected record, the user may switch to another level.

The screenshot shows the Presidential Library website interface. At the top, there is a logo for the Presidential Library and a navigation bar. The main content area displays the record for a file titled "Три записки на французском языке путешественника кавалера Гамба о русской торговле с Персией и Средней Азией". The record includes a metadata table with the following information:

Идентификатор	290da2a2-73ba-4d6e-96d7-50046bead83f
Язык текста	Французский
Название	Три записки на французском языке путешественника кавалера Гамба о русской торговле с Персией и Средней Азией
Даты	1820
Объем	51 лист
Уровень	Дело
Фонд	<a href="#">Бумаги М. М. Сперанского</a>
Опись	<a href="#">1803-1847 гг.</a>
Организация-создатель записи	РГИА
Шифр хранения	25
Форматы описания	<a href="#">Архивное описание</a> <a href="#">RUSMARC</a>

On the left side of the record, there are two callout boxes labeled "Fonds" and "Series" pointing to the corresponding fields in the metadata table. On the right side, there are four thumbnail images of the original documents.

Figure 3: Display format (Presidential Library, [www.prlib.ru](http://www.prlib.ru))

For example, Figure 3 shows the record for the file "Three notes in French of the traveller cavalier Gamba on Russian trade with Persia and Central Asia". From this view, the user may switch to viewing the record for the series or fonds. In the example, you may see links to the series "1803-1847" and the fonds "Бумаги М. М. Сперанского" (Papers of M.M. Speransky).

Every archival record may be displayed in one of the following three formats: full format (with text representation of RUSMARC tags), archival format (brief record with archive-specific fields) and full RUSMARC record for professional users.

## 5 Conclusion

Using RUSMARC for describing archival materials, like other types of materials, made it possible to create the integrated catalogue of the Presidential Library, which, in turn, made user search most effective. Hierarchical links provided in RUSMARC (UNIMARC) help keep the hierarchical structure of archival description. On the other hand, the extensive set of coded data and note fields in RUSMARC help enrich the archival description to make it more precise. Archival and traditional library materials complement each other, and their co-existence in the library catalogue makes it more efficient and user-friendly.

## References

Skvortsov, V. and Zhlobinskaya, O., 2018. 20 years of RUSMARC. *IFLA Metadata Newsletter*, 4(2), 8–12. Available at:

[https://www.ifla.org/files/assets/cataloguing/scatn/metadata\\_newsletter-20181217.pdf](https://www.ifla.org/files/assets/cataloguing/scatn/metadata_newsletter-20181217.pdf) [09.01.2020].

Zav'yalova, L. et al., 2014. *Metodicheskie rekomendatsii po katalogizatsii i indeksirovaniyu elektronnykh kopiy dokumentov [Guidelines for cataloguing and indexing digital copies of documents]*. Sankt-Peterburg: Prezidentskaya biblioteka. Available at:

<https://www.prlib.ru/item/343460> [09.01.2020].

Zhabko, E., ed., 2011a. *ISAAR (CPF): Mezhdunarodnyi standart po sozdaniyu arkhivnykh avtoritetnykh zapisei dlya organizatsii, lits i semei [International Standard Archival Authority Record for Corporate Bodies, Persons and Families]*. Sankt-Peterburg: Prezidentskaya biblioteka. Available at: <https://www.prlib.ru/item/395178> [09.01.2020].

Zhabko, E., ed., 2011b. *ISAD (G): Osnovnoi mezhdunarodnyi standart arkhivnogo opisaniya [ISAD(G): General International Standard Archival Description]*. Sankt-Peterburg: Prezidentskaya biblioteka. Available at: <https://www.prlib.ru/item/395179> [09.01.2020].

Zhabko, E., ed., 2011c. *Kodirovannoe arkhivnoe opisanie (EAD). Biblioteka tegov. Versiya 2002. [Encoded Archival Description (EAD). Tag library. Version 2002]*. Sankt-Peterburg: Prezidentskaya biblioteka. Available at: <https://www.prlib.ru/item/417799> [09.01.2020].

Zhlobinskaya, O. et al., 2013. *Printsipy i podkhody k sovmeshcheniyu predstavleniya i dostupa k bibliotechnym, arkhivnym i muzeinym resursam: sbornik nauchno-metodicheskikh materialov rabochei gruppy Prezidentskoi biblioteki [Principles and approaches to unification of presentation and access to library, archival and museum resources: a collection of methodological materials of the working group of the Presidential Library]*. Sankt-Peterburg: Prezidentskaya biblioteka. Available at: <https://www.prlib.ru/item/360874> [09.01.2020].

# O pomenu predmetnih oznak in podatka o afiliaciji v raziskovalnih knjižnicah v COBISS.SI: praksa izpolnjevanja polj bloka 6XX in podpolja 70X8

On the importance of subject cataloguing and information about author affiliation in research libraries in COBISS.SI: the practice of indexing fields 6XX and subfield 70X8

**Maja Peteh<sup>1</sup>, Primož Južnič<sup>2</sup>**

**IZVLEČEK:** Prispevek prinaša nov vpogled v katalogizacijsko prakso v sistemu COBISS.SI, posebej z vidika potreb bibliografskih podatkov za vodenje bibliografij raziskovalcev in raziskovalnih organizacij. V prispevku sta evalvirani dve vsebini katalogizacijske prakse, ki sta neobvezni, a priporočljivi. Raziskava se ukvarja z ugotavljanjem doslednosti vnosa podatka v podpolje 70X8 in analizo obstoječih možnosti za uporabo tega podatka pri pripravi bibliografij raziskovalne organizacije. Uporabljeni so trije metodološki pristopi: a) izvedba in analiza spletnega vprašalnika za knjižničarje (n = 113), zaposlene v visokošolskih in specialnih knjižnicah, o njihovem delu, b) analiza izpolnjevanja podpolja 70X8 (afiliacija avtorja) v bibliografskih zapisih, kreiranih za potrebe bibliografij raziskovalcev v sedemdesetih raziskovalnih knjižnicah, ki sodelujejo v sistemu COBISS.SI, in c) analiza beležk podatkov o kreiranih bibliografskih zapisih (n = 401) desetih knjižničarjev bibliografov. Ugotavlja se nedoslednost predmetnega označevanja bibliografskih zapisov in vnosa podatka o zaposlitvi avtorja. Vnos podatka o zaposlitvi avtorja naj bi bil obvezen za zapise, ki jih želi določena ustanova uvrstiti v bibliografijo ustanove, a ga redno vnaša le 1 % knjižničarjev bibliografov, nikoli pa ga ne vnaša 16 % knjižničarjev bibliografov. Predmetne oznake v slovenskem jeziku dosledno vnaša 83 % anketiranih knjižničarjev bibliografov, le 26 % pa jih vnaša tudi v drugem jeziku. Knjižničarji bibliografi povprečno vnesejo 4,5 predmetne oznake v slovenščini in eno v tujem jeziku. Pogosteje vnašajo nekontrolirane predmetne oznake. Zaradi posebnosti slovenskega knjižničnega sistema, ki povezuje vzajemni katalog in informacijski sistem o raziskovalni dejavnosti (SICRIS), so rezultati uporabni v slovenskem kontekstu in okoljih COBISS. Predmetne oznake v tujem, predvsem angleškem jeziku so dodana vrednost zapisu tudi zaradi vse večjih izmenjav podatkov med COBISS.SI ter vzpostavljenim Nacionalnim portalom odprte znanosti in drugimi zbirkami. Vnos predmetnih oznak sicer ni obvezen, vendar pa gre za pomemben podatek. Kljub temu ni zanemarljiv delež knjižničarjev bibliografov, ki se ne zavedajo pomena tega podatka ali pa ga ne vnašajo zaradi preobremenitve.

**KLJUČNE BESEDE:** bibliografije raziskovalcev, bibliografije ustanov, COBISS.SI, SICRIS, znanstvene knjižnice, bibliografski podatki, vsebinska obdelava, knjižničarji bibliografi

---

<sup>1</sup> Korespondenčni avtor: mag. Maja Peteh, Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za gozdarstvo in obnovljive gozdne vire, in Gozdarski inštitut Slovenije, Ljubljana, Slovenija, [maja.peteh@gozdis.si](mailto:maja.peteh@gozdis.si).

<sup>2</sup> Dr. Primož Južnič, Univerza v Ljubljani, Filozofska fakulteta, Oddelek za bibliotekarstvo, informacijsko znanost in knjigarstvo, Ljubljana, Slovenija, [primoz.juznic@ff.uni-lj.si](mailto:primoz.juznic@ff.uni-lj.si).

**ABSTRACT:** The paper gives a new insight into the cataloguing practice in the COBISS.SI system, primarily regarding the necessity of bibliographic data for the management of bibliographies for researchers and research organisations. The paper evaluates two areas in the cataloguing practice, both of them optional but recommended. The research deals with the consistency of the entry of author affiliation and with the analysis of the existing possibilities of using it in the process of preparing bibliographies for research organisations. Three methodological approaches are used: a) performance and analysis of an online questionnaire for academic and special librarians (n =113) about their work as librarians bibliographers, b) analysis of indexing subfield 70X8 (author affiliation) in bibliographic records, created for the needs of researchers' bibliographies, in seventy libraries participating in the COBISS.SI system, and c) analysis of the data logs of 401 records created by ten librarians bibliographers. The results show the inconsistency of subject indexing of bibliographic records and of the entry of the author affiliation. The entry of this information should be mandatory for the records that an institution wishes to be included in its bibliography but only 1% of librarians bibliographers regularly enter this information, whereas 16% of librarians bibliographers never enter it; 83% of the librarians bibliographers who answered the questionnaire consistently enter subject headings in Slovenian, only 26% also in other languages. On average they enter between 4.5 subject headings in Slovenian and one in another language. They tend to index uncontrolled subject headings more often. Due to the special nature of the Slovenian library system, which links the shared catalogue and the current research information system (SICRIS), the results are only applicable in the Slovenian context and in the COBISS environment. Subject headings in a foreign language, especially in English, represent an added value to the record because of the increasing data exchange between COBISS.SI and the established portal Open Science Slovenia and other databases. Even though the entry of subject headings is optional, the data is very important. However, the proportion of librarians bibliographers who are not aware of the importance of this information or do not enter it due to work overload is not negligible.

**KEYWORDS:** bibliographies of researchers, bibliographies of institutions, COBISS.SI, SICRIS, research libraries, bibliographic data, subject cataloguing, librarians bibliographers

---

## 1 Uvod

Posebnost slovenskega knjižničnega sistema je sistem COBISS z vzajemnim katalogom COBIB.SI, ki je povezan tudi z informacijskim sistemom o raziskovalni dejavnosti (Current Research Information System – SICRIS). Ta vključuje podatkovne zbirke o slovenskih raziskovalcih in raziskovalnih organizacijah (v nadaljnjem besedilu: RO) ter njihovi dejavnosti (Južnič, 2015; Peteh in Južnič, 2018). Povezanost teh dveh sistemov nam ponuja vrsto možnosti za izpise bibliografij raziskovalcev in RO ter posledično evalvacijo njihovega prispevka v znanstveni produkciji. Vodenje bibliografij ni vezano na izgradnjo kataloga (fizične) knjižnične zbirke, saj se v sistem vnašajo tudi dela, ki jih knjižnica nima v svoji zbirki; dela so lahko objavljena na spletu ali celo neobjavljena. Knjižničar bibliograf jih običajno vnaša v sistem za zaposlene v matični ustanovi knjižnice, lahko pa tudi kot storitev za zunanje naročnike. Delo več kot tretjine knjižničarjev bibliografov ni vezano na referalno delo z uporabniki in na pripravo letnih poročil, dve tretjini pa tudi ne na pripravo bibliografij (Peteh in Južnič, 2018).

Pri tem delu ima zagotavljanje kakovosti pomembno vlogo, saj podatki pomembno vplivajo na storitve, namenjene raziskovalcem, kot so izpisi osebnih bibliografij, prijave na razpise Agencije za raziskovalno dejavnost Republike Slovenije (ARRS), bibliografski kazalci uspešnosti za izvolitev v nazive v okviru univerz ipd. (Dornik et al., 2017). Zato je pomembno stalno

spremljanje kakovosti in postavljanje kazalnikov kakovosti v katalogizaciji (Dornik, 2019). Za vsebinsko pravilno razvrstitev skrbijo Osrednji specializirani informacijski centri (OSIC).

Knjižničarji so vedno bolj pomembni tudi kot partnerji pri znanstvenoraziskovalnem delu. V slovenskem okolju so nepogrešljivi v procesu evalvacije, saj skrbno beležijo informacije o objavah raziskovalcev svoje organizacije v sistem COBISS.SI. Sistem COBISS.SI je povezan s sistemom SICRIS in omogoča vpogled v bibliografije raziskovalcev ter tudi vrednotenje njihovega dela, in sicer po metodologiji, ki jo določa ARRS. Hkrati so ti podatki uporabni tudi kot vir za različne bibliometrijske raziskave (Južnič, 2015). Objave raziskovalnih rezultatov, pri katerih slovenski raziskovalci uporabljajo te podatke, so pogosto opažene in tudi ustrezno citirane.

V okviru slovenskega Nacionalnega portala odprte znanosti so vzpostavljeni repozitoriji vseh slovenskih univerz (DKUM – Digitalna knjižnica Univerze v Mariboru, RUL – Repozitorij Univerze v Ljubljani, RUP – Repozitorij Univerze na Primorskem, RUNG – Repozitorij Univerze v Novi Gorici), repozitorij raziskovalnih organizacij (DIRROS – Digitalni repozitorij raziskovalnih organizacij Slovenije) ter samostojnih visokošolskih in višješolskih izobraževalnih organizacij (ReVIS – Repozitorij samostojnih visokošolskih in višješolskih izobraževalnih organizacij). Repozitoriji so kompatibilni z navodili portala OpenAIRE in vključeni v različne spletne imenike, agregatorje ter iskalnike (npr. OpenDOAR, ROAR, BASE). Tako izpolnjujejo zahtevo pokrovitelja, Evropske komisije, ki le prek portala repozitorijev, kompatibilnih z navodili portala OpenAIRE, preverja izpolnjevanje določil iz pogodbe o sofinanciranju. Pomembno je izpostaviti, da so vzpostavljene povezave med repozitoriji in sistemom COBISS.SI, kar zelo olajša delo knjižničarjev. Ti so dopolnili seznam opravil v okviru znanstvenoraziskovalnega dela z urednikovanjem vnosov v digitalnih repozitorijih (Peteh in Južnič, 2018). Kvaliteten bibliografski zapis v sistemu COBISS predstavlja tudi osnovo metapodatkov za prenose v druge sisteme (Koler-Povh, 2016), tudi institucionalne e-repozitorije. Vzpostavljena je namreč možnost prenosa bibliografskih podatkov iz sistema COBISS v prej naštete repozitorije. V tem procesu ni vzpostavljena kvaliteta kakovosti, razvijalci sistema predpostavljajo kvaliteto zapisov v COBISS.SI. Če bibliografski zapis ne vsebuje določenih podatkov, se ti ne morejo prenesti.

Knjižničar katalogizator se mora zavedati, da z vnosom zapisov v sistem COBISS.SI:

- sodeluje pri izgradnji vzajemnega kataloga,
- sodeluje pri gradnji nacionalne bibliografije,
- zapisi postanejo predmet izmenjav podatkov med COBISS.SI in vzpostavljenim Nacionalnim portalom odprte znanosti. Bibliografski zapisi so namenjeni tudi mednarodni izmenjavi z WorldCat (OCLC), normativni zapisi se izvažajo v VIAF (Virtual International Authority File), uporabniki pa jih lahko izvažajo tudi v orodja za upravljanje referenc in citiranje. Pravilni podatki v zapisih so zato zelo pomembni in kažejo na pomembnost dela knjižničarjev katalogizatorjev.

## **2 Bibliografija raziskovalne organizacije**

V tem prispevku se ukvarjamo z bibliografijo RO, ki lahko vključuje znanstveno, strokovno, poljudno in umetniško produkcijo zaposlenih (običajno raziskovalcev), pa tudi prispevke, ki govorijo o organizaciji. Škerget in Rifl (2014) pišeta, da je »sestavljanje Slovenske bibliografije naloga NUK, za bibliografije slovenskih raziskovalcev sta odgovorna Institut informacijskih

znanosti v Mariboru – IZUM (z izobraževanjem ter z upravljanjem in vzdrževanjem baze SICRIS) in Javna agencija za raziskovalno dejavnost Republike Slovenije (ARRS).« Sistem COBISS in zbirka SICRIS se v slovenskem raziskovalnem prostoru uporabljata za izpise bibliografij (raziskovalcev). Menimo, da so v tem kontekstu za kvaliteten izpis bibliografije organizacije najbolj odgovorni knjižničarji bibliografi. Ti v postopku katalogizacije izvedejo izbor, vrednotenje, vsebinsko analizo, klasifikacijo. Postopek je opisoval že Logar (1970) in je še vedno temelj dela bibliografa, spremenil se je le medij, s katerim bibliograf ureja zbirko in medij objave. Bibliografija je, kot pravi Mišković (2010), 'končni izdelek' in hkrati 'praksa'.

Informacijski sistem SICRIS se uporablja za potrebe evalvacije raziskovalcev, raziskovalnih skupin in organizacij. Omogoča različne izpise, npr. 'Vrednotenje bibliografskih kazalcev raziskovalne uspešnosti po metodologiji ARRS'. Izpis je mogoče pripraviti a) za nabor raziskovalcev, ki so trenutno vključeni v RO, b) z izbiro 'Vrednotenje skupine izbranih raziskovalcev', kjer so že zajeti raziskovalci, trenutno vključeni v RO, dodati pa je možno tudi druge raziskovalce in c) z opcijo 'Vrednotenje skupine izbranih raziskovalcev', kjer sami definiramo, katere raziskovalce želimo vključiti. Vse omenjene izpise lahko pripravimo za poljubno obdobje, kar pa predstavlja zagato. Algoritem izpisa namreč ne upošteva možnosti, da raziskovalec ni bil del RO celotno obdobje ali celotno koledarsko leto in v izpis vključi objave raziskovalca, ki jih je ta pripravil v sklopu dela v drugi RO. Dodatno pa ne upošteva možnosti, da objave izhajajo z zamikom – tudi večletnim (npr. članki, ki izidejo po zaključku dela raziskovalca v RO). Izpisi tipa 'Vrednotenje bibliografskih kazalcev' so tako namenjeni vpogledu v trenutni raziskovalni potencial RO.

Ta problem so zaznali tudi predstavniki OSIC-ov in na ARRS, kjer ugotavljajo, da »če izpišemo vrednotenje raziskovalne uspešnosti določene organizacije za izbrano obdobje, npr. 2013–2017, v izpis ne zajamemo podatkov o objavah tistih raziskovalcev, ki npr. od leta 2016 niso več zaposleni na fakulteti, hkrati pa so v izpis zajeta dela raziskovalcev, ki so v organizaciji zaposleni na novo. SICRIS (Current Research Information System) je namreč sistem, ki kaže trenutno stanje, ne pa retrospektivnega. Če bi želeli pri organizacijah izpis del glede na kriterij vnesene šifre ustanove (podpolje 70X8), je trenutno edina možnost izpis bibliografije ustanove (BIB261) v COBISS3/Izpisi.« Zaključujejo s sklepom o vključitvi možnosti izpisa bibliografije ustanove tudi glede na podatke v podpolju 70X8 (Zapisnik ..., 2018).

ARRS poseduje podatke, kdaj je bil raziskovalec dodan v RO ali bil iz nje brisan, a žal tudi ti podatki ne držijo – preprosto zato, ker RO in vodje raziskovalnih skupin ne (po)skrbijo za ažurnost podatkov; upošteva se namreč datum, ko je RO ta podatek/spremembo sporočila.

Iz sistema COBISS (prek COBISS+ z iskalno predpono FC ali z vmesnikom COBISS3) lahko pripravimo izpis bibliografije ustanove na podlagi izpolnjenega podpolja 70X8, kamor se vnaša podatek o pripadnosti avtorja instituciji, t. i. šifro ustanove/organizacije. Podpolje je namenjeno vodenju bibliografij ustanov v sistemih COBISS in naj bi bilo obvezno v zapisih, ki jih želi določena ustanova uvrstiti v bibliografijo ustanove (COMARC/B format, 2019) in kjer se uporablja šifrant ustanov/organizacij. Popolnost izpisa bibliografije ustanove je tako odvisna od prepoznane potrebe posameznih knjižničarjev bibliografov.

### **3 Pomen vsebinske obdelave pri vnosu zapisov za vodenje bibliografij**

V sistemu COBISS so predmetne oznake (vnašamo jih v polja 600–610) in klasifikacijske oznake (vnašamo jih v polja 675–686) pomembne z vidika najdenja gradiva v katalogu. Vzajemni



katalog predvideva enoten način obdelave, torej predpostavlja uporabo »istih načel, pravilnikov in standardov« (Likar, 2003). Vnos predmetnih in klasifikacijskih oznak je priporočljiv, izjema je obvezno podpolje 675c (klasifikacijski UDK vrstilec z enotnim šifrantom) (COMARC/B format, 2019). Vsebinska analiza je odvisna od končnih uporabnikov, tipa knjižnice, knjižnične zbirke, katalogizatorjem pa so na voljo posamezni sistemi in geslovniki. Kot je ugotavljal že Klemen (2011), sta katalogizatorjem na razpolago Spletni splošni slovenski geslovník (SSSG) in spletna Univerzalna decimalna klasifikacija (UDCMRF12), ki pa še nista učinkovito povezana s sistemom za vzajemno katalogizacijo.

Za potrebe izvajanja analiz kakovosti podatkov v bibliografskih zapisih so v IZUM-u oblikovali nabor dimenzij za ugotavljanje kakovosti podatkov v bibliografskih in normativnih zapisih. Kakovostni podatki so navedeni v edinstvenem zapisu, semantično točni, točno prepisani, strukturalno in vsebinsko popolni, strukturalno in vsebinsko skladni, oblikovno dosledni, predvideni oz. neodvečni, aktualni, lahko pa so tudi dodatno informativni (Badovinac, 2018).

Celostnih analiz za blok 6 s tega vidika še ni na voljo, v rezultatih spremljanja kakovosti v IZUM-u pa zasledimo, da v približno 10 % bibliografskih zapisov manjka podatek vsaj v enem od polj 600–610 (ti podatki zagotavljajo dodatno informativnost zapisa), medtem ko je zatipkanih podatkov v poljih 600–610 razmeroma malo (Badovinac, 2019).

V novih mednarodnih katalogizacijskih načelih (Galeffi et al., 2019), ki pri nas še niso prešla v prakso, ima predmetna oznaka večjo vrednost in postane bistvena točka dostopa ('točka dostopa za predmetne oznake'). Predmetne oznake bodo tako enako pomembne kot normativna točka dostopa (avtor oz. avtorji, stvarni naslov, datum izida oz. izhajanja, standardne številke, identifikatorji in 'ključni naslovi' opisovane entitete).

Pri svojem delu se mora knjižničar katalogizator zavedati, da je zapis kakovosten, če so vanj vneseni vsi podatki, potrebni za identifikacijo publikacije, in če je v zapis vgrajenih dovolj podatkov, ki uporabniku omogočajo, da zapis najde. V Sloveniji se tako istočasno, vzajemno gradita vzajemni katalog in nacionalna bibliografija, zapisi pa so tudi predmet izmenjav podatkov med COBISS.SI ter vzpostavljenim Nacionalnim portalom odprte znanosti in drugimi zbirkami.

Predmetno označevanje zapisa od knjižničarja bibliografa zahteva poglobitev v vsebino dela v obdelavi, kar v procesu izdelave bibliografskega zapisa zahteva dodaten čas in poglobljeno poznavanje področja. Zaradi drugih delovnih obremenitev (Peteh in Južnič, 2018) smo v tem delu raziskave predpostavljali, da knjižničarji bibliografi prihranijo nekaj delovnega časa na račun vključevanja predmetnih oznak.

## 4 Metodologija

V raziskavi nas je zanimala doslednost vključevanja predmetnih oznak (polja 600–610), klasifikacijskih oznak (polja 675–686) in podatka o afiliaciji (podpolje 70X8) v bibliografske zapise v bazi podatkov COBIB.SI, ki so predmet bibliografije raziskovalcev. Predpostavljali smo, da knjižničarji bibliografi pogosteje vključujejo predmetne oznake v slovenskem jeziku kot v angleškem in da če so v delu navedene ključne besede le v angleškem jeziku, jih večinoma ne prevajajo.

Po izvedbi pilotnega vprašalnika v letu 2016 smo vnos v izpopolnjeni e-vprašalnik omogočili od 18. januarja do 21. februarja 2017. Vprašalnik je bil pripravljen z orodjem 1KA, obvestilo pa



je bilo z e-pošto posredovano prek vseh dostopnih obveščevalnih strokovnih list (strokovna knjižničarska društva, sezname NUK, liste sekcij ZBDS) in prek skupin na Facebooku. V letu 2017 je bilo v slovenskih visokošolskih in specialnih knjižnicah zaposlenih 554 delavcev za knjižničarsko delo (redno zaposleni in delavci, ki niso redno zaposleni) (BibSist, 2018), rezultati prikazujejo prakso 20 % teh delavcev. V tem času je bilo v slovenskih knjižnicah aktivnih 652 knjižničarjev katalogizatorjev (Dovoljenja za vzajemno katalogizacijo, 2018). Točnega podatka, koliko se jih ukvarja z vodenjem bibliografij, ni.

Vprašalnik je vseboval 19 vprašanj, od tega dve demografski vprašanji. Vprašalnik je odprlo 213 respondentov, od tega je bilo 113 vprašalnikov izpolnjenih v celoti ali vsaj delno zadovoljivo. Rezultati prikazujejo izkušnje in prakso 20 % vseh strokovnih delavcev, zaposlenih v visokošolskih in specialnih knjižnicah. Analiza časovnega vidika kreiranja zapisov v sistemu COBISS3 je prikazana v članku Peteh in Južnič (2018). V tem prispevku se osredotočamo le na rezultate naslednjih dveh anketnih vprašanj:

- Kako pogosto v katalog vnašate predmetne oznake v slovenskem jeziku in drugih jezikih (blok 6)?
- Kako pogosto izpolnjujete podpolje 70X8 (šifra ustanove/organizacije)?

V pilotni izvedbi ankete je sodelovalo 50 knjižničarjev. V podporo anketnemu vprašalniku smo 10 knjižničarjev bibliografov pridobili tudi za sodelovanje v drugem delu naše raziskave. Ta je obsegal beleženje podatkov o zapisih, kreiranih v izbranem obdobju. Knjižničarje bibliografe smo prosili, naj si v času od 20. januarja do 20. februarja 2017 izberejo 10 zaporednih dni in v tem času v vnaprej pripravljeno preglednico beležijo podatke o kreiranih zapisih. Sodelujoči so tako imeli možnost, da sami poiščejo obdobje, ko bodo njihove delovne obveznosti zajemale (tudi) katalogizacijo, ko ne bodo preobremenjeni z drugimi delovnimi obveznostmi in ko bodo prisotni na delovnem mestu.

V preglednici beleženja podatkov o kreiranih zapisih so bibliografi beležili 12 vrst podatkov. To so:

- skupen čas kreiranja zapisa,
- efektiven čas kreiranja zapisa,
- motnje pri delu (brez motenj, telefon, uporabnik, sodelavec, sestanek, drugo),
- tip dela,
- maska zapisa,
- podatek, ali gre za e-dostopen vir,
- zahtevnost vnosa (subjektivna ocena: zelo zahteven, srednje zahteven, nezahteven),
- način začetka izdelave zapisa (nov zapis, prevzem iz kataloga WorldCat, prevzem iz baz ISSN ali COBISS.net, uporaba predloge, priredi v nov zapis, prevzem zapisa ali prevzet zapis s popravki),
- jezik in pisava gradiva (slovenski jezik, tuj jezik, ki ga vsaj delno obvlada, tuj jezik, ki ga ne obvlada),
- vsebinska obdelava (skupno število izpolnjenih podpolj v poljih 600–610, od tega število podpolj 610a v slovenščini in drugih jezikih),
- vsebinska oprema vira (ključne besede in UDK-vrstilec so navedeni na viru, ključne besede je podal avtor ali jih bodo določili sodelavci, zadolženi za vsebinsko obdelavo gradiva in klasifikacijo),

- urejanje normativnih imen avtorjev v bazi CONOR.SI (število izpolnjenih polj avtorstva (70X) in od tega število kreiranih ali dopolnjenih normativnih zapisov in število izpolnjenih polj korporativnega avtorstva (71X).

Naštete podatke smo pridobili za 401 bibliografski zapis. Dodatno je vseh 10 sodelujočih posredovalo tudi informacije o vrsti knjižnice (visokošolska ali specialna), v kateri so zaposleni, in letu, ko so pridobili dovoljenje za vzajemno katalogizacijo. V tem članku prikazujemo le rezultate, vezane na izpolnjevanje polj bloka 6 in podpolja 70X8.

V tretjem sklopu smo se odločili sami preveriti doslednost izpolnjevanja podpolja 70X8. To smo izvedli s križnim preverjanjem izpisov iz baze podatkov COBIB.SI in iz sistema SICRIS. V raziskavo smo zajeli le izpise RO, ki imajo lastno knjižnico in so imele v času pridobivanja podatkov (junij 2016) več kot 20 raziskovalcev, kar je razvidno iz sistema SICRIS. Predpostavljamo, da so potrebe po knjižničnih storitvah v teh organizacijah dovolj visoke, da je organizacija prepoznala potrebo po organizaciji knjižnice tudi s tem, da ima zaposlenega svojega knjižničarja bibliografa, ki skrbi za bibliografijo raziskovalcev in celotne ustanove. Zajeli smo podatke 26 specialnih knjižnic oz. samostojnih inštitutov in 44 visokošolskih knjižnic oz. fakultetnih ustanov (skupaj 70). Iz baze podatkov COBIB.SI smo pridobili izpis zapisov z izpolnjenim podpoljem o tipologiji in afiliaciji avtorja, pri čemer je zadostovalo, da je ta podatek zabeležen pri vsaj enem od avtorjev. Predpostavljali smo, da je knjižničar bibliograf podatek o afiliaciji dopolnil pri vseh ali nobenem izmed avtorjev svoje ustanove. Iskanje smo omejili na obdobje 2010–2015. Iskalna zahteva je vsebovala naslednji iskalni niz: 'PY=2010:2015 AND FC=[šifriran podatek, »šifra ustanove/organizacije«, v sistemu COBISS]\* AND (TD=1\* OR TD=2\* OR TD=3\*)'. Iz izpisa smo zabeležili podatek o številu zadetkov.

Za izbranih 70 raziskovalnih organizacij smo nato pripravili primerjalni izpis bibliografije v sistemu SICRIS z metodo 'Vrednotenje bibliografskih kazalcev raziskovalne uspešnosti po metodologiji ARRS', in sicer v obdobju 2010–2015. V tem primeru podatek v podpolju 70X8 ne vpliva na izpis podatkov, saj algoritem zajame zapise, če je raziskovalec v trenutku priprave izpisa vključen v RO. Na osnovi izpisa smo zabeležili podatek o skupnem številu vključenih del.

## 5 Rezultati

### 5.1 Vključevanje predmetnih oznak v bibliografski zapis

Na vprašanje o doslednosti predmetnega označevanja bibliografskih zapisov je odgovorilo 82 od 113 sodelujočih v raziskavi. Kar 84 % sodelujočih knjižničarjev v bibliografske zapise za objavljena dela dosledno vnese predmetne oznake v slovenskem jeziku. Od tega 1 % predmetnega označevanja opravijo sodelavci, saj imajo v knjižnici bibliografski opis ločen od vsebinske obdelave. Med sodelujočimi knjižničarji bibliografi jih 4 % navede predmetne oznake le, če ima vir ključno dokumentacijsko informacijo, kjer so navedene tudi ključne besede, ali pa podatke za vsebinsko obdelavo posreduje avtor, 9 % knjižničarjev bibliografov pa bibliografske zapise le občasno opremi s predmetnimi oznakami (tabela 1).

Predmetne oznake v angleškem jeziku dosledno vnaša 26 % vprašanih knjižničarjev, 24 % jih navede le, če so navedeni na viru ali jih posreduje avtor, 37 % pa občasno. Delež knjižničarjev bibliografov, ki nikoli ne vnašajo predmetnih oznak v angleščini, je 13 % (tabela 1).

Tabela 1: Vnos predmetnih oznak v polja 600–610 v slovenskem jeziku in drugih jezikih (n = 82)

Vnos predmetnih oznak	Slovenski jezik		Drugi jeziki	
	Frekvenca (n)	Delež (%)	Frekvenca (n)	Delež (%)
Dosledno vnašam	61	74	21	26
Določam le objavljenim delom (tip 1* in 2*)	7	9	0	0
Občasno vnašam	9	11	30	37
Če so navedene v delu ali jih posreduje avtor	3	4	20	24
Nikoli	1	1	11	13
Vnaša sodelavec, zadolžen za vsebinsko obdelavo	1	1	0	0
<b>Skupaj</b>	<b>82</b>	<b>100</b>	<b>82</b>	<b>100</b>

Tabela 2 prikazuje rezultate beleženja podatkov o kreiranih zapisih in prikazuje podatek, kdo je določil predmetne oznake. Največkrat je knjižničar bibliograf tisti, ki zapisu določi predmetne oznake (48,4 %). V 26,2 % je predmetne oznake posredoval avtor ob oddaji dela za vnos v osebno bibliografijo, v 24,9 % pa so bile ključne besede navedene na viru.

Tabeli 1 in 2 skladno dokazujeta, da knjižničar bibliograf v visokošolskih in specialnih knjižnicah kreira celoten zapis in vsebinska obdelava večinoma ni ločena od katalogizacije.

Tabela 2: Vir predmetnih oznak v poljih 600–610 (n = 401)

Vrsta gradiva (tip)	Število zapisov (n)	Na viru (%)	Določi avtor (%)	Določi bibliograf (%)
Znanstveni članki (1.01–1.03)	23	47,8	43,5	8,7
Strokovni in poljudni članki (1.04–1.05)	66	7,6	40,9	51,5
Prispevki v kongresnih zbornikih (1.06–1.13)	28	53,6	32,1	14,3
Poglavja v monografijah (1.16–1.17) <sup>1</sup>	8	12,5	0,0	62,5
Drugi prispevki v serijskih publikacijah (1.19–1.25)	46	4,3	8,7	87,0
Poročila, elaborati ipd. (2.12–2.15)	75	4,0	20,0	76,0
Zaključna dela (2.08–2.11)	58	46,6	27,6	25,8
Kongresni zborniki (2.3* in 2.25)	15	13,3	13,3	73,4
Druge monografije (2.01–2.07, 2.16–2.24)	19	21,1	15,8	63,1
Neobjavljena dela (3*)	63	47,6	30,2	22,2
<b>Povprečje (%)</b>	/	<b>24,9</b>	<b>26,2</b>	<b>48,4</b>

<sup>1</sup> Dodatno je v 25,0 % zapisov predmetne oznake določil sodelavec, ki je zadolžen za vsebinsko obdelavo in ni bibliograf.

Pregled tabel s podatki o kreiranem 401 bibliografskem zapisu desetih knjižničarjev bibliografov pokaže, da je 396 (98,9 %) zapisov opremljenih s predmetnimi oznakami. V polja 600–610 so povprečno vnesli 6,7 predmetne oznake, od tega večino (78,4 %) v polje 610, ki je namenjeno prostemu oblikovanju predmetnih oznak. V polje 610 vnesejo povprečno 4,7 predmetne oznake, v drugih jezikih pa 1,4 predmetne oznake. Priporočila avtorjem člankov v znanstvenih revijah in navodila študentom pri oblikovanju zaključnih del običajno navajajo 3–5 ključnih besed. V primeru angleških revij je ta zahteva omejena na ključne besede (*keywords*) v angleščini, v primeru del v drugih jezikih (npr. slovenščini) pa navajajmo ključne besede v jeziku publikacije in enako število ključnih besed v angleščini. Podroben pregled po tipologijah

zapisov kaže močno izstopanje zapisov za znanstvene članke. Tu je povprečje predmetnih oznak v drugih jezikih 4,5, kar pripisujemo visokemu deležu opremljenosti člankov s ključnimi besedami. To povprečje je v skladu s priporočili avtorjem. Tudi sicer iz povprečja o številu izpolnjenih podpolj v poljih 600–610 izstopajo gradiva, kjer je vsebina pogosteje že opremljena s ključnimi besedami (tabela 3).

Če iz analize izvzamemo dela s tipologijo 3\* (neobjavljena in izvedena dela), ugotovimo, da je med 401 zapisom 252 (74,6 %) zapisov za dela v slovenščini. Pri zapisih za dela v slovenskem jeziku (n = 252) jih je 249 (98,8 %) opremljenih z vsaj eno predmetno oznako v slovenskem jeziku in 34 (13,5 %) tudi z vsaj eno predmetno oznako v drugem jeziku. Pri zapisih (n = 86) za dela, napisana v drugih jezikih, jih je 71 (82,6 %) opremljenih z vsaj eno predmetno oznako v slovenskem jeziku in 40 (46,5 %) z vsaj eno predmetno oznako v drugem jeziku.

Tabela 3: Predmetne oznake v poljih 600–610 v slovenščini in v drugih jezikih (n = 82)

Vrsta gradiva (tip)	Povprečno število vseh podpolj 600–610	Delež podpolj 610a (%)	Povprečno število podpolj 610a v slovenščini	Povprečno število podpolj 610a v drugih jezikih
Znanstveni članki (1.01–1.03)	9,6	81,0	4,6	4,5
Strokovni in poljudni članki (1.04–1.05)	5,7	93,1	5,2	0,2
Prispevki v kongresnih zbornikih (1.06–1.13)	8,4	70,2	4,3	1,8
Poglavja v monografijah (1.16–1.17)	8,0	70,3	5,6	1,5
Drugi prispevki v serijskih publikacijah (1.19–1.25)	4,9	75,4	3,6	0,4
Poročila, elaborati ipd. (2.12–2.15)	5,6	76,2	3,6	0,7
Zaključna dela (2.08–2.11)	6,1	98,9	5,8	0,2
Kongresni zborniki (2.3* in 2.25)	8,9	82,1	5,3	3,0
Druge monografije (2.01–2.07, 2.16–2.24)	6,7	65,6	4,8	0,8
Neobjavljena dela (3*)	7,2	58,5	4,0	0,5
<b>Povprečje</b>	<b>6,6</b>	<b>78,3</b>	<b>4,5</b>	<b>0,8</b>

Na področju klasifikacije UDK je slika precej drugačna. Če so ključne besede povprečno v 24,9 % navedene na viru in jih v povprečju 26,2 % določi avtor ob oddaji dela za vnos v osebno bibliografijo (tabela 2), je UDK v virih (ki ga primarno vnašamo v podpolje 675a) naveden le v 4,2 %. Večino predstavljajo zaključna dela, nekaj je tudi (domačih) znanstvenih člankov. Tako v kar 95,8 % ta naloga pripada knjižničarju bibliografu ali drugemu knjižničarju, zadolženemu za določanje klasifikacijskih oznak.

## 5.2 Vključevanje podatka o afiliaciji v bibliografski zapis

Glede na odgovore knjižničarjev bibliografov (n = 113) v anketnem vprašalniku podatek o afiliaciji (podpolje 70X8), ki je pogoj za izdelavo bibliografije ustanove, redno vnaša le 1 % sodelujočih pri anketiranju, nikoli 16 %, 55 % vprašanih samo za raziskovalce matične organizacije, 28 % sodelujočih pa, če je podatek naveden na publikaciji ali če raziskovalce pozna.

Na primeru sedemdesetih knjižnic je prikazano stanje o:

- številu izpolnjenih podpolj 70X8 pri vsaj enem izmed avtorjev dela, in sicer za dela, nastala v obdobju 2010–2015, kot je zavedeno v COBIB.SI, in
- številu del, vključenih v bibliografijo RO glede na vključene raziskovalce (tabela 4).

Tabela 4: Primerjava izpisa bibliografij RO iz COBISS3 in SICRIS (obdobje 2010–2015)

Stanje v izpisu	Visokošolska knjižnica	Specialna knjižnica	Skupaj
Število zajetih knjižnic	44	26	<b>70</b>
Število zapisov z izpolnjenim podpoljem 70X8 (COBIB.SI)	214.659	84.167	<b>298.826</b>
Število zapisov v izpisih 'Vrednotenje bibliografskih kazalcev raziskovalne uspešnosti po metodologiji ARRS'	268.286	101.825	<b>370.111</b>

Izpis iz vmesnika COBISS3 na podlagi izpolnjenega podpolja 70X8 prikaže okoli 20 % manj enot kot izpis iz sistema SICRIS (bibliografija RO).

## 6 Razprava in zaključki

V prispevku spremljamo doslednost pri vnosu predmetnih oznak predvsem zaradi potrebe po najdbi gradiva v katalogu COBIB.SI tako za domače kot za tuje govoreče uporabnike. Menimo, da so predmetne oznake v tujem (predvsem angleškem) jeziku dodana vrednost zapisu tudi zaradi vse večjih izmenjav podatkov med COBISS.SI ter vzpostavljenim Nacionalnim portalom odprte znanosti in drugimi zbirkami.

Rezultati raziskave so pokazali, da večina (84 %) knjižničarjev bibliografov, ki so odgovorili na naš vprašalnik, v zapise, namenjene za vodenje bibliografij, vnaša predmetne oznake v slovenskem jeziku in tako prispeva k boljši najdljivosti in kakovosti zapisov v slovenskem vzajemnem katalogu. Okoli 15 % vprašanih običajno ne izvede vsebinske analize vira, razen če so ključne besede navedene v delu ali jih je posredoval avtor. Zapisi iz COBIB.SI se prenašajo v tuje kataloge in e-repozitorije, kar slovenski znanosti ponuja preprosto priložnost za promocijo. S tega vidika nas lahko skrbi tudi dejstvo, da 13 % vprašanih knjižničarjev bibliografov v zapise nikoli ne vključuje predmetnih oznak v angleškem jeziku. Čeprav vnos predmetnih oznak ni obvezen, le priporočljiv, lahko zaključimo, da se ta delež knjižničarjev bibliografov ne zaveda pomena tega podatka. Veljalo bi jih opozoriti, da so ti zapisi predmet mnogih izmenjav in je zato vse zapise, tudi zapise za gradivo v slovenščini, treba opremiti tudi s predmetnimi oznakami v angleščini. Zavedamo se, da so knjižničarji bibliografi mnogokrat podvrženi časovnemu pritisku, saj se od njih zahteva takojšen vnos dela. Poleg tega so obremenjeni še z drugim delom in se, tudi zavestno, odpovedujejo podrobni vsebinski obdelavi.

Raziskovali smo možnosti izpisa bibliografij ustanov v sistemu COBISS. Iz kataloga COBIB.SI je to možno, če je dosledno izpolnjeno podpolje 70X8 (afiliacija avtorja), kar je obvezen podatek za zapise, ki jih določena ustanova želi uvrstiti v bibliografijo ustanove. Ugotavljamo, da tega podatka 16 % anketirancev nikoli ne vnaša, 28 % pa le, če je podatek naveden na viru. Rezultati naše raziskave kažejo na nedoslednost pri vnašanju tega podatka, zato se sprašujemo, ali se slovenske RO zavedajo pomena promocije lastnih dosežkov prek letnih in priložnostnih

bibliografij. Kvaliteta izpisa je tesno povezana s kvalitetnim delom knjižničarja bibliografa. Da bi zagotovil popolnost bibliografije organizacije, bi knjižničar bibliograf moral tudi redno spremljati zapise za dela raziskovalcev svoje ustanove, ki so jih kreirali drugi katalogizatorji, in jih dopolniti s podatkom o afiliaciji.

Prek sistema SICRIS ni možno pripraviti izpisa bibliografije RO (ki bi vključil dela raziskovalcev, povezanih z organizacijo v izbranem obdobju, v času priprave izpisa pa več ne, in hkrati le tista dela izbranih raziskovalcev, ki so jih ti objavili v sklopu dela v tej organizaciji, ne pa v okviru vzporednega dela druge). Možnost, ki jo običajno uporabljamo, je vrednotenje raziskovalne uspešnosti določene organizacije ali skupine, a to ni prava retrospektivna bibliografija organizacije. Algoritem v izpis vključi le dela raziskovalcev, ki so trenutno člani raziskovalne skupine ali ustanove, in sicer ne glede na to, ali so dela nastala v sklopu njihovega dela v tej ali kakšni drugi organizaciji. Poleg tega ne vključi z zamikom objavljenih del raziskovalcev, ki so v organizaciji delovali v preteklosti ali le del obdobja.

## Zahvala

Avtorja se želiva zahvaliti vsem knjižničarjem, ki so izpolnili anketo. Pri oblikovanju anketnega vprašalnika so sodelovali člani izvršnega odbora Sekcije za visokošolske knjižnice pri ZBDS. Poudariti želiva pomoč Nataše Knap (Narodna in univerzitetna knjižnica) in Leje Borovnjak (Teološka fakulteta Univerze v Ljubljani) v fazi priprave in promocije ankete ter Kristine Pritekelj (Filozofska fakulteta Univerze v Ljubljani), ki je opravila lektorski pregled. Pri pripravi vprašalnika je koristne napotke posredovala tudi dr. Ema Dornik (IZUM). Za koristne napotke se zahvaljujema tudi recenzentom.

## Reference

Badovinac, B., 2018. Nabor dimenzij za opredelitev kakovosti podatkov v bibliografskih in normativnih zapisih. *Organizacija znanja*, 23(1/2), 2–10.

Badovinac, B., 2019. »Pikice in vejice« pod drobnogledom: spremljanje kakovosti zapisov v letu 2018. *Blog COBISS*, 27. 5. 2019. Dostopno na: <https://blog.cobiss.si/2019/05/27/pikice-in-vejice-pod-drobnogledom/> [29. 1. 2020].

*BibSiST online: statistični podatki o knjižnicah*, 2018. Ljubljana: Narodna in univerzitetna knjižnica, Center za razvoj knjižnic. Dostopno na: <http://bibsist.nuk.uni-lj.si/statistika/index.php> [8. 10. 2018].

*COMARC/B format: za bibliografske podatke*, 2019. Maribor, IZUM. Dostopno na: <https://izobrazevanje.izum.si/EntryFormDesktopDefault.aspx?tabid=38&type=manual&manual=1> *Comarc B svn* [20. 1. 2020].

Dornik, E., Badovinac, B., Kos, J. in Farkaš, B., 2017. Sistem zagotavljanja kakovosti COBIB.SI: izbrane aktivnosti za leto 2016. *Knjižnica*, 61(1–2), str. 191–205.

Dornik, E., 2019. Zagotavljanje kakovosti COBIB.SI: kako uspešni smo? V: *Knjižnice – obvladovalke podatkov? Kongres Zveze bibliotekarskih društev Slovenije, Maribor, 25.–27. september 2019*. Ljubljana: Zveza bibliotekarskih društev Slovenije. Str. 35–50.

*Dovoljenja za vzajemno katalogizacijo*, 2018. Maribor: IZUM. Dostopno na: <http://izobrazevanje.izum.si/EntryFormDesktopDefault.aspx?tabid=360&type=licence> [8. 10. 2018].

Galeffi, A., Bertolini, M.V., Bothmann, R.L., Escolano Rodríguez, E. in McGarry, D., 2019. *Izjava o mednarodnih katalogizacijskih načelih (ICP)*. Ljubljana: Narodna in univerzitetna knjižnica. Dostopno na: [https://www.nuk.uni-lj.si/sites/default/files/dokumenti/2019/ICP\\_2017.pdf](https://www.nuk.uni-lj.si/sites/default/files/dokumenti/2019/ICP_2017.pdf) [20. 1. 2020].

Južnič, P., 2015. Pomen urejenih bibliografij za spremljanje in vrednotenje znanstvenoraziskovalnega dela. *Organizacija znanja*, 20(1), str. 10–14.

Klemen, A., 2011. Ustreznost vsebinskih opisov dokumentov v online dostopnem knjižničnem katalogu. *Knjižnica*, 56(1/2), str. 55–75.

Koler-Povh, T., 2016. Vloga knjižničarja pri izgradnji institucionalnega repozitorija: izziv ali nuja? V: *Upravljanje znanja v knjižnicah: posvetovanje sekcij ZBDS, Maribor, 2016*. Dostopno na: [http://www.zbds-zveza.si/sites/default/files/dokumenti/2013/izvlecek\\_koler2016.pdf](http://www.zbds-zveza.si/sites/default/files/dokumenti/2013/izvlecek_koler2016.pdf) [8. 10. 2018].

Likar, T., 2003. Enotna obdelava knjižničnega gradiva – pogoj za izmenjavo in skupno uporabo bibliografskih zapisov. *Knjižnica*, 47(1/2), str. 7–34.

Logar, J., 1970. *Uvod v bibliografijo*. Ljubljana: Narodna in univerzitetna knjižnica.

Mišković, V., 2010. *Na začetku je (bila) knjiga: bibliografija na Slovenskem*. Ljubljana: Narodna in univerzitetna knjižnica.

Peteh, M. in Južnič, P., 2018. Vrednotenje bibliografske storitve v luči raziskovalne dejavnosti: pomen in vpliv vodenja bibliografije raziskovalcev za knjižničarje v slovenskih raziskovalnih knjižnicah. *Knjižnica*, 63(1), str. 117–137.

Škerget, D. in Rifl, B., 2014. Spletni viri in slovenska bibliografija: študija primera. *Knjižnica*, 58(1/2), str. 35–56.

*Zapisnik sestanka s predstavniki OSIC-ev in ARRS-ja [z dne 15. 11. 2018]*, 2018. Maribor, IZUM. Dostopno na: [https://www.cobiss.si/zapisniki/delovne-skupine/Zapisnik\\_Sestanek\\_z\\_OSIC-i\\_in%20ARRS\\_2018-11-15.pdf](https://www.cobiss.si/zapisniki/delovne-skupine/Zapisnik_Sestanek_z_OSIC-i_in%20ARRS_2018-11-15.pdf) [20. 2. 2018].



# Improving WorldCat quality

## Izboljševanje kakovosti kataloga WorldCat

Jay Weitz<sup>1</sup>

**ABSTRACT:** OCLC's WorldCat approaches 475 million bibliographic records. Many of those records have been created manually by members of OCLC's worldwide cooperative. Others have been added to WorldCat en masse from institutions large and small, from national libraries, from cultural heritage institutions, or from rural public libraries. The focus of this article is quality control of the bibliographic database, historically and currently, in four interrelated aspects: keeping OCLC-MARC validation in harmony with an ever-changing MARC 21, the specific effort to phase out OCLC-defined Encoding Levels in favour of those defined in MARC 21, a history of the automated Duplicate Detection and Resolution (DDR) software, and our work on updating Bibliographic Formats and Standards (BFAS) as reflected in Chapter 4 »When to Input a New Record«.

**KEYWORDS:** WorldCat, OCLC-MARC, MARC 21, quality assurance, data quality control

**IZVLEČEK:** OCLC-jev katalog WorldCat se približuje 475 milijonom bibliografskih zapisov. Veliko teh zapisov so ročno kreirali člani OCLC-jeve kooperative po vsem svetu, mnoge druge pa so v katalog WorldCat dodale velike in manjše ustanove, nacionalne knjižnice, ustanove s področja kulturne dediščine ali podeželske splošne knjižnice. Članek se osredotoča na kontrolo kakovosti bibliografske baze podatkov, in sicer s historičnega in sedanjega zornega kota, s štirih povezanih vidikov. Ti vidiki so: ohranjanje validacije OCLC-MARC v skladu z nenehno spreminjajočim se formatom MARC 21, konkretno prizadevanje po postopni zamenjavi nivojev kodiranja, ki jih je določil OCLC, s tistimi, ki so opredeljeni v MARC 21, zgodovina programske opreme za samodejno odkrivanje in reševanje dvojnikov in naše delo o ažuriranju bibliografskih formatov in standardov (Bibliographic Formats and Standards (BFAS)), kakor se kaže v četrtem poglavju tega dela (»When to Input a New Record«), ki govori o tem, kdaj vnesti nov zapis.

**KLJUČNE BESEDE:** WorldCat, OCLC-MARC, MARC 21, zagotavljanje kakovosti, spremljanje kakovosti podatkov

---

## 1 Introduction

At the time of this writing, OCLC's WorldCat (<https://www.worldcat.org/>) approaches 475 million bibliographic records. Many of those records have been created manually by members of OCLC's worldwide cooperative. Others have been added to WorldCat en masse from institutions large and small, from national libraries, from cultural heritage institutions, or from rural public libraries. Since WorldCat was launched in 1971, participants in the cooperative have built it record-by-record into the most comprehensive collection of data about library collections.

---

The article is based on the following presentations: *How the OCLC MARC update process works* (OCLC, November 2018) and *When to input a new record* (OCLC, January 2019) (both available at [https://help.oclc.org/WorldCat/Metadata\\_Quality/AskQC/Previous\\_AskQC\\_office\\_hours](https://help.oclc.org/WorldCat/Metadata_Quality/AskQC/Previous_AskQC_office_hours)), *Improving WorldCat quality: resolving to reduce duplicates: a history of duplicate detection and resolution* (OCLC, January 2020) (available at <https://www.youtube.com/watch?v=KEJqMqGzHzs>) and *Focus group excitation: OCLC Encoding Levels* (November 2019, with Cynthia Whitacre).

<sup>1</sup> Jay Weitz, Senior Consulting Database Specialist, OCLC Metadata Policy, Ohio, United States, [weitzj@oclc.org](mailto:weitzj@oclc.org).



Behind the scenes, a small group of us at OCLC have been trying our best to continuously improve the quality of data in WorldCat through manual and automated means. Considering both the size of WorldCat and its inexorable growth, our work is daunting, but we try not to be discouraged. We are WorldCat Metadata Quality, about a dozen of us who concentrate on the bibliographic and authority databases, one who works on the WorldCat Registry, and one who works on the WorldCat Knowledge Base. The focus of this article is quality control of the bibliographic database, historically and currently, in four interrelated aspects: keeping OCLC-MARC validation in harmony with an ever-changing MARC 21, the specific effort to phase out OCLC-defined Encoding Levels in favor of those defined in MARC 21, a history of the automated Duplicate Detection and Resolution (DDR) software, and our work on updating Bibliographic Formats and Standards (BFAS) (2020) as reflected in Chapter 4 »When to Input a New Record«.

## 2 MARC 21 and OCLC-MARC evolving

What is now known as MARC 21 did not spring fully formed from the brow of the programmer and analyst Henriette Avram (1919–2006) at the Library of Congress in 1968. It evolved slowly and painstakingly, code-by-code, field-by-field, subfield-by-subfield over the past fifty plus years. What we now call WorldCat became available in 1971, and by that time MARC was already changing.

Between 1973 and 2013, most of the changes made to MARC went through a group called the MARC Advisory Committee (MAC), which included the Committee on Representation in Machine-Readable Form of Bibliographic Information, mercifully shortened to and known familiarly as MARBI. MARBI was an interdivisional committee of the American Library Association (ALA) with representation from the divisions then known as the Association for Library Collections and Technical Services (ALCTS), the Library and Information Technology Association (LITA), and the Reference and User Services Association (RUSA). In 2013, MAC was revamped, no longer sponsored by any ALA division (MARC development, 2019). MAC continued to advise the MARC Steering Committee, consisting of representatives from the Library of Congress (LC), Library and Archives Canada (LAC), British Library (BL), and Deutsche Nationalbibliothek (DNB), and to serve as a discussion forum on the MARC formats and MARC data. Just like MARBI before it, the current MAC has met at each ALA Annual Conference and Midwinter Meeting, although the status of ALA Midwinter is now in flux. Unlike MARBI, MAC members, from various national libraries, library organizations, and specialist communities, all have a vote in making changes to MARC. The *MAC Terms of Reference* make clear that any user of MARC 21 may submit discussion papers or proposals, regardless of one's affiliation with any of the constituent entities (MARC Advisory Committee, 2016).

One convenient way to explain the OCLC-MARC Update process is to follow a particular MARC element through the whole sequence of events, from idea to reality. The creation of a new MARC element, be it a new field, new indicator, new subfield, or what have you, begins with an idea. The element we'll follow actually began at OCLC, as a result of the work on *Faceted Application of Subject Terminology (FAST)* (2020) within OCLC Research.

Without going into a lot of detail here, FAST heightened awareness of a longstanding ambiguity in MARC 21, the need to differentiate subject access points for named events that *cannot* be regarded as responsible agents (such as earthquakes or wars) from named events

that *can* be regarded as responsible agents (such as conferences or meetings). Those within OCLC who were most familiar with the issues got together to draw up a discussion paper for the MARC Advisory Committee (MAC). Historical links to MARC Discussion Papers from 1995 to the present are available on the MARC Standards website at <http://www.loc.gov/marc/mac/list-dp.html>. In the specific *MARC Discussion Paper No. 2016-DP09* (2016) being considered, »Coding Named Events in the MARC 21 Authority and Bibliographic Formats«, the context of the problem was laid out with two possible options for solution. In December 2015, OCLC submitted the discussion paper for consideration at the MAC meetings during the 2016 ALA Midwinter Meeting. As is documented in the »Status/Comments« section of the paper, a straw poll of MAC members revealed a clear preference for one of the options, as well as other suggestions for improving the eventual proposal.

Taking the results and recommendations emerging from the MARC Advisory Committee, the OCLC stakeholders revised the discussion paper into a *MARC Proposal No. 2016-05* (2016), »Defining New X47 Fields for Named Events in the MARC 21 Authority and Bibliographic Formats«. It addressed the concerns raised in the MAC discussion and provided much more detail about the set of proposed set of fields. Again, as can be seen in the »Status/Comments« section, the proposal was approved with the proviso that OCLC would »generate and distribute a list of LCSH headings which are modelled as events in FAST.«

Some months following the MAC meetings at ALA, the Library of Congress announced a new *MARC 21 Update* with the official versions of new and changed MARC elements incorporated into the respective current »base edition« of MARC 21: Bibliographic (February 1999), Authority (October 1999), Holdings (January 2000), Classification (January 2000), and Community Information (January 2000). Historical links to all of the MARC 21 Format Updates from 2000 to the present are available on the MARC Standards website at <http://www.loc.gov/marc/status.html>.

In the past, usually during the third quarter of the calendar year, OCLC would issue a *Technical Bulletin* that announced the OCLC-MARC Bibliographic, Authority, and Holdings format and MARC Code changes to be implemented at that time. Most of the changes were from the two most recent *MARC 21 Updates* and all MARC Codes announced by LC in *Technical Notices* issued since the most recent OCLC-MARC Update. Additionally, we often included other changes requested by members of the OCLC cooperative and suggested by OCLC staff.

Ordinarily, within a few weeks of the release of the OCLC *Technical Bulletin*, we would install the OCLC-MARC Update and announce the implementation via messages on an array of discussion lists, the OCLC Connexion Message of the Day, and elsewhere. As soon as an OCLC-MARC Update was implemented, we would begin the process of making changes to the OCLC documents *Bibliographic Formats and Standards* (2020) and *OCLC-MARC Local Holdings Format and Standards* (2018). In general, changes to WorldCat indexing occur on a schedule independent of the rest of the OCLC-MARC Updates. As a result, changes to the *Searching WorldCat Indexes* (2019) document are not made until later, once the appropriate changes have been made to the indexes.

The OCLC document *Authorities: Format and Indexes* (2018) gives the full information about the valid authority fields and the indexes for each of the files. OCLC now maintains what amounts to two sets of validation rules for its authority files.

More familiar is the long-established set of rules that govern OCLC's version of the traditional Library of Congress-Name Authority Cooperative (LC-NACO) Authority File. These validation rules, covering LC names and LC subjects, conform to *LC Guidelines Supplement to the MARC 21 Format for Authority Data* (2018), popularly known as *The Blue Pages*. Validation changes to the LC-NACO Authority File have to be coordinated among LC, OCLC, and each of the other NACO nodes. At the time of this writing, LC and the NACO nodes including OCLC are in the process of rolling out authority validation changes that have been delayed now for several years.

The other, newer, and less-familiar set of OCLC Authority validation rules covers all of the non-LC authority files that are made available only through OCLC's Record Manager (Record, 2020):

- **Canadiana (Autorités de noms Canadiana en français)**

Source: Bibliothèque et Archives Canada.

The Canadiana Name Authorities in French is used by Library and Archives Canada (LAC) and other Canadian libraries when creating bibliographic descriptions in French.

- **GND Germany Authority File**

Source: Deutsche Nationalbibliothek (German National Library).

GND is an Integrated Authority File that contains over 9 million records for Persons, Corporate bodies, Conferences and Events, Geographic Information, Topics and Works.

- **Māori Subject Headings File**

Source: Ngā Upoko Tukutuku.

Māori Subject Headings provide subject access in te reo Māori to materials for and/or about Māori.

- **MeSH (Medical Subject Headings)**

Source: U.S. National Library of Medicine.

Subject authority file: 630,000 records.

- **NTA Names (Nederlandse Thesaurus van Auteursnamen)**

Source: Koninklijke Bibliotheek (National Library of the Netherlands).

Name authority file: 2,571,933 records representing only personal names.

The OCLC document *Authorities: Format and Indexes* (2018) also gives the full information about the valid authority fields and the indexes for each of these files. Changes from OCLC-MARC Updates may take some time to filter out to this document. As with bibliographic indexing, authority indexing occurs on a schedule independent of the rest of the OCLC-MARC

Updates. As a result, changes to the *Authorities: Format and Indexes* document are not made until later, once the appropriate changes have been made to the authority indexes.

New MARC Codes are announced by the Library of Congress in irregularly scheduled *Technical Notices*, on the average of about 11 or 12 per year. Each LC *Technical Notice* includes the proviso: »The codes should not be used in exchange records until 60 days after the date of this notice to provide implementers time to include newly-defined codes in any validation tables.« In recent years, OCLC has tried to validate new MARC Codes at the next opportunity for installation of validation changes, usually at least once each quarter. That may vary. Sometimes validation occurs more quickly than the sixty-day moratorium but may occasionally take longer.

Beginning in 2018, OCLC has also issued *WorldCat Validation Release Notes* at the time of each installation. The release notes may be found on the OCLC website at:

[https://help.oclc.org/Librarian\\_Toolbox/Release\\_notes](https://help.oclc.org/Librarian_Toolbox/Release_notes). At the same time, OCLC's schedules for changes to validation and for the publication of validation release notes have grown more flexible and frequent. The result is timelier implementation of new MARC elements and codes, as well as quicker notification that they may be used in WorldCat. Older OCLC *Technical Bulletins* remain available on the OCLC website at: [https://help.oclc.org/WorldCat/Cataloging\\_documentation/Technical\\_Bulletins](https://help.oclc.org/WorldCat/Cataloging_documentation/Technical_Bulletins).

### 3 OCLC Encoding Levels

Several of us in Metadata Quality, including Senior Metadata Operations Manager Cynthia Whitacre and my fellow Senior Consulting Database Specialist Robert Bremer, have been working on a special project to phase out the OCLC-defined alphabetic Encoding Level values in favor of the numeric Encoding Level values defined in MARC 21 itself. This is a central aspect of our long-term effort to bring OCLC-MARC into closer harmony with MARC 21, which will ease any future transition to a post-MARC environment. We realized, however, that the OCLC Encoding Levels represent a long tradition built into the practices and workflows of some members of the OCLC cooperative. With that in mind, we wanted to get a firm grasp on how libraries have used Encoding Levels within their cataloging and process workflows with the goal of implementing this change as smoothly and with as little disruption as we could manage.

The *MARC 21 Format for Bibliographic Data* (2019) defines Encoding Level (MARC Leader/17) as a »One-character alphanumeric code that indicates the fullness of the bibliographic information and/or content designation of the MARC record«. In the earliest days of MARC, only the Library of Congress (LC) was authorized to create bibliographic records using any of the MARC-defined Encoding Levels, including blank and the numeric codes.

As a result, when OCLC members first began creating bibliographic records in WorldCat in 1971, OCLC was obliged to implement its own set of alphabetic Encoding Levels (see ELv1 at Bibliographic, 2020) (Table 1).

Table 1: MARC 21 Encoding Levels and OCLC-MARC Encoding Levels (ELvl)

<b>MARC 21 Encoding Levels</b>	<b>OCLC-MARC Encoding Levels (ELvl)</b>
blank – Full level	I – Full-level input by OCLC participants
1 – Full level, material not examined	K – Minimal-level input by OCLC participants
2 – Less-than-full level, material not examined	M – Added from a batch process
3 – Abbreviated level	
4 – Core level	
5 – Partial (preliminary) level	
7 – Minimal level	
8 – Prepublication level	

Over the decades, LC loosened its control over many of the Encoding Levels defined in MARC proper. Gradually, the OCLC-defined alphabetic Encoding Levels have become redundant, as well as one of the major areas in which OCLC-MARC remains different from MARC 21 proper.

With valued assistance from market analysts in OCLC's Library Services to the Americas area, we realized that talking with focus groups was the logical way to gain an understanding of our users' practices, needs, concerns, and questions. The analysts organized, scheduled, and ran our focus groups after writing up a market research proposal that included our objectives, the methodology, a timeline, and a budget.

In November 2018, we sent out invitations to various discussion lists for two »virtual discussions« on Encoding Levels, explaining briefly what we had in mind and what information we were looking for. The response from the OCLC cooperative was so overwhelming that we ended up scheduling four focus group sessions. That also allowed us to focus the groups even more closely: one each for public libraries, academic libraries, Association of Research Libraries members (ARLs), and special libraries.

In each of the four sessions, we presented some historical background, an explanation of how important the institutions' input would be in our decision-making, a refresher on the differences between MARC 21 and OCLC-MARC Encoding Levels, and a series of questions that included:

- What sorts of distinctions do you make between MARC 21 and OCLC-MARC Encoding Levels?
- Encoding Levels aside, what other bibliographic elements play a part in determining your workflows, such as particular library identifiers in field 040 or authentication codes in field 042?
- How would this affect your copy cataloging and original cataloging processes?
- How could OCLC help make this a painless process?

In the end, we had 27 librarians from 24 institutions participate in the four focus groups, from a diverse range of libraries in each category. All were incredibly invested in the topic and offered thoughts and opinions in answer to the questions posed. The market analysts

produced a detailed summary of the focus group discussions, which has served to guide our plans.

We learned, or had reinforced, three major things through the focus groups. First, that catalogers were incredibly open to this change, which will affect their workflows and practices much less than we had expected. Second, that they really expect OCLC to provide good documentation and training surrounding this change when we are ready to make it. Third, that they definitely want a visual indication in bibliographic records to tell them that a record has been added to WorldCat via a batch process. Currently, the OCLC-defined Encoding Level M provides that information, but there is no equivalent among the standard MARC Encoding Levels. OCLC needs to figure out an alternative method of conveying this information within the bibliographic record.

In addition to the focus groups, we spoke with a number of our OCLC colleagues in an attempt to understand all of the current uses of Encoding Levels in various OCLC products and services. As one example, the WorldCat Cataloging Partners collections (formerly known as PromptCat), offered through Collection Manager, use Encoding Levels for selection of records to deliver to libraries. Coordinating the Encoding Level changes with other OCLC areas will help us ensure a smooth transition.

Our colleague Robert Bremer has drafted requirements for converting existing records in WorldCat. In April 2020, we opened up the use of the standard MARC Encoding Levels to all cataloging members for both creating new records and editing existing records as the first member-facing step toward the ultimate goal of switching to the standard MARC 21 Encoding Levels (WorldCat, 2020). We have also begun to talk about our plans at various face-to-face meetings, such as the American Library Association (ALA) and Program for Cooperative Cataloging (PCC), to get members of the cooperative used to the changes. We recognize that this will be a real cultural shift for catalogers, but we've also gotten the strong sense that catalogers are eager for the changes to happen.

#### **4 Resolving to reduce duplicates: a short history of duplicate detection and resolution**

What we now call WorldCat was unleashed onto the world on August 26, 1971. But not until a dozen years later in 1983 did OCLC develop the capability to manually merge duplicate bibliographic records. At first, this was an overnight batch operation, but it eventually evolved into an instantaneous process period.

A 1987 user survey conducted by OCLC found that duplicate records were considered to be the most serious of seven quality problems in WorldCat. That prompted us, during the late 1980s and early 1990s, to work with the OCLC Office of Research to develop the first version of automated Duplicate Detection and Resolution (DDR) software. Beginning in June 1991, OCLC used DDR to match WorldCat bibliographic records in the books format against themselves to find and merge duplicates.

That original DDR dealt only with bibliographic records in the books format. It used a series of algorithms that compared fourteen descriptive elements gathered from more than 50 fixed and variable fields and subfields. The fourteen elements included: cataloging library (040 subfield \$c); LCCN; ISBN; government document classification number; media (Form); author;

title; statement of responsibility; edition statement; place of publication and publisher; publication date; number of pages or volumes; size; and series statement. Additionally, there was a set of about a dozen or so »flagged« conditions that prevented merges in situations that we determined to be too risky. Several of the flags resulted in the setting aside of all records for microforms, which were not deduplicated. Other flags took into consideration the degree of internal consistency within each record and between pairs of records, particularly discrepancies of dates, places, publishers, and edition statements. Setting aside the flagged records, the algorithms calculated a numerical level of similarity between records. Record pairs that were not flagged and set aside and had a similarity of 0.94 (out of 1.00) were deduplicated. The retained record was chosen according to an early version of what we now call the Record Retention Hierarchy, which considered Encoding Level (Leader/17), Cataloging Source (008/39), certain values in field 040 subfield \$c (such as DLC and NLM), and other factors. Between mid-1991 and mid-2005, DDR had been run through the bibliographic database sixteen times, resulting in the elimination of some 1,592,586 duplicate records in the Books format.

The original version of DDR was designed to run on the old PRISM platform, which was decommissioned in mid-2005. An entirely new version of DDR was needed to run on the Connexion platform. So, in 2005, a project was begun to re-invent the DDR software for the new environment and to expand its capabilities to deal with all types of bibliographic records, not just books. There were also the intentions to allow targeted deduplication, to make it an ongoing background process rather than an occasional process that had to be run against the entire WorldCat catalog each time, and to take advantage of technological advances since the early 1990s.

The project involved roughly five years of rigorous planning, careful development, exhaustive testing, and extensive record cleanup. In May 2009, we put the new software into limited production, processing small targeted subsets of WorldCat. Between May 2009 and January 2010, we processed about 500,000 records in this way, resulting in about 15,000 duplicates merged, *every one of which was examined individually for accuracy*. Each correct merge confirmed some aspect of our work, but even more importantly, we learned something from every incorrect merge. We pulled the records apart, fine-tuned the software, and retested so that each particular problem would not occur again if it could be prevented.

Once we were confident that the new DDR was performing accurately, the full processing of WorldCat began in late January 2010 in two parallel processes. One process, which we called »walking the database«, began with OCLC Record #1 and traversed all of WorldCat. That completed in September 2011, processing over 166 million bibliographic records, resulting in 5,126,132 duplicates being deleted. The other process looked at each day's newly added records plus records updated in ways that could affect matching. This processing continues now every day, although now there is a deliberately built-in delay of about a week that allows records to »settle in«. Through the end of February 2020, DDR has processed nearly 843 million records, resulting in over 50.1 million duplicate bibliographic records deleted.

It's interesting to note that the percentages of duplicates found and merged in both the 500,000 targeted records during 2009 and in »walking the database« during 2010 and 2011 were both roughly the same, at 3%. The overall percentage of duplicates merged since 2009 now stands at 5.94%.

As mentioned earlier, the original DDR compared about fourteen descriptive elements drawn from over fifty fixed and variable fields and subfields. Current DDR, now roughly a decade old, is much more sophisticated. Although this is an oversimplification of a complex process, there are now at least two dozen different points of comparison taken into consideration. By my really rough estimate, I've figured that of the over 270 defined MARC 21 fixed and variable fields, data in well over half of them can figure into DDR in some way. Obviously, many of the two dozen or so comparison points draw data from multiple parts of a bibliographic record and involve manipulation of data in ways designed to distinguish both variations that should be equated and distinctions that must be recognized. There are special comparison points designed specifically for certain bibliographic formats. Just as examples: scale is considered for maps; aspects of instrumentation and the presence of parts are considered for scores; screen presentation formats, region information, and color broadcast systems are considered for videorecordings.

Many of the augmented comparison points rely on the vastly expanded indexing capabilities that were introduced with Connexion. All sorts of detailed analyses and comparisons of data that were impossible in earlier years can now be performed. Because now all bibliographic formats are being processed, not just books, matching elements that were not relevant previously are now taken into consideration, and most bibliographic elements had to be re-examined within a wider context of additional bibliographic formats.

Not only does the current DDR consider more comparison points overall, but the work it performs within each comparison point is more complex. To give you just one small but typical example, in the comparison of titles alone, well over 60 different conditions are looked for, each one of which triggers its own particular reaction in the software. These conditions involve things as obvious as equating various ways to express numbers; and equating ampersands, plus signs, and the word »and« to trying to account for single-character differences in such geographic names as 'Lane County' versus 'Lake County'.

As mentioned earlier, the original DDR had a set of about a dozen or so »flagged« conditions that prevented merges in specific risky situations. Current DDR doesn't have »flags« in the same way, but there are several categories of bibliographic records that are exempted from DDR processing, including:

- Digital Collection Gateway records (identified by certain codes in field 029 subfield \$t).
- SCIPRO records for art and rare book sales catalogs (identified by code 'scipro' in field 042).
- Records for photographs (identified by either of two Material Types, 'pht' for photograph and 'pic' for picture).
- All records with dates of publication earlier than 1801; in consultation with the American Library Association Rare Books and Manuscripts Section's Bibliographic Standards Committee, OCLC is planning to change this cutoff date to 1830.
- Maps records with dates of publication earlier than 1901, determined in consultation with the ALA Map and Geospatial Information Round Table (MAGIRT).
- Records cataloged under any of 25 rare/archival descriptive conventions identified in field 040 subfield \$e.



In our work on both versions of DDR, we always tried to err on the side of caution, deciding not to merge when we could not be sure. Of course, DDR is an automated process, and no matter how painstakingly we have planned, designed, developed, and tested, no such process can ever be as exacting as a human cataloger. In a database as large and diverse as WorldCat, no one can anticipate all of the possible permutations of data that might be encountered. And although we are constantly making changes to our algorithms based on reports of erroneous merges, we long ago learned that every »improvement« made to the algorithms has some corresponding cost. Sometimes a change means losing some otherwise perfectly legitimate matches. Other times, a change will result in erroneous matching in an unanticipated situation. Given the sheer size of WorldCat, the wide expanse of cataloging practices, and the seemingly infinite variety of errors that both human and automated creators of bibliographic records can make, DDR is imperfectible, in spite of our best efforts.

Just as humans may create duplicate records in WorldCat, however, humans can help to eliminate duplicates. Metadata Quality staff have been merging duplicates, as noted earlier, since 1983. But since fiscal year 2014, members of the OCLC cooperative who also participate in the Program for Cooperative Cataloging (PCC) have been working as part of the *Member Merge Project* (MMP) (2019). MMP participants receive special training through a series of webinars with Metadata Quality staff and/or others involved in MMP covering merging documentation and policies, followed by a rigorous review period that results in independence. During the first half of fiscal year 2020, MMP participants have averaged about 1500 merges per month.

## **5 Bibliographic Formats and Standards and »When to Input a New Record«**

*Bibliographic Formats and Standards* (BFAS) (2020) is OCLC's documentation of OCLC-MARC, its implementation of the Library of Congress's *MARC 21 Format for Bibliographic Data* (2019). BFAS serves as a guide to bibliographic record structure, input standards, and coding practices for the WorldCat bibliographic database. When *Resource Description and Access* (RDA) was implemented in 2013, OCLC's Metadata Quality staff undertook a comprehensive multiyear project to review, revise, reorganize, and update BFAS. Among many other things, our intention has been to add references to RDA and account for RDA in bibliographic examples.

As described earlier in this article, BFAS is kept up to date when new MARC fields, subfields, and other elements are implemented as part of a MARC Update. Just as important, the ongoing BFAS revision has also included updates to the five narrative chapters that outline WorldCat online cataloging, special cataloging guidelines for particular types of resources and situations, bibliographic quality assurance in the cooperative environment of WorldCat, and »When to Input a New Record«.

OCLC's »When to Input a New Record«, Chapter 4 of *Bibliographic Formats and Standards* (BFAS) (2020), has long served to provide a common basis for decision-making in the creation of the WorldCat bibliographic database by participants in the OCLC cooperative. »When to Input ...« has also been the public reflection of how OCLC's matching algorithms (including Duplicate Detection and Resolution (DDR) and the automated loading of records) are generally intended to work. As part of our ongoing thorough revision of BFAS, in October 2017 we made available an updated version of »When to Input a New Record«. Of course, it is impossible to

cover every possible case in a document such as »When to Input ...« but we have tried to account for as many of the most common ones as possible.

OCLC's »When to Input a New Record« first appeared in July 1983 as a new eleven-page appendix added to the 1982 Second Edition of the document *Bibliographic Input Standards*, which was one of the predecessor publications that would be combined into *Bibliographic Formats and Standards* in 1993. In 2004, the Association for Library Collections and Technical Services (ALCTS) first published *Differences Between, Changes Within: Guidelines on When to Create a New Record* (2007), which was intended to supplement the descriptive cataloging rules of AACR2. The document, which was revised in 2007 and was maintained by an ALCTS task force, provides guidance to the cataloger who has found copy that is a close match to the item in hand about whether to use that copy or to create a new bibliographic record.

*Differences Between, Changes Within* is a valuable supplement to OCLC's »When to Input ...«, but does not replace it for members of the OCLC cooperative. On most major points, the two documents agree. Because of the unique cooperative nature of WorldCat and its application of a master record concept, however, there are several areas in which OCLC has chosen to differ. OCLC has requested that users follow OCLC practice in these instances.

OCLC's series of presentations with the overall title of *Cataloging Defensively* (2020) can come in handy in cases where you have determined that a separate bibliographic record is justified but you need to make sure that your new record is properly differentiated from other similar records. Included on the *Cataloging Defensively* web page are presentations devoted specifically to maps, sound recordings, musical scores, video recordings, and edition statements, as well as the more general presentation from 2010.

The cardinal rule of »When to Input a New Record« is »When in doubt, do not create a duplicate; use an existing record«. WorldCat has always relied upon the expertise of individual catalogers making reasonable judgments about bibliographic resources. We know that human catalogers are able to make more informed choices than any automated process. This human factor is the main reason that this cardinal rule for »When to Input a New Record« is the exact opposite of the cardinal rule for our automated Duplicate Detection and Resolution (DDR) and automated matching, which is to err on the side of either letting a potential duplicate remain in, or adding one to, WorldCat when there is uncertainty.

Errors in a record do not justify the creation of a duplicate. We encourage catalogers to correct the existing record when able or to report the errors to OCLC via [bibchange@oclc.org](mailto:bibchange@oclc.org). Differences indicative of a distinct bibliographic item usually (but not always) occur in more than one field. If a difference occurs in a single field, catalogers are asked to do their best to determine whether there are two separate bibliographic items with only one significant difference or whether the difference is an error or a difference of opinion.

For better or worse, there is a lot of redundancy built into a MARC Bibliographic record. An internally consistent record will often have corroborating data in multiple fields, such as cartographic scale and coordinates in coded form in field 034 and in more human-friendly form in field 255. Catalogers must be alert for contradictory data within the record, such as discrepancies in dates between the fixed field and the 26X field. Responsible catalogers will always consider the record in its entirety, and when in doubt, will not create a duplicate record, but instead use the existing record.

## 6 Conclusion

The small staff of OCLC's WorldCat Metadata Quality, with the help of members of the OCLC cooperative, work together to try to improve the quality of data in WorldCat through both manual and automated means. Keeping WorldCat validation current with the evolving specifications of MARC 21, working to bring OCLC-MARC into closer alignment with MARC 21 in such areas as Encoding Levels, eliminating bibliographic duplicates, and revising such OCLC documentation as *Bibliographic Formats and Standards* (2020) are just a few of the ongoing efforts that keep the metadata quality in WorldCat.

## References

- Authorities: format and indexes*, 2018. Ohio: OCLC. Available at: [https://help.oclc.org/Metadata\\_Services/Authority\\_records/Authorities\\_Format\\_and\\_indexes](https://help.oclc.org/Metadata_Services/Authority_records/Authorities_Format_and_indexes) [20.03.2020].
- Bibliographic formats and standards*, 2020. Ohio: OCLC. Available at: <https://www.oclc.org/bibformats/en.html> [20.03.2020].
- Cataloging defensively*, 2020. Ohio: OCLC. Available at: [https://help.oclc.org/WorldCat/Cataloging\\_documentation/Cataloging\\_defensively](https://help.oclc.org/WorldCat/Cataloging_documentation/Cataloging_defensively) [20.03.2020].
- Differences between, changes within: guidelines on when to create a new record*, 2007. Chicago: ALCTS. Available at: <http://www.ala.org/alcts/sites/ala.org.alcts/files/content/resources/org/cat/differences07.pdf> [20.03.2020].
- FAST (Faceted Application of Subject Terminology)*, 2020. Ohio: OCLC. Available at: <https://www.oclc.org/research/themes/data-science/fast.html> [20.03.2020].
- LC guidelines supplement to the MARC 21 format for authority data*, 2018. 2002 Edition with subsequent updates. Washington: LC. Available at: <https://www.loc.gov/catdir/cpsolcmarcsuppl.pdf> [20.03.2020].
- MARC 21 format for bibliographic data*, 2019. Washington: LC. Available at: <https://www.loc.gov/marc/bibliographic/> [20.03.2020].
- MARC Advisory Committee*, 2016. Available at: [http://www.loc.gov/marc/mac/MAC\\_ToR.html](http://www.loc.gov/marc/mac/MAC_ToR.html) [20.03.2020].
- MARC development*, 2019. Washington: LC. Available at: <http://www.loc.gov/marc/mac/index.html> [20.03.2020].
- MARC discussion paper no. 2016-DP09*, 2016. Available at: <https://www.loc.gov/marc/mac/2016/2016-dp09.html> [20.03.2020].
- MARC proposal no. 2016-05*, 2016. Available at: <https://www.loc.gov/marc/mac/2016/2016-05.html> [20.03.2020].
- Member merge project*, 2019. Ohio: OCLC. Available at: [https://help.oclc.org/WorldCat/Metadata\\_Quality/Member\\_Merge](https://help.oclc.org/WorldCat/Metadata_Quality/Member_Merge) [20.03.2020].

*Record manager authorities guide*, 2020. Ohio: OCLC. Available at:  
[https://help.oclc.org/Metadata\\_Services/WorldShare\\_Record\\_Manager/Authority\\_records/Record\\_Manager\\_Authorities\\_Guide/01Introduction](https://help.oclc.org/Metadata_Services/WorldShare_Record_Manager/Authority_records/Record_Manager_Authorities_Guide/01Introduction) [30.03.2020].

*OCLC-MARC local holdings format and standards*, 2018. Ohio: OCLC. Available at:  
[https://help.oclc.org/Metadata\\_Services/Local\\_Holdings\\_Maintenance/OCLC\\_MARC\\_local\\_holdings\\_format\\_and\\_standards](https://help.oclc.org/Metadata_Services/Local_Holdings_Maintenance/OCLC_MARC_local_holdings_format_and_standards) [20.03.2020].

*Searching WorldCat indexes*, 2019. Ohio: OCLC. Available at:  
[https://help.oclc.org/Librarian\\_Toolbox/Searching\\_WorldCat\\_Indexes](https://help.oclc.org/Librarian_Toolbox/Searching_WorldCat_Indexes) [20.03.2020].

*WorldCat Validation release notes, April 2020*, 2020. Ohio: OCLC. Available at:  
[https://help.oclc.org/Metadata\\_Services/WorldShare\\_Record\\_Manager/WorldCat\\_Validation\\_release\\_notes\\_and\\_known\\_issues/2020\\_Release\\_notes/090WorldCat\\_Validation\\_release\\_notes\\_April\\_2020](https://help.oclc.org/Metadata_Services/WorldShare_Record_Manager/WorldCat_Validation_release_notes_and_known_issues/2020_Release_notes/090WorldCat_Validation_release_notes_April_2020) [04.03.2020].

# Korisnici u COBISS okruženju – iskustva Narodne biblioteke Užice

## Users in the COBISS environment – experiences at the Public Library of Užice

**Gordana Bacotić<sup>1</sup>, Biljana Ristović<sup>2</sup>**

**SAŽETAK:** Narodna biblioteka Užice anketirala je 2019. godine svoje korisnike s ciljem da provjeri koliko oni poznaju proces pretraživanja kataloga u aplikaciji COBISS/OPAC (sada COBISS+). Istraživanje se sastojalo iz sledećih segmenata: profil korisnika, najčešće korišćene pristupne tačke, korišćenje ekranskih saveta i olakšica pri pretraživanju i korisnička procena dobijenih rezultata. Uprkos tome što se obuka korisnika za pretraživanje u COBISS/OPAC kontinuirano sprovodila, rezultati istraživanja su pokazali da mogućnosti, koje je nudio COBISS/OPAC, nisu bile dovoljno iskorišćene. Korišćenje bibliotečkih kataloga zahteva veće znanje i angažovanje nego pretraživanje na internetu. Katalog se koristi za pronalaženje već poznatih jedinica građe u fondu Biblioteke, a ne za složena tematska pretraživanja koja su prepuštena bibliotekarima.

**KLJUČNE REČI:** COBISS/OPAC, katalog, korisnici, pretraživanje, anketa, Narodna biblioteka Užice

**ABSTRACT:** In 2019, the Public Library of Užice conducted a survey to assess the level of user knowledge of the search process in COBISS/OPAC (now COBISS+). The survey consisted of the following segments: user profile, most commonly used access points, use of screen tips and ease of searching, and user evaluation of the search results obtained. Although user training for searching COBISS/OPAC was carried out continuously, the survey results have shown that the options in COBISS/OPAC were not sufficiently used. Using library catalogues requires more knowledge and engagement than searching the Internet. The catalogue is used for finding already known items in the library's holdings, and not for topical searches, which are left to librarians.

**KEYWORDS:** COBISS/OPAC, catalogue, users, search, survey, Public Library of Užice

---

## 1 Uvod

Razvoj informacionih tehnologija doneo je značajne promene u bibliotečkom poslovanju, a posebno u izradi kataloga i njihovom korišćenju. Između ostalog, pojava uzajamne katalogizacije omogućila je veliku uštedu u vremenu pri obradi bibliotečke građe; lisni katalozi su zamenjeni elektronskim kojima se pristupa i preko interneta, što bi trebalo da predstavlja veliku olakšicu, ne samo za bibliotekare, već i za korisnike bibliotečkih usluga. U lisnim katalozima, zapisima se pristupalo samo preko autora, naslova i predmeta (elementi za način ulaganja listića u katalog), dok u onlajn katalogu različiti elementi opisa mogu biti pristupne tačke (Svenonius, 2007, str. 44). Uprkos tome što bibliotečki katalozi pružaju različite mogućnosti pretraživanja, njihovo je korišćenje daleko manje od očekivanog, jer je njihova upotreba za korisnike složenija od upotrebe internet pretraživača koji su postali neizbežni u svakodnevnom profesionalnom i privatnom životu. Pored očekivanja da će se povećavati broj

---

<sup>1</sup> Kontakt autor: Gordana Bacotić, viši diplomirani bibliotekar, Narodna biblioteka Užice, Užice, Srbija [gbacotic@gmail.com](mailto:gbacotic@gmail.com).

<sup>2</sup> Biljana Ristović, viši diplomirani bibliotekar, Narodna biblioteka Užice, Užice, Srbija, [biljana.ristovic76@gmail.com](mailto:biljana.ristovic76@gmail.com).

korisnika kataloga, naročito zbog toga što su se katalogi tokom godina razvijali, menjali i donekle prilagođavali korisničkom načinu razmišljanja, prema iskustvu zaposlenih u Narodnoj biblioteci Užice, korisnici se i dalje u velikoj meri oslanjaju na pomoć bibliotekara.

U Srbiji je 2019. godine više od 220 biblioteka bilo uključeno u COBISS sistem koji bibliotekarima i korisnicima omogućava onlajn pristup, ne samo lokalnim bazama podataka, već i uzajamnoj bibliografsko-kataloškoj bazi COBIB i specijalizovanim bazama podataka.

Narodna biblioteka Užice deo je sistema uzajamne katalogizacije COBISS od 2004. godine i prva je javna biblioteka koja se priključila ovom sistemu (Ršumović, 2006, str. 206), čime je njen fond postao vidljiv i pretraživ preko interneta. Sve do polovine 2019. godine korišćena je aplikacija COBISS/OPAC koja je kasnije zamenjena aplikacijom COBISS+.

Pretraživanje onlajn kataloga COBISS/OPAC bilo je prilagođeno različitim zahtevima i stepenu znanja korisnika (Šokica, 2005, str. 34). Postojala su tri načina pretraživanja COBISS/OPAC: osnovno, izborno i komandno, što je uz određene modifikacije ostalo i u COBISS+. Osnovno pretraživanje je najjednostavnije i prilagođeno je korisnicima koji pretražuju rečima ili sintagmama koje se nabrajaju. Za ovo pretraživanje nije potrebno veće predznanje, jer podseća na okvir za pretragu (*search box*) internet pretraživača. Izborno pretraživanje je nešto složenije, podrazumeva korišćenje Bulovih operatora i pruža mogućnost kombinovanja niza parametara. Podrazumevani operator je »/« (*And*), a za primenu ostalih operatora potrebno je veće predznanje i angažovanje korisnika.

Komandno pretraživanje je najsloženija vrsta pretraživanja koja podrazumeva upotrebu indeksa za pretraživanje (prefiksi i sufiksi) i kontekstualnih operatora *With* i *Near* (u COBISS+ aplikaciji nema operatora *With*). »Mogućnosti pretraživanja u komandnom načinu rada su velike i njih koriste uglavnom bibliotekari.« (Šokica, 2005, str. 40). Za ovu vrstu pretraživanja potrebno je obučiti korisnike, a ponekad se traži i poznavanje bibliografsko-kataloške obrade.

Kada je reč o pretraživanju, prelazak na COBISS+ verziju doneo je više promena. Prema prvim komentarima korisnika u Narodnoj biblioteci Užice, novi interfejs je donekle teži za snalaženje, jer su neke opcije na početnom ekranu manje upadljive (na primer, link za izborno i komandno pretraživanje). Katalog, koji je prilagođen korisniku, trebalo bi da omogući jasna i uvek dostupna uputstva tokom procesa pretraživanja. Ako čitanje uputstava podrazumeva višestruko otvaranje i zatvaranje prozora, to ometa i usporava pretraživanje, posebno ako je uputstvo detaljno. U verziji COBISS/OPAC saveti za pretraživanje nalazili su se s desne strane prozora za pretraživanje, dok su u COBISS+ manje uočljivi, a takođe nije eksplicitno naglašeno da link *više informacija*, koji postoji na pristupnoj strani, vodi ka savetima za pretraživanje. Čitanje uputstva podrazumeva izlazak iz prozora za pretraživanje, što dodatno otežava pretragu.

Između ostalog, prednosti COBISS+ verzije su mogućnost dodavanja novih polja u prozoru za pretragu u izbornom pretraživanju, fasetna navigacija kojom se filtriraju dobijeni rezultati pretraživanja, a takođe brojne mogućnosti personalizovane pretrage povezane s profilom *Moj COBISS*.

Neki suštinski problemi s kojima se korisnici susreću prilikom pretrage kataloga ostali su, međutim, nepromenjeni, a time se ovaj rad i bavi.

## 2 COBISS i obuka korisnika u Narodnoj biblioteci Užice

Prelazak s lisnih na elektronske kataloge počeo je u Narodnoj biblioteci Užice sredinom devedesetih godina 20. veka. U početku je korišćen Uneskov softver za bibliotečko poslovanje (DOS verzija) koji je bio previše zahtevan za korisnike. U praksi, korisnici nikada nisu samostalno pretraživali katalog korišćenjem ovog sistema. Prema iskustvu zaposlenih, ni prelazak na znatno jednostavniju COBISS/OPAC aplikaciju nije doprineo značajnijem osamostaljivanju korisnika pri pretraživanju.

Odmah posle obuke bibliotekara za rad u COBISS/OPAC okruženju, započela je obuka zainteresovanih korisnika za samostalnu pretragu COBISS/OPAC, u cilju njihovog osposobljavanja za samostalno pronalaženje i pravilno korišćenje izvora informacija. Edukacija korisnika je permanentan proces koji se odvija na dva načina:

- individualno (u prostorijama Biblioteke);
- grupno (za učenike osnovnih i srednjih škola) u školskim informatičkim kabinetima.

Individualna obuka za pretraživanje u COBISS/OPAC odvijala se, uz redovan posao, u svim odeljenjima i bila je uslovljena vremenom koje su na raspolaganju imali korisnik i bibliotekar. Uglavnom se svodila na nekoliko osnovnih koraka: upoznavanje s lokalnom i uzajamnom bibliografsko-kataloškom bazom i razlikama između njih, s prozorom i opcijama za pretraživanje, s podacima o smeštaju i dostupnosti građe. Korisnicima, koji su tražili stručnu literaturu, objašnjena je razlika između ključne reči i predmetne odrednice. Takođe se skretala pažnja na skraćivanje termina zvezdicom (\*), na upotrebu Bulovih operatora, na vezu s bazama punog teksta. Korisnicima smo najčešće pokazivali izbornu pretraživanje jer je, za razliku od osnovnog, bilo više detaljno i pregledno, a manje komplikovano od komandnog. Zainteresovani su imali mogućnost da dobiju i detaljna uputstva od bibliotekara, a svima smo ukazivali na činjenicu da postoje saveti za pretraživanje. Korisnici su najčešće pokazivali interesovanje za učenje pretraživanja, ali i relativno kratkoročno strpljenje za detaljnija objašnjenja.

Grupna obuka bila je namenjena prvenstveno završnim razredima osnovnih i srednjih škola, a organizovana je i za nastavno osoblje i školske bibliotekare. Sprovedena je preko projekta *Pronađi i koristi informaciju* koji je podržalo Ministarstvo kulture i informisanja Republike Srbije 2014. godine. Ova obuka nastavljena je i po završetku projekta i ujedno postala redovna aktivnost u Narodnoj biblioteci Užice. Do sada je njome obuhvaćeno preko 2000 učenika užičkih osnovnih i srednjih škola (Davidovski i Marinčić, 2019, str. 139). U okviru jednog od redovnih poslova Matične službe Narodne biblioteke Užice, sprovedena je edukacija učenika osnovnih i srednjih škola na teritoriji Zlatiborskog okruga za pretraživanje COBISS/OPAC, pod nazivom *Pronađi pravu informaciju*.

Obuka za pretraživanje u COBISS/OPAC bila je segment i u okviru drugih projekata: Ministarstvo kulture i informisanja podržalo je 2014. godine projekat *Biblioteka – mesto susreta svih* za učenike osnovnih škola u seoskim ogradnicama (Ribaševina, Karan i Lunovo Selo); na Dečjem odeljenju realizovan je projekat *Biblioteka – mesto za učenje* (Ministarstvo kulture i informisanja Republike Srbije 2018. godine); Grad Užice podržao je 2012. godine projekat *Profil, anfas, poluprofil* namenjen učenicima osnovne škole na području Knjižnog ogranka Sevojno.

Grupna obuka je izvođena u kabinetima za informatiku tokom jednog školskog časa po unapred utvrđenom programu. Kada škola nije bila povezana sa internetom, što se retko događalo, učenicima je COBISS/OPAC predstavljen preko pripremljene prezentacije. Grupna obuka je bila strukturisana i detaljna, s predviđenim vremenom za vežbu, međutim, motivisanost učenika često je bila na nižem nivou u odnosu na korisnike koji su pokazivali interesovanje za individualnu obuku.

U okviru projekta *Mladi u biblioteci* koji je finansirao Grad Užice, Narodna biblioteka Užice je pokrenula i YouTube kanal, kao vid promovisanja svojih usluga, gde je jedan od video zapisa posvećen pretraživanju COBISS/OPAC (<https://www.youtube.com/watch?v=AwqyGq6oh8c>).

Svi oblici edukacije pokrenuti su zato što je iskustvo u radu pokazalo da je mali procenat korisnika pretraživao COBISS/OPAC i što su korisnici bili relativno nesamostalni pri pretraživanju. Osim toga, urađena je i empirijska studija u cilju sticanja što preciznijeg uvida u iskustva korisnika pri pretraživanju baze podataka.

### 3 Koncept i metodologija

Krajem 2018. i u prvoj polovini 2019. godine, u Narodnoj biblioteci Užice sprovedeno je istraživanje među korisnicima u cilju utvrđivanja njihovog poznavanja procesa pretrage COBISS/OPAC i zadovoljstva rezultatima pretrage. Konkretni cilj istraživanja bilo je ispitivanje snalaženja prosečnog korisnika usluga Narodne biblioteke Užice u COBISS/OPAC okruženju.

U istraživanju su postavljene sledeće hipoteze:

1. COBISS/OPAC se najčešće koristi za pretraživanje po parametrima poznatim korisniku (korisnici pretražuju da li biblioteka poseduje određeni naslov ili dela određenog autora);
2. korisnici ne koriste sve mogućnosti pretraživanja (Bulovi operatori, skraćivanje);
3. pretraživanje po temi suviše je složeno za korisnika;
4. korisnici nisu zadovoljni rezultatima pretrage COBISS/OPAC.

Korišćena je metoda ankete. Prva faza istraživanja podrazumevala je usmeno ispitivanje korisnika starijih od 15 godina na Pozajmnom odeljenju za odrasle, s ciljem da se ustanovi da li pretražuju COBISS/OPAC. One, koji su se pozitivno izjasnili, zamolili smo da popune anketu. Po sistemu slučajnog uzorka obuhvaćeno je 100 korisnika različitog pola, uzrasta, obrazovanja i interesovanja. Anketiranje je sprovedeno od decembra 2018. do septembra 2019. godine i trajalo je duže od predviđenog vremena, jer se mali broj korisnika izjasnio da pretražuje u COBISS-u.

Anketnom obrascu u elektronskom obliku (onlajn Google upitnik), korisnici su pristupali u prostorijama Biblioteke, preko računara koji je namenjen korisnicima. Anketa je sadržala 15 pitanja zatvorenog tipa i sastojala se iz sledećih segmenata: profil korisnika (pol, zanimanje), korišćenje saveta i olakšica pri pretraživanju (zvezdica, Bulovi operatori), najčešće korišćene pristupne tačke (poseban osvrt na ključnu reč i predmetnu odrednicu), korisnička procena procesa pretraživanja i zadovoljstvo dobijenim rezultatima.

Odgovor na svako pitanje bio je obavezan. Bibliotekar je prisustvovao anketiranju i davao dodatna objašnjenja onda kada je korisnik to tražio. Izvršena je osnovna kvantitativna



statistička analiza odgovora. U deskriptivnoj analizi, kao statistička tehnika korišćeni su procenti.

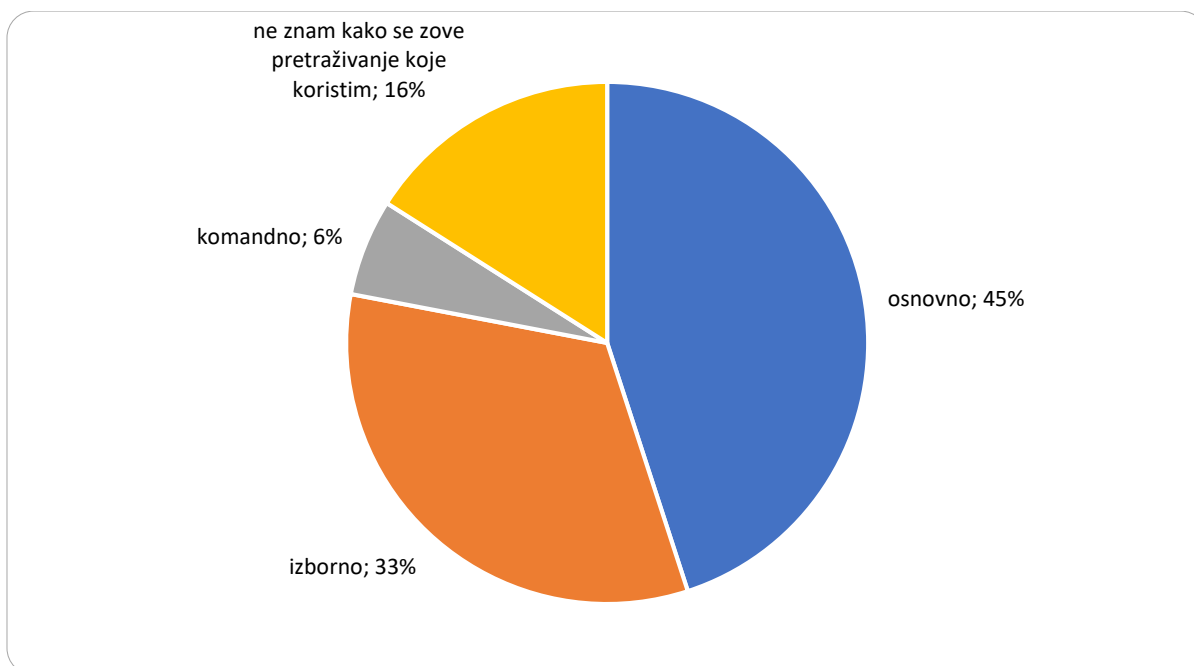
#### 4 Rezultati istraživanja

U istraživanju je učestvovalo 100 ispitanika, od kojih je 65% bilo ženskog i 35% muškog pola. Ovaj odnos približan je polnoj strukturi ukupnog broja korisnika Biblioteke (66% prema 34%). Anketirano je 37 zaposlenih, 28 studenata, 19 učenika, 11 nezaposlenih i 5 penzionera.

Na pitanje kako su naučili da koriste COBISS/OPAC, 45% je odgovorilo da je naučilo samostalno, a 55% uz pomoć bibliotekara. Najveći procenat ispitanika, koji su tvrdili da su samostalno naučili da pretražuju, spada u kategorije studenata i zaposlenih, koji su u najproduktivnijem periodu učenja i usavršavanja u struci.

Kada je reč o korišćenju saveta za pretraživanje, koji su se u COBISS/OPAC nalazili pored maske za pretraživanje, od ukupnog broja ispitanika, 37% nije primetilo savete za pretraživanje koji se nalaze na stranici na kojoj se pretražuje, a 41% ih nije koristilo, iako su znali da postoje, dok je samo 22% koristilo ovu vrstu pomoći.

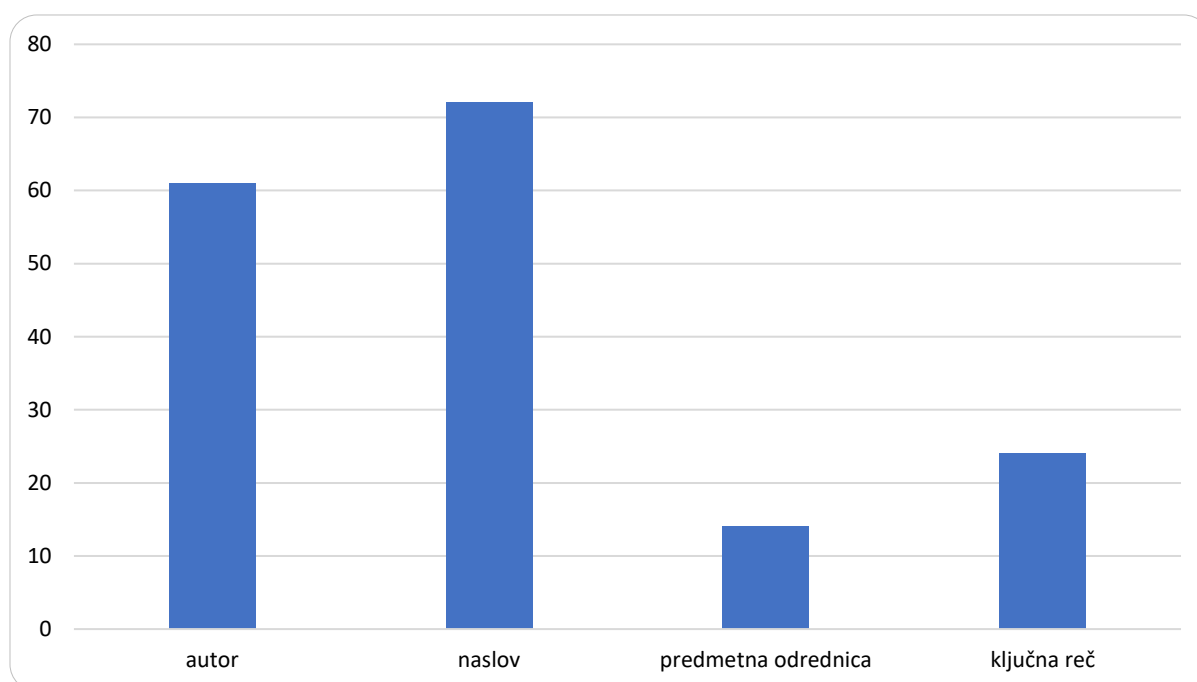
U vezi s načinom pretraživanja u COBISS/OPAC, ispitanici su se u većini izjasnili da koriste osnovno pretraživanje (slika 1). Ispostavilo se da 16% ispitanika ne zna kako se zove pretraživanje koje koriste, a kroz razgovor je utvrđeno i da se nazivi pretraživanja mešaju. Zaključak je da većina koristi izborno pretraživanje na koje i bibliotekari upućuju prilikom obuke, jer je najpreglednije i usklađeno je s mogućnostima korisnika. Ovaj rezultat, takođe, ukazuje na činjenicu da obaveštenja na interfejsu u verziji COBISS/OPAC nisu bila dovoljno uočljiva (npr. različita boja, ikone), kako bi na sebe skrenula pažnju.



Slika 1: Načini pretraživanja u COBISS/OPAC (n=100, u %)

Kada je reč o oblastima koje ispitanici pretražuju, najviše su zastupljene društvene nauke (51%), zatim sledi beletristika (41%), a najmanje se pretražuju teme iz prirodnih i primenjenih nauka (8%). Fond javnih biblioteka čini 50% beletristike koja se prema godišnjim izveštajima najčešće i pozajmljuje. Ova oblast je najtraženija, ali prema rezultatima ankete nije i najčešće pretraživana, jer se pri izboru beletristike čitaoci najčešće oslanjaju na preporuku i pomoć bibliotekara. Iako su liste najčitanijih knjiga i preporuke za čitanje dostupni u svim medijima, korisnici i dalje veruju preporuci bibliotekara, jer je ona prilagođena njihovom individualnom ukusu. Kroz višegodišnju interakciju, bibliotekar upoznaje afinitete korisnika, a istovremeno ima uvid u kvalitetne naslove koji nisu medijski praćeni.

Na pitanje o najčešće korišćenim pristupnim tačkama prilikom izbornog pretraživanja u COBISS/OPAC, ispitanici su imali mogućnost višestrukog odgovora. Rezultati pokazuju da korisnici najčešće pretraživanjem proveravaju da li biblioteka poseduje određeni naslov ili dela određenog autora, dok je pretraživanje po temi ili oblasti (preko predmetne odrednice ili ključne reči) znatno ređe (slika 2).



Slika 2: Najčešće korišćene pristupne tačke prilikom pretraživanja u COBISS/OPAC (n=100)

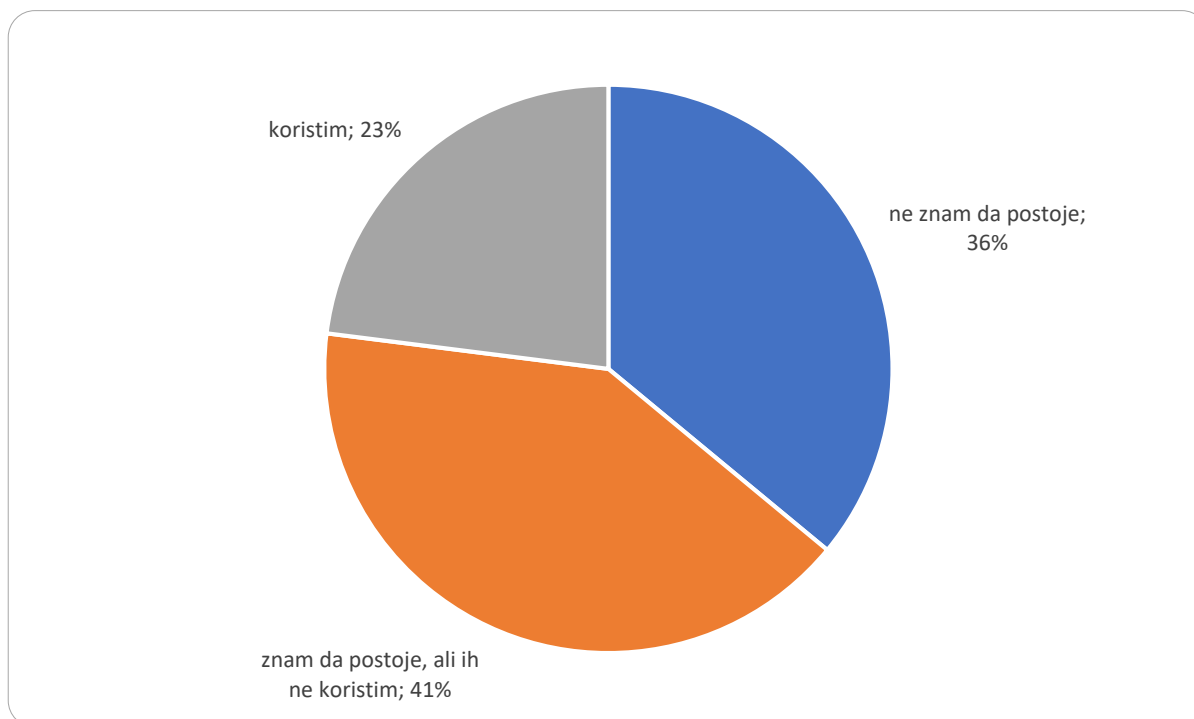
Sledeća tri pitanja u anketi odnosila su se na upotrebu predmetne odrednice kao pristupne tačke pri pretraživanju. Na pitanje *Koristite li opciju "predmetna odrednica" pri pretraživanju* 29% ispitanika odgovorilo je da koristi, 47% da ne koristi, a 24% izabralo je opciju *ne znam šta je predmetna odrednica*, što znači da više od dve trećine anketiranih ne koristi predmetnu odrednicu kao pristupnu tačku. Među onima koji koriste predmetnu odrednicu, najbrojnija su lica iz kategorije studenata i zaposlenih.

Kada je reč o korišćenju ključne reči i predmetne odrednice pri pretraživanju, kada su autor i naslov dela nepoznati, 89% ispitanika se odlučuje za ključnu reč, a samo 11% za predmetnu odrednicu, što je dvostruko manje od procenta ispitanika koji su se u prethodnom pitanju

izjasnili da koriste predmetnu odrednicu. Samo je jedan ispitanik iz kategorije studenata dao prednost predmetnoj odrednici u odnosu na ključnu reč, a ista osoba najčešće pretražuje oblast prirodnih i primenjenih nauka.

Treće, test pitanje, odnosilo se na procenu odnosa broja pogodaka dobijenih upotrebom ključne reči i broja pogodaka dobijenih upotrebom predmetne odrednice. Ispitanici su odgovarali na pitanje *Više pogodaka može se dobiti pretragom preko ključne reči ili predmetne odrednice?* Većina odgovora (80%) prednost je dala ključnoj reči. Zanimljiv je podatak da od 11% ispitanika, koji su u prethodnom pitanju (*Kada vam je nepoznat autor i naslov dela češće pretražujete ključne reči ili predmetne odrednice?*) dali prednost predmetnoj odrednici u odnosu na ključnu reč, samo 2% smatra da se više pogodaka dobija preko ključne reči, što ukazuje na činjenicu da većina ne razume koncept predmetne odrednice.

Naredna dva pitanja odnosila su se na korišćenje olakšica pri pretraživanju. Tokom obuke, koju individualno ili grupno sprovode bibliotekari, najviše pažnje se poklanja skraćivanju pojma zvezdicom i *upotrebi* Bulovih operatora. Kada je reč o korišćenju zvezdice za skraćivanje pojma za pretraživanje, ispitanici su se izjasnili na sledeći način: 25% koristi zvezdicu, a 75% je ne koristi. Provera korišćenja Bulovih operatora formulisana je na način blizak korisnicima: *Znate li da postoje opcije i, ili, ne pri pretraživanju i da li ih koristite?* Zanimljivo je da čak 64% ispitanika poznaje tu opciju, a da je relativno malo (23%) koristi (slika 3).

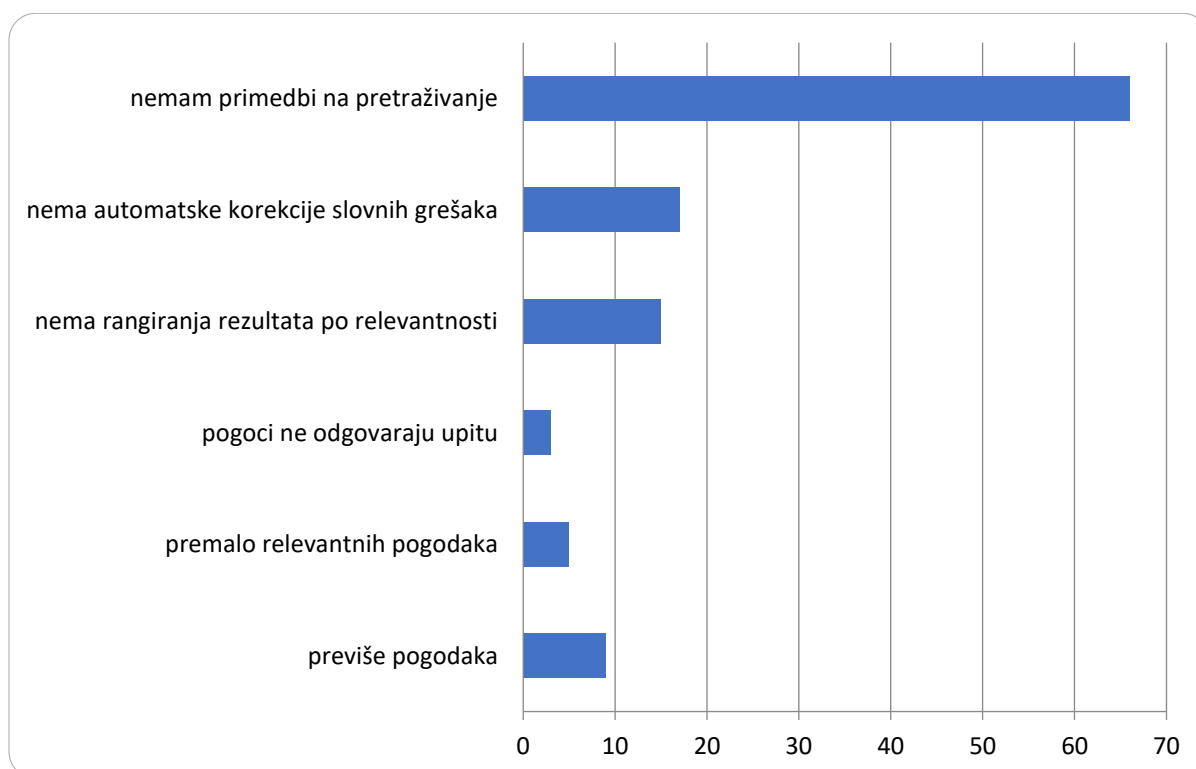


Slika 3: Korišćenje Bulovih operatora u COBISS/OPAC (n=100, u %)

Poslednji segment ankete odnosio se na primedbe u vezi s pretraživanjem i na subjektivni utisak korisnika o prilagođenosti pretraživanja njihovim mogućnostima i potrebama. Apsolutna većina korisnika (92%) zadovoljna je rezultatima pretraživanja u COBISS/OPAC, 8% je delimično zadovoljno, a nezadovoljnih nema. Analiza odgovora anketiranih korisnika, koji

su delimično zadovoljni pretragom, dala je sledeće rezultate: 2 koristi uputstva, 2 koristi Bulove operatore, 3 koristi skraćivanje zvezdicom.

U pretposljednem pitanju, ispitanici su se izjašnjavali o eventualnim nedostacima pretraživanja u COBISS/OPAC. Na osnovu dosadašnjeg iskustva u radu s korisnicima, primedbe su formulisali ispitivači, a na raspolaganju je bilo više odgovora (slika 4).



Slika 4: Primedbe na pretraživanje COBISS baze u COBISS/OPAC (n=100)

Nepostojanje automatske korekcije slovnih grešaka najčešća je primedba, jer su korisnici navikli da ta opcija postoji na internet pretraživačima. Sledeća primedba odnosi se na rangiranje rezultata po relevantnosti. U COBISS/OPAC postojala je opcija *Sortiranje* i podrazumevala je grupisanje rezultata po tri kriterijuma (autor, naslov i godina). Korisnici, međutim, očekuju rangiranje prema relevantnosti (na primer, ne očekuju samo sortiranje autora prema abecedi ili godini izdavanja rastuće i opadajuće, već i prema njihovoj rangiranosti na citatnim listama, što je veoma važno kada je reč o stručnim i naučnim radovima).

Ispitanici, čija je primedba iskazana sa *previše pogodaka*, uglavnom pri pretraživanju koriste ključnu reč.

Niko od ispitanika koji su, kao primedbu, istakli *premalo relevantnih pogodaka* i *pogoci ne odgovaraju upitu* ne koristi Bulove operatore, a samo 2 koristi pretraživanje preko predmetne odrednice, što ukazuje na to da korisnici nisu zainteresovani da istraže sve mogućnosti koje im program nudi. Najčešća primedba, koju je uputilo 8% ispitanika delimično zadovoljnih pretragom, odnosi se na nedostatak automatske korekcije slovnih grešaka (4), na nedostatak

rangiranja rezultata po relevantnosti (3), na previše dobijenih pogodaka (3) i na to da pogoci ne odgovaraju upitu (1).

Poslednje pitanje odnosilo se na procenu težine procesa pretraživanja u COBISS/OPAC. Većina ispitanika (69%) smatra da je pretraživanje jednostavno, 31% smatra da je primereno ličnim mogućnostima, a niko ne smatra da je COBISS/OPAC komplikovan za pretragu.

## 5 Diskusija

Analizom rezultata ankete dolazimo do zaključka da su odgovori istog ispitanika često u koliziji, što je rezultat nedovoljnog znanja o mogućnostima pretrage, neupuštanja u detaljnije upoznavanje baze, oslanjanja na pomoć bibliotekara, kao i prilagođenosti baze stručnim licima (bibliotekarima). Mogućnosti koje nudi COBISS/OPAC nisu u dovoljno iskorišćene, s jedne strane zbog složenosti kataloga, a s druge strane zbog nezainteresovanosti korisnika za istraživanje.

Bez obzira na to što su se saveti za pretraživanje u COBISS/OPAC verziji nalazili na istom ekranu kao i prozor za pretraživanje, nije ih primetila čak trećina korisnika, što ukazuje na njihovu nezainteresovanost za unapređivanje sopstvenih veština u procesu pretraživanja. S druge strane, nazivi pretraživanja nisu bili dovoljno grafički istaknuti, tako da se nije mogao dobiti precizan odgovor na pitanje koje se pretraživanje najčešće koristi. Istraživači pretpostavljaju da je, kao najpreglednije i uglavnom preporučivano tokom obuke, najčešće korišćeno *Izborna pretraživanje*.

Rezultati ankete su potvrdili prvu hipotezu da se COBISS/OPAC najčešće koristi za pretraživanje po parametrima koji su poznati korisniku (autor, naslov). To je za prosečnog korisnika efikasan način pretraživanja jer ne zahteva previše udubljanja u tehnike pretraživanja, niti iziskuje mnogo vremena.

Što se tiče olakšica pri pretraživanju, u praksi se pokazalo da se od Bulovih operatora koristi samo *AND* koji je podrazumevan i primenjuje se automatski, bez intervencije korisnika, a to praktično znači da se Bulovi operatori ne koriste. U radu s korisnicima primećeno je da se zvezdica takođe najčešće pogrešno koristi jer se, umesto na osnovu ili koren, stavlja na kraj reči, čime se gubi njen smisao, ako je upotrebljena u polju za predmetnu odrednicu ili ključnu reč.

Ovim je potvrđena početna hipoteza da korisnici ne koriste ni najjednostavnije olakšice za pretraživanje (Bulovi operatori, skraćivanje).

Odgovori na grupu pitanja, koja se odnose na pretraživanje teme ili oblasti (bez unapred poznatog autora ili naslova), pokazali su da se ključna reč koristi skoro dvostruko više od predmetne odrednice. Ključna reč je jedna od ponuđenih pristupnih tačaka na početnoj maski u izbornom pretraživanju, a predmetna odrednica se bira iz padajućeg menija. Sam koncept ključne reči bliži je korisniku od pojma predmetne odrednice. Preko ključnih reči može se pretražiti sadržaj skoro svih polja u bibliografskom zapisu (naslov, napomene, sažetak – koji se vrlo retko unosi, predmetne odrednice, slobodno oblikovane predmetne odrednice), čime se dobija veliki broj pogodaka, među kojima je malo relevantnih (veliki odziv, ali mala preciznost) (Ristović i Bacotić, 2018, str. 76).

Kao pristupna tačka, predmetna odrednica se koristi najčešće u situaciji kada se pretražuje nepoznata jedinica građe, a pri tom je jedini poznati parametar tema. Sam pojam predmetne odrednice suviše je apstraktan za korisnike. »Način na koji katalogizator formuliše temu podrazumeva upotrebu određenih alata (tezaursi, normativne baze, referensna literatura, katalog nacionalne biblioteke i dr.), ali se ipak dešava da usvojeni termin nije u skladu sa načinom razmišljanja prosečnog korisnika.« (Ristović i Bacotić, 2018, str. 68). Korisnik formuliše upit prirodnim jezikom, dok je predmetna odrednica formulisana kontrolisanim jezikom. Slobodno oblikovana predmetna odrednica najbliža je načinu razmišljanja korisnika, mada se retko koristi u COBISS.SR. (Prema istraživanju, koje je urađeno 2018. godine, od ukupnog broja zapisa u COBIB.SR (3.223.643), 114.486 zapisa ima popunjeno polje 610 za slobodno oblikovanu predmetnu odrednicu (3,5 %)) (Ristović i Bacotić, 2018, str. 73).

Rezultati ovog segmenta istraživanja potvrdili su hipotezu da je pretraživanje po predmetu suviše složeno za prosečnog korisnika.

Rezultati ankete potvrdili su tri hipoteze:

1. COBISS/OPAC se najčešće koristi za pretraživanje po parametrima poznatim korisniku (autor, naslov i sl.);
2. korisnici ne koriste sve mogućnosti pretraživanja (Bulovi operatori, skraćivanje);
3. pretraživanje po temi suviše je složeno za korisnika.

Nije potvrđena poslednja hipoteza:

4. korisnici nisu zadovoljni rezultatima pretrage COBISS/OPAC.

Razlog zbog kojeg većina ispitanika smatra da je pretraživanje lako i pri tom su zadovoljni rezultatima pretraživanja, leži upravo u činjenici da se najčešće pretražuju unapred poznati izvori. Za ovakav način pretraživanja dovoljno je samo izbeći greške u pisanju, tako da je razumljivo što je najčešća primedba na pretraživanje nedostatak automatske korekcije slovnih grešaka. Korisnici nemaju strpljenja ni vremena za upoznavanje sa svim opcijama koje nude bibliotečki pretraživači, tako da od raspoloživih opcija za pretraživanje koriste samo bibliotekari, za koje se ispostavilo da i pored napredne tehnologije predstavljaju neizostavnu kariku u pronalaženju relevantnih informacija.

## 6 Završno razmatranje

Unapređivanje bibliotečkih kataloga predstavlja temu koja je često zastupljena u stručnoj literaturi. Smatra se da većina onlajn bibliotečkih kataloga ne koristi sve mogućnosti Web okruženja koje bi katalogi *treće generacije* trebalo da ponude korisnicima (Golub, 2003, str. 3). Prema istraživanjima, idealan pretraživač trebalo bi da ima sledeće karakteristike: jednostavnu navigaciju sa grafički jasno istaknutim smernicama; proveru upita formulacijom *Jeste li mislili...*, postojanje ključne reči i predmetne odrednice na početnoj strani za pretraživanje; isticanje pojmova, koje je korisnik uneo u pretraživač, pri ispisu rezultata; nizanje rezultata prema relevantnosti; predlaganje sličnih naslova prema predmetu, sadržaju i čitanosti; korisničko označavanje (evaluacija publikacije od drugih korisnika); ispis bibliografskih jedinica obogaćenih grafičkim, tekstualnim sadržajem (slikom omota, sažetkom ili sadržajem iz publikacije); jednostavan pristup drugim bazama podataka sa kojima je biblioteka povezana itd. (Wynne and Hanscom, 2011, str. 180).

Korisnik očekuje da dođe do informacije brzo i jednostavno, sa što manje *klikova*, a to govori u prilog tome da bibliotečki katalogi ne mogu biti izolovani, već treba da budu povezani sa što više relevantnih baza podataka. U odnosu na druge izvore podataka na internetu, prednost biblioteke leži u činjenici da ona predstavlja portal koji vodi do proverene i pouzdane informacije i koji se mora nametnuti kao nezaobilazni put u potrazi za znanjem.

## Reference

Davidovski, B. i Marinčić, B., 2019. Primena novih tehnologija u radu sa decom i mladima u Narodnoj biblioteci Užice. U: Čeliković, K. ur. *Dani hrvatske knjige i riječi: zbornik radova 2018*. Subotica: Hrvatska čitaonica i Zavod za kulturu vojvođanskih Hrvata. Str. 133–150.

Golub, K., 2003. *Predmetno pretraživanje u knjižničnim katalogima s web-sučeljem*. Magistarski rad. Zagreb: Sveučilište u Zagrebu, Filozofski fakultet.

Svenonius, E., 2007. *Intelektualna osnova organizovanja informacija*. Beograd: Narodna biblioteka Srbije i Clio.

Wynne, S. C. i Hanscom, M. J., 2011. The effect of next-generation catalogs on catalogers and cataloging functions in academic libraries. *Cataloging & Classification Quarterly*, 49(3), str. 179–207.

Ristović, B. i Bacotić, G., 2018. Slobodno oblikovana predmetna odrednica i potrebe korisnika. *Bibliotekar*, 60(1), str. 67–79.

Ršumović, I. ur., 2006. *Narodna biblioteka Užice: 1856-2006: 150 godina postojanja*. Užice: Narodna biblioteka.

Šokica, N., 2005. Pretraživanje uzajamnog kataloga - COBISS/OPAC. *Infoteka*, 6, str. 33–41.

# Ažuriranje zapisov v VIAF in priporočila za kakovostnejši prikaz slovenskih normativnih zapisov v VIAF

## Updating records in VIAF and recommendations for higher quality display of Slovenian authority records in VIAF

**Andreja Krajnc Vobovnik<sup>1</sup>**

**IZVLEČEK:** V prispevku so opisani ažuriranje slovenskih normativnih zapisov v VIAF, vpis identifikatorjev v bazo podatkov CONOR.SI in njihov prikaz v COBISS+. Pri večini normativnih zapisov je ažuriranje v VIAF potekalo odlično, le pri nekaj zapisih za osebna imena so se pri združevanju v skupke pojavile napake, ki smo jih uvrstili v štiri skupine in jih primerjali s tipičnimi problemi pri vpisu, opisanimi v dokumentu *VIAF Guidelines*. Na osnovi vpisa zapisov v VIAF in analize napačnih povezav smo oblikovali splošna priporočila za kreiranje normativnih zapisov za osebna imena, da bi bilo pri vpisu in ažuriranju zapisov v VIAF v prihodnosti čim manj napak, ki nastanejo v normativni bazi podatkov CONOR.SI pri katalogizaciji. Opisana so tudi priporočila za ravnanje z normativnimi zapisi, ki so že bili poslani v VIAF. Namen prispevka je tudi ozaveščanje katalogizatorjev o pomenu normativne kontrole na globalnem nivoju in poudarjanje, da njihov intelektualni prispevek pri tem ni zanemarljiv.

**KLJUČNE BESEDE:** normativni zapisi, osebna imena, imena korporacij, CONOR.SI, VIAF, ISNI, Slovenija

**ABSTRACT:** The article describes the process of updating Slovenian authority records in VIAF, the entry of identifiers into the CONOR.SI database and their display in COBISS+. For most authority records, the updating in VIAF went smoothly, whereas in some records for personal names errors occurred during the process of putting them into clusters, which were then sorted into four groups and compared with the typical entry-related issues described in the *VIAF Guidelines*. Based on the entry of records in VIAF and the analysis of erroneous links, we compiled general recommendations for creating authority records for personal names in order to avoid errors in the future, which occur in the CONOR.SI authority database during cataloguing. The aim of this article is also to educate cataloguers on the importance of authority control on a global level and to emphasise that their intellectual contribution is not negligible.

**KEYWORDS:** authority records, personal names, corporate body names, CONOR.SI authority database, VIAF, ISNI, Slovenia

---

## 1 Uvod

Eden izmed glavnih konceptov normativne datoteke VIAF je ohranitev nacionalnih ali lokalnih oblik imena v normativnih zapisih, ki so kreirane na osnovi različnih katalogizacijskih pravil. Uporabnik namreč želi prikaz izvornih verzij normativnih zapisov v različnih jezikih, pisavah in črkovanjih. Ista oseba iz različnih virov se v VIAF identificira s primerjavo normativnih točk dostopa ter z variantnimi in sorodnimi točkami dostopa. Ko se imena ujemajo, se primerjajo še dodatni podatki, ki potrdijo ujemanje. Pri osebah so ti podatki letnica rojstva, letnica smrti,

---

<sup>1</sup> Andreja Krajnc Vobovnik, strokovna sodelavka, Institut informacijskih znanosti (IZUM), Maribor, Slovenija, [andreja.krajncvobovnik@izum.si](mailto:andreja.krajncvobovnik@izum.si).



obdobje delovanja in podatki iz bibliografskih zapisov, predvsem naslovi del (Bennett et al., 2006).

V VIAF se soočajo s težavami pri povezovanju določenih vrst imen v skupke (angl. *clusters*) zaradi velike raznolikosti oblik imen v različnih jezikih, zato spodbujajo članice VIAF, naj jim sporočijo napake, ki jih nato ročno popravijo. Pri združevanju normativnih zapisov se pogosto srečujejo z naslednjimi tipičnimi problemi (VIAF Guidelines, 2019):

- *Pomešana imena (pomešani homonimi)*: Pojavijo se, ko se naslovi del, ki bi morali biti povezani z različnimi ločenimi entitetami, vsi povežejo samo z eno od teh entitet.
- *Manjkajoči naslovi del in datumi*: VIAF potrebuje za diferenciacijo med entitetami dovolj podatkov; če so osebna imena brez datumov ali naslovov del, je primerjanje oteženo.
- *Neobičajno kodirani podatki*: Ustanova pri katalogizaciji uporablja netipična lokalna katalogizacijska pravila.
- *Razlikovanje z uporabo jezikovno odvisnih podatkov*: Pri osebnih imenih je kvalifikator v podpolju 200c – *Dodatki k imenu (ne datumi)* (v formatu COMARC/A) oblikovan v jeziku katalogizacije. Pri večini primerjav se ta podatek zaradi jezikovnih različic običajno ne upošteva.
- *Podvojeni zapisi*: V VIAF se trudijo, da bi vključevali en zapis iste članice VIAF v isti skupek. VIAF občasno objavi sezname potencialnih podvojenih zapisov, ki so uporabni pri razreševanju tako za VIAF kot za sodelujočo ustanovo. Podvojeni zapisi pogosto povzročijo, da se skupek kasneje razcepi.
- *Klasična imena, imena kraljev, kraljic in papežev*: Zaradi velike raznolikosti oblik imen v različnih jezikih je združevanje zelo zahtevno. V VIAF probleme, na katere opozorijo članice VIAF, rešujejo ročno.

V prispevku probleme podrobneje opišemo in jih ponazorimo s primeri iz baze podatkov CONOR.SI. Opisani problemi se nanašajo samo na osebna imena, ker smo v VIAF poslali premalo normativnih zapisov za imena korporacij, da bi lahko podrobneje analizirali napake pri združevanju. S sodelavci VIAF (VIAF Team) smo odlično sodelovali; napake, povezane z napačnim združevanjem v skupke, so hitro odpravili. Nekaterih izmed predstavljenih primerov zato ne moremo več najti v VIAF.

V nadaljevanju prispevka želimo katalogizatorjem predstaviti, na kaj naj bodo pozorni pri določanju normativnih točk dostopa in urejanju normativnih zapisov, da bi zagotovili boljšo kakovost tako baze podatkov CONOR.SI kot tudi slovenskih normativnih zapisov v VIAF. V prihodnje bi radi odpravili razloge, zaradi katerih pri združevanju slovenskih normativnih zapisov v skupke nastajajo tipične napake, ki smo jih prepoznali pri preverjanju vključenih zapisov in se ujemajo z nekaterimi tipičnimi problemi, opisanimi v dokumentu *VIAF Guidelines*.

## 2 Ažuriranje zapisov iz baze podatkov CONOR.SI v VIAF

Skoraj leto dni po inicialnem pošiljanju smo novembra 2019 za VIAF prvič pripravili podatke za ažuriranje. Poleg tega smo tudi prvič poslali normativne zapise za imena korporacij. V izboru je bilo 64.869 normativnih zapisov, od tega jih je bilo 3.523 novih (3.449 za osebna imena in 74 za imena korporacij). V VIAF je bilo tokrat vpisanih 5,05 % vseh normativnih zapisov iz baze podatkov CONOR.SI, s katerimi je povezanih 2.081.065 pripadajočih bibliografskih zapisov, tj. 38,5 % bibliografskih zapisov iz baze podatkov COBIB.SI. Delež obeh vrst zapisov se je pri prvem

ažuriranju glede na inicialni vpis le za malenkost povečal. Pri normativnih zapisih se je povečal za 0,05 %, pri bibliografskih pa za 0,5 %.

Za ažuriranje zapisov v VIAF pošiljamo trenutno stanje ustreznih normativnih zapisov in pripadajočih bibliografskih zapisov. Izbor zapisov za osebna imena je ostal enak kot pri inicialnem vpisu (Krajnc Vobovnik, 2018). Izbor zapisov za imena korporacij je podoben izboru za osebna imena, razlikuje se v nekaj podrobnostih, ki so vezane na značilnosti normativnih zapisov za imena korporacij. Po pravilu v COMARC/A je normativni zapis dokončno popravljen, ko je v podpolju *001a – Status zapisa* vpisana koda »c« – *popravljen zapis* in je v podpolju *100b – Koda za status normativne točke dostopa* koda »a« – *sprejeta*. Za korporacije smo izbor seveda omejili na zapise s kodo »b« – *ime korporacije* v podpolju *001c – Vrsta entitete*. Normativne zapise smo iz formata COMARC/A konvertirali v format MARC 21/A in tako kot pri osebnih imenih izločili podatke, ki niso za javno rabo (*830 – Splošna katalogizatorjeva opomba*), in podatke v bloku *9XX – Blok za nacionalno rabo*, ki niso namenjeni mednarodni izmenjavi. Nato smo za vsak normativni zapis za ime korporacije, ki ustreza zgornjim pogojem, v bazi podatkov COBIB.SI poiskali pripadajoče bibliografske zapise (po številki normativnega zapisa v podpoljih *71X3*). Pri tem smo izločili normativne zapise, ki v bazi podatkov COBIB.SI nimajo nobenega bibliografskega zapisa, zapise za dogodke (izvedena dela), zapise, označene za brisanje, prve vnose zapisa, CIP-zapise in zapise za raritete. V izbor za pošiljanje je prišlo samo 74 zapisov za ime korporacije, kar predstavlja komaj 0,5 % vseh normativnih zapisov za korporacije, čeprav smo normativno kontrolo korporacij vzpostavili prej, jeseni 2018.

Pred ažuriranjem zapisov, ki so že bili poslani v VIAF, smo v bazi podatkov CONOR.SI preverili, ali še imajo vpisane kodirane podatke, na katerih temelji izbor normativnih zapisov. Ugotovili smo, da se je pri 19 normativnih zapisih pojavila sprememba, zaradi katere se ne bi več uvrstili v izbor. Spremenjeni so bili naslednji podatki: Pri treh zapisih je koda za status normativne točke dostopa (podpolje *100b*) spremenjena iz *sprejeta* (koda »a«) v *začasna* (koda »c«). Iz zapisov nismo mogli ugotoviti vzroka za to spremembo. Pri 16 zapisih pa je koda za status zapisa (podpolje *001a*) spremenjena iz *popravljen zapis* (koda »c«) nazaj v *nov zapis* (koda »n«).

Omenjenih zapisov pred pošiljanjem nismo ponovno preverjali ali popravljali. Smo pa vse katalogizatorje, ki so v 16 zapisih spremenili status zapisa, opozorili na pomen kod, ki vplivajo na izbor zapisov. Ugotovili smo, da so zapisi, ki so že bili vpisani v VIAF in so pri ažuriranju zaradi spremembe kodiranih podatkov izpadli iz novega izbora, brisani tudi iz VIAF. To želimo pri programskem vpisu v VIAF v prihodnje preprečiti oz. zmanjšati na najmanjšo možno raven, kar je eden izmed namenov tega pisanja. Pred pošiljanjem bomo take zapise preverjali in katalogizatorje zaprosili za dodatno preverjanje normativnih zapisov in za morebitne popravke.

### 3 Vpis identifikatorjev v bazo podatkov CONOR.SI in prikaz v COBISS+

Pri preverjanju stanja zapisov, ki so na strani VIAF Dataset (<http://viaf.org/viaf/data/>) mesečno objavljeni za celoten VIAF, smo ugotovili, da so bili tako kot pri inicialnem vpisu tudi pri prvem ažuriranju v VIAF vpisani vsi poslani normativni zapisi. Pri tem so zapisi pridobili identifikator VIAF ID, polovica od teh zapisov pa tudi ISNI (angl. *International Standard Name Identifier*). Oba smo vpisali v bazo podatkov CONOR.SI: v polje *035 – Kontrolne številke zapisov iz drugih sistemov* smo vpisali VIAF ID, v polje *010 – Mednarodni standardni identifikator imen (ISNI)* pa

ISNI. Datum vpisa se programsko vpiše kot zadnji podatek v sistemsko generiranem polju (*Conversion*) (slika 1).

<b>ID=403811 LN=0000001577 PN V1 20.12.2001 CONOR::IZUM Updated: 27.03.2020</b>	
<b>CONOR::IZUM_ANDREJA Copied: 0 Adopted by: NUK Conversion 18.05.2019</b>	
<b>001</b>	<b>ac - popravljen zapis bx - normativni zapis ca - osebno ime</b>
<b>010</b>	<b>a0000000108646146</b>
<b>035</b>	<b>a(VIAF)4968133</b>
<b>100</b>	<b>ba - sprejeta cslv - slovenski gba - latinica</b>
<b>101</b>	<b>aslv - slovenski</b>
<b>102</b>	<b>asvn - Slovenija</b>
<b>106</b>	<b>a1 - ne uporablja se kot predmetna oznaka</b>
<b>120</b>	<b>ab - moški ba - razlikovano osebno ime</b>
<b>152</b>	<b>aPPIAK - Pravilnik i priručnik za izradbu abecednih kataloga</b>
<b>19011</b>	<b>a1948 b04 c13</b>
<b>200</b>	<b>1 aJančar bDrago</b>

Slika 1: Identifikatorja VIAF ID in ISNI v normativnem zapisu za Draga Jančarja (Vir: CONOR.SI, 2020)

Na osnovi vpisa v normativne zapise smo v javno dostopni bazi podatkov CONOR.SI v COBISS+ omogočili prikaz obeh identifikatorjev in vzpostavili povezavo do servisov VIAF in ISNI (slika 2).

Pri preverjanju vključenih zapisov smo ugotovili, da je identifikator ISNI pri inicialnem vpisu v VIAF pridobilo 48,4 % poslanih normativnih zapisov za osebna imena (29.712 zapisov), po prvem ažuriranju pa že 50,3 % poslanih zapisov (32.639 zapisov), od tega samo 6 zapisov za ime korporacije (8 % poslanih zapisov za ime korporacije), med njimi za dve slovenski ustanovi, in sicer za Nacionalno in univerzitetno knjižnico Ljubljana (ISNI [0000 0001 2286 3964](#)) in Mestni muzej Ljubljana (ISNI [0000 0001 1015 6429](#)). Visok delež pridobljenih ISNI za osebna imena je spodbuden, saj je ena izmed poglobitvenih prednosti vpisa slovenskih normativnih zapisov v VIAF tudi brezplačna pridobitev ISNI (Krajnc Vobovnik, 2018).

Uporabniški	GARE
Osebno ime	Jančar, Drago
Variante osebnega imena	Jancár, Drago (drugo); Jancar, Drago (drugo); D. J.; Jančar, Drago; Giantsar, Ntragko
Država	Slovenija
Jezik, ki ga upor. oseba/korp.	slovenski
Opombe	Pisatelj, esejist, dramatik in publicist. Dobil je več slovenskih in mednarodnih nagrad, med drugim leta 1993 Prešernovo nagrado za življenjsko delo. Od leta 2001 je redni član Slovenske akademije znanosti in umetnosti.
ISNI	0000 0001 0864 6146
VIAF ID	4968133
CONOR.SI-ID	403811

Slika 2: Prikaz identifikatorjev VIAF ID in ISNI v normativnem zapisu za Draga Jančarja (Vir: COBISS+, 2020)

## 4 Problemi pri vpisu normativnih zapisov za osebna imena v VIAF

Za vse normativne zapise, ki so vpisani v VIAF, smo iz datoteke VIAF Dataset pripravili seznam VIAF ID-jev, ki pripadajo posameznim CONOR.SI-ID-jem.

Na tem seznamu smo preverili, ali obstajajo tudi podvojeni VIAF ID-ji. Analiza je pokazala, da je imelo podvojen VIAF ID samo 11 slovenskih normativnih zapisov. Pri podrobnejšem preverjanju smo ugotovili, da jih lahko razvrstimo v štiri skupine:

1. podvojeni zapisi v bazi podatkov CONOR.SI (4 zapisi),
2. napačne povezave z normativnimi zapisi v bibliografskih zapisih v bazi podatkov COBIB.SI (2 zapisa),
3. razlike v katalogizacijskih pravilih (1 zapis),
4. napačno združeni zapisi v VIAF (4 zapisi).

Normativni zapisi za isto entiteto, ki so jih kreirale različne članice VIAF, se v VIAF združujejo v skupke, kjer je vsaka članica predstavljena z lastnim simbolom, običajno z nacionalno zastavo. Zato se dva zapisa z istim VIAF ID-jem, ki ju je prispevala ista članica, pojavita v istem skupku in ju lahko pri pregledovanju hitro opazimo že na osnovi podvojenega simbola (slika 3).

### 4.1 Podvojeni zapisi v bazi podatkov CONOR.SI

Ena izmed pridobitev vpisa normativnih zapisov v VIAF je tudi seznam potencialnih podvojenih zapisov v bazi podatkov CONOR.SI, ki smo ga dobili iz datoteke VIAF Dataset. Pred pošiljanjem nismo pričakovali, da bomo v VIAF poslali podvojene zapise, saj smo v izbor zajeli samo normativne zapise s sprejeto točko dostopa (*100b = »a« – sprejeta normativna točka dostopa*) in dokončno popravljene zapise, tj. zapise brez slovničnih in vsebinskih napak (*001a = »c« – popravljen zapis*), med katerimi ne bi smelo biti podvojenih zapisov. Pri vseh štirih podvojenih zapisih smo ugotovili, da gre za tuje avtorje. Pri treh so avtorji z dvema osebnima imenoma, kjer je v enem normativnem zapisu drugo osebno ime vpisano v izpisani obliki, v drugem pa samo z začetnico. Pri enem podvojenem zapisu pa gre za različen vpis nemške črke.

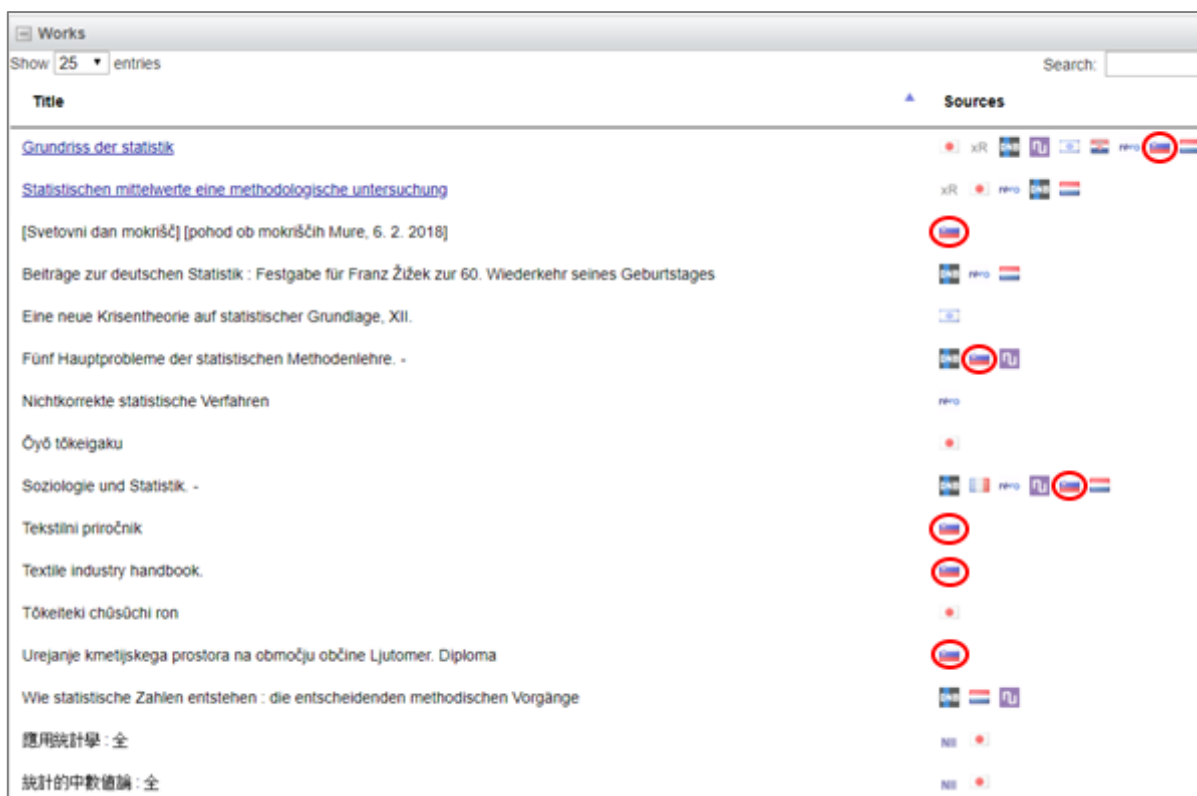
Primer avtorja z dvema osebnima imenoma je ameriški matematik, fizik, astronom in filozof Charles Sanders Peirce. V bazi podatkov CONOR.SI sta bila dva normativna zapisa: Peirce, Charles Sanders, 1839–1914 (CONOR.SI-ID 52600675), in Peirce, Charles S., 1839–1914 (CONOR.SI-ID 196990819). Oba zapisa sta v VIAF (slika 3) pravilno združena v isti skupek (<http://viaf.org/viaf/89203252>). V bazi podatkov COBIB.SI je bila številka prvega normativnega zapisa vpisana v 15 bibliografskih zapisih, številka drugega pa v 7 bibliografskih zapisih. Avtorjevo ime je v različnih bibliografskih zapisih navedeno v obeh oblikah. Pri preverjanju oblike imena v polju 200 – *Naslov in navedba odgovornosti* smo ugotovili, da je izpisana oblika imena navedena v 8 bibliografskih zapisih, oblika z začetnico pa v 12 zapisih, v dveh zapisih pa je avtor naveden samo s priimkom. Pri urejanju zapisov smo obdržali zapis z izpisano obliko imena Peirce, Charles Sanders, 1839–1914 (CONOR.SI-ID 52600675). Po slovenskih katalogizacijskih pravilih damo prednost izpisani obliki imena, če avtor uporablja tako začetnice oz. krajšave imena kot izpisano obliko imena (ZNAČKA, 2001). Skrajšana oblika imena je navedena kot variantna oblika v polju 400 – *Variantna točka dostopa*.





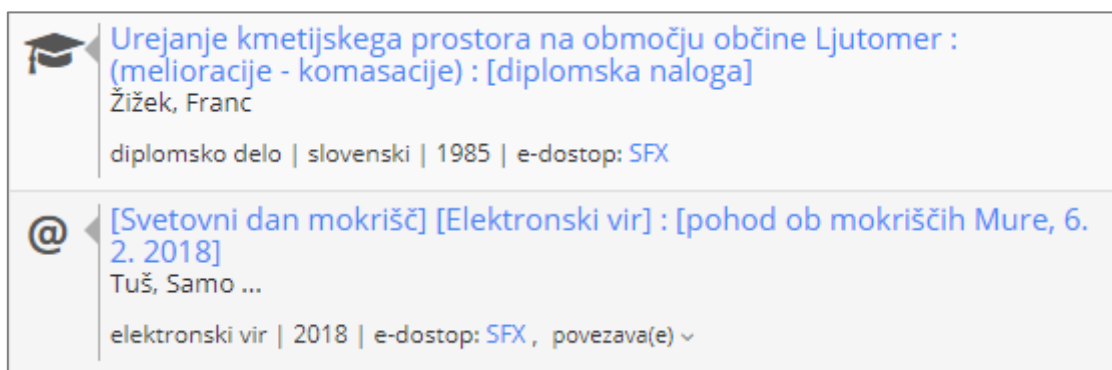
Slika 4: Skupek za normativno točko dostopa Žižek, Franc z dvema normativnima zapisoma iz baze podatkov CONOR.SI (Vir: VIAF, 2019)

Pri pregledu bibliografskih zapisov v VIAF smo ugotovili, da so vpisana dela treh različnih avtorjev – soimenjakov, kar je nastalo na osnovi napačnih povezav v bibliografskih zapisih v bazi podatkov COBIB.SI. Na sliki 5 (zavihek *Works*) so prikazani v VIAF pomešani bibliografski zapisi vseh treh avtorjev, kjer je videti, kot da gre za eno osebo, ki je delovala na več različnih področjih. V resnici gre za tri različne avtorje, ki so delovali na različnih področjih in v različnih časovnih obdobjih. Na slikah od 6 do 8 so prikazani sezname bibliografskih zapisov, povezanih s posameznimi avtorji, v bazi podatkov COBIB.SI po popravkih.



Slika 5: Združena dela dveh avtorjev zaradi vpisa napačnih normativnih točk dostopa v bibliografskih zapisih v bazi podatkov COBIB.SI (Vir: VIAF, 2019)





Slika 6: Deli, povezani z normativno točko dostopa Žižek, Franc (Vir: COBIB.SI, 2019)



Slika 7: Deli, povezani z normativno točko dostopa Žižek, Franc, 1944- (Vir: COBIB.SI, 2019)



Slika 8: Dela, povezana z normativno točko dostopa Žižek, Franc, 1876–1938 (Vir: COBIB.SI, 2019)

### 4.3 Razlike v katalogizacijskih pravilih

Zaradi različnih katalogizacijskih pravil sodelujočih ustanov sta se pri enem paru normativnih zapisov združila dva normativna zapisa za osebno ime avtorja, ki je spremenil ime. Ne gre za psevdonim, ampak za spremembo osebnega imena v hebrejsko obliko. V bazi podatkov

CONOR.SI je za prvotno osebno ime Fischl, Viktor, 1912–2006 (CONOR.SI-ID 24032099) kreiran en normativni zapis, za ime v hebrejščini Dagan, Avigdor, 1912–2006 (CONOR.SI-ID 84234339) pa drug normativni zapis. »Če tuji avtor uporablja za različna področja ustvarjanja različne psevdonime oz. pravo ime in enega ali več psevdonimov, upoštevamo ustrezno normativno zbirko ali nacionalno bibliografijo, kar pomeni, da lahko izdelamo več enotnih značnic.« (ZNAČKA, 2001, str. 13), to je normativnih točk dostopa. V takih primerih kreiramo dva ločena normativna zapisa, ki sta med sabo povezana prek polja 500 – *Sorodna točka dostopa – osebno ime*. V VIAF sta v isti (mešani) skupek združena oba naša zapisa za obe imeni – Viktor Fischl in Avigdor Dagan (<http://viaf.org/viaf/83981379>) (slika 9).

Obstaja pa tudi samostojni skupek samo za prvotno obliko imena Viktor Fischl (<http://viaf.org/viaf/93611328>) (slika 10).



Slika 9: Mešani skupek za Viktor Fischl, 1912–2006 in Avigdor Dagan, 1912–2006 (Vir: VIAF, 2020)



Slika 10: Skupek za normativno točko dostopa Fischl, Viktor, 1912–2006 (Vir: VIAF, 2020)

Iz VIAF smo dobili pojasnilo, da se je mešani skupek za ta primer oblikoval zaradi različnih interpretacij oblik osebnega imena. Nekatere članice VIAF obravnavajo eno izmed imen kot variantno obliko, kjer je osebno ime Fischl, Viktor, 1912–2006, ali osebno ime Dagan, Avigdor, 1912–2006, vpisano v polje 400 – *Variantna točka dostopa – osebno ime*, druge članice VIAF, med njimi tudi slovenske knjižnice, ki dopolnjujejo bazo podatkov CONOR.SI, pa obravnavajo obe obliki osebnega imena kot sorodni točki dostopa in ju vpišejo v polje 500 – *Sorodna točka dostopa – osebno ime*. Zaradi različne interpretacije sta se oba slovenska normativna zapisa znašla v istem (mešanem) skupku. Take primere v VIAF ročno ločujejo glede na vsebino polja 200 – *Normativna točke dostopa – osebno ime*, vendar pa ne oblikujejo treh skupkov, enega za eno obliko, drugega za drugo obliko in tretjega mešanega. Dogovorili smo se, da bodo v



VIAF ročno prenesli normativni zapis za Fischl, Viktor, 1912–2006, na skupek za Fischl, Viktor (<http://viaf.org/viaf/93611328>), a ob objavi prispevka ta popravek še ni viden v VIAF.

#### 4.4 Napačno združevanje v VIAF

Pri združevanju v skupke so se napačno združili 4 pari normativnih zapisov. Pri vseh primerih je napaka nastala pri soimenjakih oziroma pri avtorjih z zelo podobnimi osebnimi imeni. Zapis za osebno ime Dekleva, Jože, 1944- (CONOR.SI-ID 1849187), se je napačno združil v isti skupek z zapisom za osebno ime Dekleva, Jože, 1930–2018 (CONOR.SI-ID 182115) (slika 11). Prvi avtor piše o urbanizmu, drugi pa je bil športni novinar. V VIAF so bila zaradi združitve obeh identitet združena tudi njuna dela s področij urbanizma in športa. V skupku je tudi oblika imena brez letnice, ki so ga prispevale druge članice VIAF.

The screenshot shows a VIAF record for 'Dekleva, Jože, 1944-' and 'Dekleva, Jože, 1930-2018'. Below the name, there is a section for 'Works' with a search bar and a 'Sources' column. The works listed include '[Uvodnik]', 'Bele arene, 1983:', 'Diskusijski prispevek [na sekciji D: koliko prostora javnosti]', 'Družbeni značaj ekološke krize : elementi strategije humanizacije okolja', 'Dušan Kuret - devetdesetletnik', 'Evgen Bergant 1933-2004', 'Ljubljana 2000 : [dogovor o skupnih temeljih dolgoročnih planov]', 'Ljubljana dva tisoč', 'Lokacijski model za planiranje stanovanjske gradnje na nivoju občinskih prostorskih planov', and 'Methods of spatial planning at local-cmmunal level.'. In the 'Sources' column, several small icons are visible, with a red circle highlighting the top one and a larger red oval highlighting a vertical column of icons on the right side of the list.

Slika 11: Napačno oblikovan skupek za normativno točko dostopa Dekleva, Jože (Vir: VIAF, 2019)

<b>ISNI:</b>	0000 0003 8377 0109 <a href="http://isni.org/isni/0000000383770109">http://isni.org/isni/0000000383770109</a>
<b>Name:</b>	Dekleva, Jože
<b>Creation class:</b>	Language material Text
<b>Creation role:</b>	author editor
<b>Related names:</b>	Simmie, James Simmie, James (1941-) Šimmie, James Martin (1941-) Štravs, Luka
<b>Titles:</b>	Bele arene, 1983: Urejanje prostora 2 : pregled novejših raziskav Urejanje prostora na občinski ravni Yugoslavia in turmoil : after self-management?
<b>Notes:</b>	
<b>Sources:</b>	VIAF LC NSK BOWKER NTA

Slika 12: Napačno oblikovan zapis za identiteto Jože Dekleva (Vir: ISNI, 2020)

Na osnovi napačne združitve v VIAF je dvema identitetama napačno dodeljen en ISNI 0000 0003 8377 0109 v bazi podatkov ISNI. V bazo podatkov ISNI, ki je osnovana na normativnih podatkih iz VIAF, se je prenesla napaka iz VIAF in se dodala še nova, saj v zapis ni prenesena letnica rojstva. Tudi v ISNI sta se združili obe identiteti, tako da so dela obeh avtorjev pomešana (slika 12).

Na naš predlog so v VIAF napako odpravili. Nepravilno oblikovani skupek so razdružili in oblikovali dva ločena skupka (<http://viaf.org/viaf/271522732>, <http://viaf.org/viaf/43236294>) (slika 13). Spremembi v VIAF pa ni sledila sprememba v ISNI, kjer se zapis do objave tega prispevka še ni spremenil.

Heading	Type	Sample Title
1 <a href="#">Dekleva, Jože, 1944-</a> xA <a href="#">Dekleva, Jože</a>	Personal	Diskusijski prispevek [na sekciji B: urejanje prostora med naročnikom in izvajal ... Yugoslavia in turmoil : after self-management?
2 <a href="#">Dekleva, Jože, 1930-2018</a> xA <a href="#">Dekleva, Jože</a> <a href="#">Jože Dekleva</a>	Personal	Bele arene, 1983: Bele arene, 1983:

Slika 13: Razdružena skupka za soimenjaka z imenom Jože Dekleva (Vir: VIAF, 2020)

## 5 Podvojeni ISNI

V bazo podatkov ISNI prispeva podatke več kot 50 ustanov z različnih področij (VIAF, nacionalne knjižnice, repozitoriji, ustanove, ki popisujejo vire s področja glasbe, in ustanove za upravljanje z avtorski pravicami); ti podatki se povezujejo z uporabo različnih algoritmov za ujemanje zapisov (How ISNI works, 2020). Težave pri povezovanju zapisov v bazi podatkov ISNI smo zasledili pri analizi vpisanih zapisov, ko smo ugotovili, da sta 107 normativnim zapisom dodeljena dva identifikatorja ISNI, nekaterim pa celo trije. Delež zapisov, pri katerih so se pojavile težave, je sicer samo 0,36 % vseh ISNI, ki smo jih pridobili s pošiljanjem zapisov v VIAF, a je precej velik za dodatno ročno preverjanje in usklajevanje podatkov v CONOR, VIAF

in ISNI. Seznam podvojenih identifikatorjev ISNI, za katere smo menili, da so verjetno napačno dodeljeni, smo poslali v VIAF v dodatno preverjanje in usklajevanje z bazo podatkov ISNI.

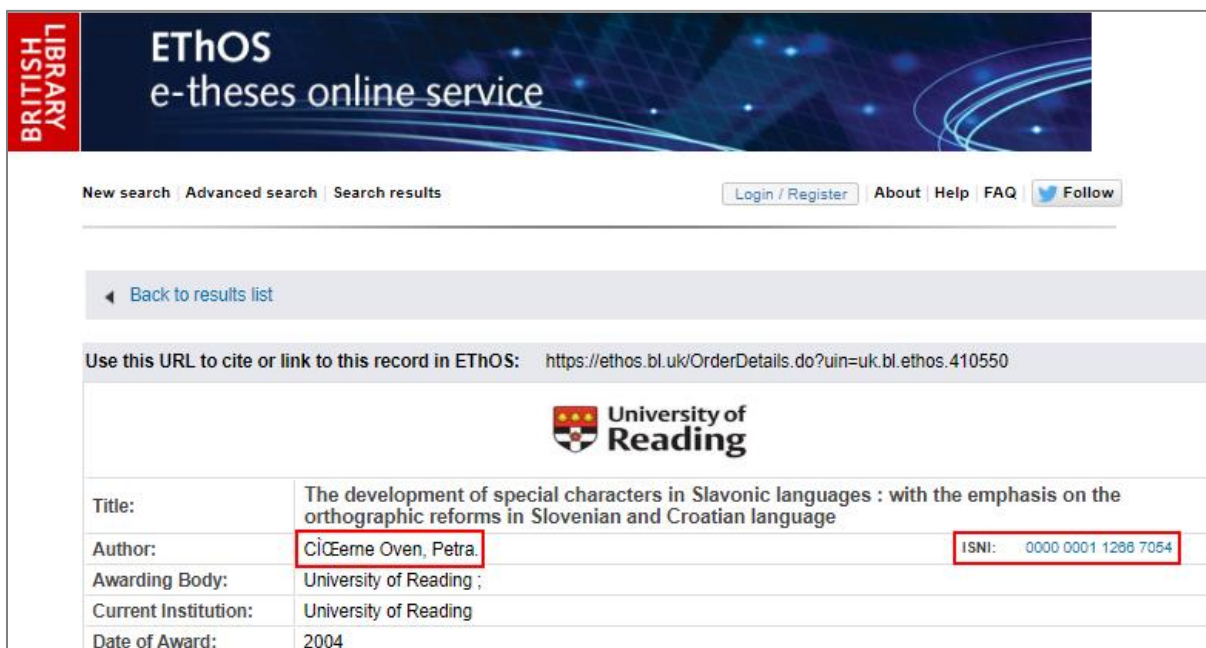
Ob uvedbi novega polja 010 – *Mednarodni standardni identifikator imen (ISNI)* v format COMARC/A smo se odločili, da bo polje neponovljivo. Takrat še nismo imeli vpogleda, koliko različnih identifikatorjev ISNI bo v resnici dodeljeno isti entiteti. Odločili smo se, da bomo pri več dodeljenih ISNI za isto entiteto ročno preverjali, kateri identifikator bomo vpisali v normativni zapis v bazo podatkov CONOR.SI. ISNI se mora nanašati na normativno točko dostopa, ki je navedena v polju 200 – *Normativna točka dostopa – osebno ime*. Pri podrobnejšem preverjanju zapisov v bazi podatkov ISNI smo ugotovili, da se večina nanaša na isto identiteto. Večina takih primerov se nanaša na tuje avtorje, nekateri primeri pa na slovenske. Med njimi je tudi dr. Petra Černe Oven, oblikovalka in predavateljica na Akademiji za likovno umetnost in oblikovanje, ki so ji dodeljeni kar trije različni identifikatorji: ISNI 0000 0003 8500 0589, ISNI 0000 0003 8726 940X in ISNI 0000 0001 1266 7054 (slika 14).

<p><b>ISNI:</b> 0000 0003 8500 0589  <a href="http://isni.org/isni/0000000385000589">http://isni.org/isni/0000000385000589</a></p> <p><b>Name:</b> Černe Oven, Petra            Oven, Petra Cerne</p> <p><b>Creation class:</b> Language material            Text</p> <p><b>Creation role:</b> author            creator</p> <p><b>Related names:</b> Arhitekturni muzej v Ljubljani            Kisin, Peter            Predan, Barbara            Šuštaršič, Maja</p> <p><b>Titles:</b> BIO 22            Dvaindvajseti bienale industrijskega oblikovanja            Edward Wright, c2007:            Twenty-second biennial of industrial design</p> <p><b>Notes:</b></p> <p><b>Sources:</b> VIAF LC            BOWKER            NLN</p>	<p>Černe Oven, Petra. </p> <p>Oven, Petra Cerne </p> <p>VIAF ID: 279276747 (Personal)            Permalink: <a href="http://viaf.org/viaf/279276747">http://viaf.org/viaf/279276747</a></p> <p>Preferred Forms</p> <p>4xx's: Alternate Name Forms (7)</p> <p>Works</p> <p>Selected Co-authors</p> <p>Countries and Regions of Publication (10)</p> <p>Publication Statistics</p> <p>Selected Publishers (14)</p> <p>About</p> <p>Personal Information            Nationality or associated country:            svn - </p> <p>Language:            slv - Slovenian </p> <p>External Links  <a href="#">ISNI - 0000 0003 8500 0589</a>  <a href="#">ISNI - 0000 0003 8726 940X</a>  <a href="#">WorldCat Identities - lccn-nb2010018321</a></p> <p>Record Views</p> <p>History of VIAF ID:279276747 (9)</p>
<p><b>ISNI:</b> 0000 0003 8726 940X  <a href="http://isni.org/isni/000000038726940X">http://isni.org/isni/000000038726940X</a></p> <p><b>Name:</b> Černe Oven, Petra            Oven, Petra Cerne</p> <p><b>Creation class:</b> cre            Language material</p> <p><b>Creation role:</b> contributor</p> <p><b>Related names:</b> Vodaine, Jean (is Vladimir Kaučič, 1921-2006.)</p> <p><b>Titles:</b> Jean Vodaine</p> <p><b>Notes:</b></p> <p><b>Sources:</b> NTA            TEL</p>	
<p><b>ISNI:</b> 0000 0001 1266 7054  <a href="http://isni.org/isni/0000000112667054">http://isni.org/isni/0000000112667054</a></p> <p><b>Name:</b> Černe Oven, Petra</p> <p><b>Creation class:</b> Text</p> <p><b>Creation role:</b> author</p> <p><b>Related names:</b> University of Reading</p> <p><b>Titles:</b> development of special characters in Slavonic languages, The ; with the emphasis on the orthographic reforms in Slovenian and Croatian language</p> <p><b>Notes:</b> Thesis (Ph.D.)--University of Reading, 2004  <a href="http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.410550">http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.410550</a></p> <p><b>Sources:</b> ETHOS            OCLCT</p>	

Slika 14: Trije zapisi za Petro Černe Oven v ISNI (Vir: ISNI, 2020) (levo) in skupek za normativno točko dostopa Černe Oven, Petra z dvema identifikatorjema ISNI (Vir: VIAF, 2020) (desno)

V datoteki VIAF Dataset sta dva identifikatorja (ISNI 0000 0003 8726 940X in ISNI 0000 0003 8500 0589), ki sta v VIAF vpisana v isti skupek (slika 14) (<http://viaf.org/viaf/279276747>). Prvi ISNI za osebno ime brez šumnikov (Cerne Oven, Petra) izvira iz VIAF, drugi ISNI za osebno ime s šumniki pa iz nizozemske nacionalne knjižnice (koda NTA).

Tretji ISNI (0000 0001 1266 7054) je dodeljen za osebno ime s šumniki, ki izvira iz britanskega nacionalnega repozitorija doktorskih disertacij EThOS (slika 15).



The screenshot shows the EThOS e-theses online service interface. At the top, there is a navigation bar with links for 'New search', 'Advanced search', and 'Search results'. Below this, there is a search bar and a 'Back to results list' link. The main content area displays the title of the thesis: 'The development of special characters in Slavonic languages : with the emphasis on the orthographic reforms in Slovenian and Croatian language'. The author is listed as 'ČiČerme Oven, Petra', and the ISNI number is '0000 0001 1266 7054'. The awarding body and current institution are both listed as 'University of Reading', and the date of award is '2004'.

Title:	The development of special characters in Slavonic languages : with the emphasis on the orthographic reforms in Slovenian and Croatian language	
Author:	ČiČerme Oven, Petra	ISNI: 0000 0001 1266 7054
Awarding Body:	University of Reading ;	
Current Institution:	University of Reading	
Date of Award:	2004	

Slika 15: Zapis za Petro Černe Oven (Vir: EThOS, 2020)

Po standardu ISO 27729 se identifikator ISNI dodeli samo za eno javno identiteto osebe ali organizacije. Če ima ena oseba ali organizacija (entiteta) več kot eno identiteto, je vsaka identificirana s svojim ISNI. Alternativno črkovanje, alternativne predstavitve, različni nabori znakov, pisave ali jeziki sami po sebi ne upravičujejo dodelitve različnih identifikatorjev ISNI (ISO 27729, 2013). Opisani primer za Petro Černe Oven, ko so bili njeni identiteti dodeljeni kar trije različni ISNI, kaže, da vsak zapis izvira iz druge baze podatkov. Algoritem za povezovanje ni povezal istih oblik osebnega imena iz dveh različnih baz podatkov (nizozemske nacionalne knjižnice in EThOS), in tudi ne oblik s šumniki in brez šumnikov iz treh različnih baz podatkov. Vzpostavitev enoličnega identifikatorja za isto identiteto ni uspela pri enostavnih različicah osebnega imena s šumniki in brez njih, zahtevnejših primerov, kot so druge pisave in različna transliteracija (npr. cirilica) normativnih oblik z dodatki k imenu v jeziku ustanove, pa nismo preverjali oz. takšnih primerov nismo zasledili.

## 6 Priporočila za kakovostnejši prikaz slovenskih normativnih zapisov v VIAF

Na podlagi analize vpisa in ažuriranja v VIAF smo katalogizatorjem pripravili osnovna priporočila za oblikovanje normativnih zapisov za osebna imena v bazi podatkov CONOR.SI, ki bodo prispevala k izboljšanju prikaza slovenskih normativnih zapisov v VIAF. Pri določanju normativne točke dostopa je treba upoštevati katalogizacijska pravila in navodila (gl. ZNAČKA), pri oblikovanju drugih delov normativnega zapisa pa navodila v priročniku o formatu COMARC/A.

### 6.1 Priporočila za oblikovanje normativnih zapisov za osebna imena

**Dokončno urejeni zapis.** Zapis je dokončno urejen, kadar znamo določiti enotno točko dostopa in smo zbrali dovolj podatkov o osebnem imenu. Ko zapis dokončno uredimo, v

podpolje 001a – Status zapisa vpišemo kodo »c« – popravljen zapis in brišemo kodo »3« – nepopoln zapis v podpolju 001g – Koda popolnosti zapisa (Kos, 2019).

**Pomembnost kodiranih podatkov.** Pri dokončno urejenem normativnem zapisu navedemo vse ustrezne kode, zlasti pa tiste, na katerih temelji izbor zapisov za VIAF. To so:

- 001a = »c« – popravljen zapis,
- 001c = »a« – osebno ime,
- 100b = »a« – sprejeta,
- 120b = »a« – razlikovano osebno ime.

Če je le mogoče, kodiranih podatkov, na katerih temelji izbor zapisov za VIAF, ne spreminjamo več.

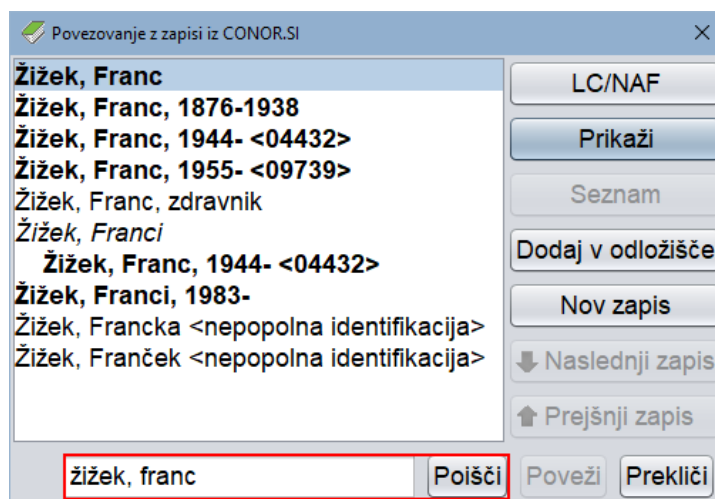
**Različne oblike imena.** V normativnem zapisu navedemo vse variantne oblike osebnega imena (polje 400 – Variantna točka dostopa – osebno ime) iz različnih virov, ki imajo zapise v COBIB.SI.

**Obogateni zapisi.** Zapise obogatimo tudi z javno dosegljivimi podatki o biografiji in delovanju osebe (polje 340 – Opomba o biografiji in delovanju), ki jih lahko preverimo v slovenskih virih ali pri avtorju. Zapise lahko obogatimo tudi z dodatnimi informacijami v drugih poljih, v polju 856 – Elektronska lokacija in dostop recimo lahko navedemo slovenske spletne biografske leksikone, npr. Štajerci.si ([www.stajerci.si](http://www.stajerci.si)), kot to že delajo v Mariborski knjižnici (slika 16).

The screenshot displays the Štajerci.si website interface. On the left, a sidebar lists various identifiers and links, including a red arrow pointing to the URL <http://www.xn--tajerci-pqb.si/osebe/jan%C5%B2ekovi%C4%8D-zora/4/>. The main content area features a search bar, navigation tabs (Začetek, Napredno iskanje, Poklici in dejavnosti, Kraji delovanja, O leksikonu), and a profile for Zora Janžekovič. The profile includes a photograph, birth and death information, and a detailed biographical text in Slovenian. The text describes her as a plastic surgeon and her involvement in the Maribor Red Cross during the war.

Slika 16: Biografski podatki za Zoro Janžekovič (Vir: Štajerci.si, 2020) (levo) in normativni zapis za Zoro Janžekovič (Vir: CONOR.SI, 2020) (desno)

**Soimenjaki.** Pri osebnih imenih nastane največ napak pri soimenjakih, zato je treba že pri kreiranju in urejanju normativnih zapisov posebno pozornost posvetiti soimenjakom. Poleg pazljivega oblikovanja normativne točke dostopa je treba zapise opremiti s pojasnilom o entitetah z enakimi ali podobnimi imeni. To naredimo z vpisom opombe »Ne zamenjaj:« v polju 820 – Podatki o rabi ali pomenskem obsegu. Razen tega je treba pri povezovanju v bibliografskih zapisih v poljih 70X skrbno preveriti pravilnost povezave do normativnega zapisa za osebno ime avtorja (slika 17).



Slika 17: Izpis seznama soimenjakov pri vnosu imena avtorja v bibliografski zapis (Vir: COBIB.SI, 2020)

**Podvojeni zapisi.** Da bi se izognili kreiranju več zapisov za isto osebo, tj. podvojenim zapisom, je treba pred kreiranjem normativnega zapisa iskati po več različnih kriterijih, preveriti točke dostopa v drugih zapisih in dosledno preverjati priporočene popravke ob shranjevanju zapisa. Uporabimo tehnike iskanja, s katerimi lahko preprečimo nastanek podvojenih normativnih zapisov (Kos, 2018).

**Povezovanje normativnih zapisov z bibliografskimi zapisi.** Tiste bibliografske zapise, pri katerih manjka podatek v podpolju 70X3 – *Številka normativnega zapisa*, je treba povezati z normativnimi zapisi za osebna imena. V bazi podatkov COBIB.SI je bilo konec leta 2019 še okrog 23 % bibliografskih zapisov, v katerih polja 70X niso povezana z normativnimi zapisi za osebna imena (IZUM, 2020). Priporočamo povezovanje bibliografskih zapisov z normativnimi predvsem za slovenske avtorje (metoda *Poveži z značnicami v bibliografskih zapisih*, ki je opisana v priročniku COBISS3/Katalogizacija, pogl. 9.4.3).

## 6.2 Priporočila za ravnanje z normativnimi zapisi, ki so že bili poslani v VIAF

V normativnih zapisih v bazi podatkov CONOR.SI, ki so že bili poslani v VIAF, ne spreminjamo kodiranih podatkov, na katerih je temeljil izbor zapisov za VIAF, in ne popravljamo in ne brišemo identifikatorja VIAF ID, ker se vpisuje programsko in je osnova za povezavo iz COBISS+ do skupka v VIAF. Identifikator ISNI lahko spremenimo, če ugotovimo, da pripada drugi bibliografski identiteti. Eni osebi je lahko dodeljen več kot en ISNI, in sicer za vsako javno identiteto svoj (recimo psevdonim Prežihov Voranc in pravo ime Lovro Kuhar imata vsak svoj ISNI). Pri preverjanju pravilnosti ISNI pa moramo upoštevati, da se v bazi podatkov CONOR.SI identifikator ISNI nanaša na identiteto, povezano z osebnim imenom, ki je vpisano v polju 200 – *Normativna točka dostopa – osebno ime*.

## 7 Razprava

Vpis normativnih zapisov v VIAF pomeni vključitev slovenskih normativnih podatkov v mednarodni okvir, glavna prednost pa je pridobitev identifikatorja ISNI. Z vpisom smo za 50,3 % poslanih zapisov brezplačno pridobili identifikator ISNI, ki je sicer plačljiv. Poleg ISNI smo v normativne zapise v bazi podatkov CONOR.SI programsko vpisali tudi VIAF ID. Na tej



osnovi smo lahko v COBISS+ vzpostavili povezave do servisov VIAF in ISNI v zapisih, ki so vpisani v VIAF.

Pri prvem ažuriranju, novembra 2019, so bili v VIAF vpisani tudi normativni zapisi za imena korporacij. Teh zapisov je bilo zelo malo, le 74 ali komaj 0,5 % vseh zapisov za imena korporacij v bazi podatkov CONOR.SI, čeprav smo normativno kontrolo korporacij vzpostavili leto prej, jeseni 2018. Delež zapisov za osebna imena, ki smo jih prispevali v VIAF, je sicer desetkrat večji, a še zmeraj znaša le 5 % vseh zapisov za osebna imena v bazi podatkov CONOR.SI. Glede na to, da je normativna kontrola za osebna imena v sistemu COBISS.SI vzpostavljena že 17 let, je tudi ta delež bistveno premajhen. Glavni razlog je treba iskati pri katalogizatorjih, ki posvečajo premalo pozornosti preverjanju in urejanju normativnih podatkov. Večina jih kreira zapise z nesprejeto obliko osebnega imena ali kratke normativne zapise. Skrajni čas je, da katalogizatorji spremenijo svoje razmišljanje pri kreiranju in redigiranju normativnih zapisov, ki velja za najzahtevnejše delo v postopkih katalogizacije. Da lahko normativna točka dostopa pridobi status *sprejeta*, je treba preveriti identiteto avtorja, da pa lahko zapis dobi status *popravljen zapis*, pa je treba urediti normativni zapis, da je dokončno popravljen, tj. brez slovničnih in vsebinskih napak (Krajnc Vobovnik, 2018). Prednostna naloga katalogizatorjev bi morala biti, da predvsem za slovenske avtorje normativne zapise dokončno uredimo in opremimo z biografskimi podatki in s povezavami na spletne biografske vire.

Glede na vrste problemov, navedenih v *VIAF Guidelines* (2019), smo pri vpisu v VIAF zasledili le napačne povezave med normativnimi in bibliografskimi zapisi ter podvojene normativne zapise (tabela 1). Vsi problemi pri združevanju v skupke v VIAF, ki smo jih zasledili pri analizi napačno povezanih normativnih zapisov iz baze podatkov CONOR.SI, izhajajo iz napake, ko se pri katalogizaciji nepravilno identificirajo soimenjaki.

*Tabela 1: Vrste napak pri vpisu slovenskih normativnih zapisov v VIAF glede na tipologijo problemov pri združevanju v skupke v VIAF*

<b>Problemi po VIAF Guidelines (2019)</b>	<b>Problemi pri vpisu zapisov CONOR.SI v VIAF (2019)</b>
mešana imena (mešani homonimi)	/
manjkajoči naslovi del in datumi	napačne povezave z normativnimi zapisi v bibliografskih zapisih v bazi podatkov COBIB.SI
neobičajno kodirani podatki	/
razlikovanje z uporabo jezikovno odvisnih podatkov	/
podvojeni zapisi	podvojeni zapisi
klasična imena, imena kraljev, kraljic in papežev	/

Tudi če odpravimo vse napake, ki izvirajo iz izvornih podatkov v sistemu COBISS.SI, se napačnemu združevanju v VIAF ne moremo v celoti izogniti. VIAF namreč združuje zapise, kreirane po različnih katalogizacijskih pravilih, in eden izmed ciljev je tudi ohranitev nacionalnih in lokalnih oblik imena v normativnih zapisih. Zato lahko pričakujemo, da se bodo napake pojavljale še naprej. Pomembno pa je, da se zavedamo, katerim napakam se pri katalogizaciji lahko izognemo in katerim težje.

Več dodeljenih ISNI za isto identiteto, ki smo jih odkrili pri preverjanju VIAF Dataset, pomeni, da se izgubita pomen in namen enoličnega identifikatorja za javno identiteto neke osebe. Gre za nekonsistentnost baze podatkov ISNI, zato bomo pri razreševanju še naprej sodelovali s sodelavci VIAF, da uredijo podvojene identifikatorje predvsem za slovenske avtorje v VIAF in v ISNI.

Pri vodenju projekta vpisa normativnih zapisov v VIAF smo glede na analizo vhodnih in izhodnih podatkov pripravili še nekaj dodatnih priporočil za katalogizacijo normativnih zapisov. Priporočila se nanašajo predvsem na osebna imena in so katalogizatorjem splošno znana. V prispevku izpostavljamo, kaj najbolj vpliva na uspešnost združevanja v skupke v VIAF. Poleg tega želimo spodbuditi katalogizatorje, da predvsem za slovenske avtorje kreirajo dokončno urejene normativne zapise tako za osebna imena kot za imena korporacij, saj so samo dokončno urejeni normativni zapisi (s statusom *popravljen zapis*) primerni za izmenjavo na mednarodnem nivoju.

## 8 Zaključek

Še naprej bomo pošiljali normativne zapise za osebna imena in imena korporacij v VIAF. Trenutno je v VIAF vpisanih samo 5,05 % vseh normativnih zapisov v bazi podatkov CONOR.SI, s katerimi smo zajeli 38,5 % bibliografskih zapisov v bazi podatkov COBIB.SI. Glede na inicialni vpis je to le malenkostno povečanje, naš cilj pa je, da vsako leto v VIAF pošljemo čim več čim boljših normativnih zapisov, in to ne samo za osebna imena, ampak tudi za imena korporacij. S priporočili za oblikovanje normativnih zapisov v CONOR.SI ozaveščamo katalogizatorje o pomenu normativne kontrole na globalnem nivoju, o vplivu nesporno in nedvoumno oblikovane normativne točke dostopa ter o pomenu vpisanih osnovnih podatkov o delovanju osebe predvsem za slovenske avtorje. Enako velja tudi za normativne zapise za imena korporacij, predvsem slovenskih, in za čim večje povezovanje le-teh z bibliografskimi zapisi.

Kakovostna, konsistentna in bogata nacionalna normativna baza podatkov, ki je v največji možni meri povezana z bibliografsko bazo podatkov, je osnova za vzpostavitev modela IFLA LRM (angl. *Library Reference Model*) v sistemu COBISS.SI. V vmesniku COBISS+, ki smo ga objavili novembra 2016, smo oblikovali skupke zapisov z namenom, da identificiramo bibliografske odnose v katalogu. Skupke smo oblikovali za dela in njihove pojavne oblike, ki smo jih pridobili na osnovi različnih algoritmov in jih prikazali kot »Vse izdaje in prevodi« (Krajnc Vobovnik in Mazić, 2017). Od začetka leta 2020 pripravljamo osnove za vzpostavitev normativnih zapisov za dela in pojavne oblike, ki jih bomo vpisali v bazo podatkov CONOR.SI ter kasneje tudi v VIAF.

## Reference

Bennet, R., Hengel-Dittrich, C., O'Neill, E. T. in Tillet, B. B., 2006. VIAF (Virtual International Authority File): Linking Die Deutsche Bibliothek and Library of Congress Name Authority File. V: *World Library and Information Congress: 72nd IFLA General Conference and Council*, August 2006, Seoul, Korea. Dostopno na: <https://archive.ifla.org/IV/ifla72/papers/123-Bennett-en.pdf> [18. 3. 2020].

*COBISS3/Katalogizacija*, 2020. Maribor, Institut informacijskih znanosti. Dostopno na: [https://izobrazevanje.izum.si/EntryFormDesktopDefault.aspx?tabid=38&type=manual&manual=1\\_COBISS3\\_Katalogizacija\\_svn](https://izobrazevanje.izum.si/EntryFormDesktopDefault.aspx?tabid=38&type=manual&manual=1_COBISS3_Katalogizacija_svn) [18. 3. 2020].



*How ISNI works*, 2020. Dostopno na: <http://www.isni.org/how-isni-works> [18. 3. 2020].

ISO (International Organization for Standardization), 2013. *Slovenski standard SIST ISO 27729:2013. Informatika in dokumentacija – Mednarodni standardni identifikator imen (ISNI) – Information and Documentation – International standard name identification (ISNI) – Information et documentation – Code international normalisé des nomes (ISNI)*. Ljubljana: Slovenski inštitut za standardizacijo.

IZUM, 2020. *Letno poročilo o delu IZUM za leto 2019*. (Priloga 3: Statistika bibliografskih zapisov, povezanih z normativno bazo podatkov CONOR.SI.). Maribor, Institut informacijskih znanosti.

Kos, J., 2019. Kako sestavimo normativni zapis za osebno ime. *Blog COBISS*, 8. 10. 2019. Dostopno na: [https://blog.cobiss.si/2019/10/08/sestavimo\\_normativni\\_zapis/](https://blog.cobiss.si/2019/10/08/sestavimo_normativni_zapis/) [18. 3. 2020].

Kos, J., 2018. Tehnike iskanja v CONOR.SI pri kreiranju normativnih zapisov za osebna imena avtorjev. *Blog COBISS*, 22. 6. 2018. Dostopno na: <https://blog.cobiss.si/2018/06/22/tehnike-iskanja-conor/> [18. 3. 2020].

Krajnc Vobovnik, A., 2018. Vpis slovenskih normativnih zapisov za osebna imena v VIAF. *Organizacija znanja* (23), 1–2, str. 11–20. Dostopno na: [https://www.cobiss.si/oz/HTML/OZ\\_2018\\_1\\_2\\_final/14/index.html](https://www.cobiss.si/oz/HTML/OZ_2018_1_2_final/14/index.html) [18. 3. 2020].

Krajnc Vobovnik, A. in Mazić G., 2017. Združevanje bibliografskih zapisov v COBISS+. *Knjižnica* (61), 1/2, str. 87–100. Dostopno na: <https://knjiznica.zbds-zveza.si/knjiznica/article/view/6163> [18. 3. 2020].

*VIAF Guidelines*, 2019. Dostopno na: <https://www.oclc.org/content/dam/oclc/viaf/VIAF%20Guidelines.pdf> [18. 3. 2020].

*ZNAČKA: priročnik za določanje značnic pri katalogizaciji*, 2001. Ljubljana, Narodna in univerzitetna knjižnica. Dostopno na: <http://home.izum.si/izum/e-prirocniki/znacka.pdf> [18. 3. 2020].

# Digitalna humanistika: modna muha ali naš (bibliografski) vsakdan?

## Digital humanities: a whim or a (bibliographical) routine?

Sonja Svolfjšak<sup>1</sup>

**IZVLEČEK:** Članek predstavlja in analizira izbrane projekte in raziskovalne iniciative s področja digitalne humanistike, v katere so aktivno vključene knjižnice in druge dediščinske ustanove. Ti ambiciozni globalni projekti dopolnjujejo obstoječa znanja o splošnih in specifičnih družbenih in kulturnozgodovinskih fenomenih ter pisni kulturni dediščini z uporabo spletnih aplikacij in orodij ter kolaborativnega, lokacijsko neodvisnega raziskovalnega pristopa. Posredno članek ponuja vpogled v preplet analognosti in digitalnosti na področju digitalne humanistike in v možnosti, ki jih ta ponuja raziskovalcem in knjižnicam kot upravljavkam z metapodatki in digitalnimi objekti – tudi kot potencialnim pobudnikom, gostiteljem in aktivnim partnerjem v projektih in mednarodnih raziskovalnih skupnostih in iniciativah.

**KLJUČNE BESEDE:** digitalna humanistika, knjižnice, raziskave, projekti

**ABSTRACT:** The article presents and analyses selected digital humanities projects and research initiatives in which libraries and other heritage institutions participate actively. These ambitious global projects supplement the existing knowledge on general and specific social and cultural-historical phenomena in the field of written cultural heritage with the help of web-based applications, tools and the collaborative location-independent research approach. Indirectly, the article also gives insight into the analogue and the digital aspects of digital humanities, as well as into the opportunities that the field offers to researchers, libraries, and other heritage institutions – also as potential initiators, hosts and active partners in such projects, as well as in the international research communities and initiatives.

**KEYWORDS:** digital humanities, libraries, research, projects

---

## 1 Uvod

V množici virov, ki so dostopni v strokovni literaturi, predvsem pa na spletu, je po mnenju avtorice tega prispevka, eno najkonciznejših definicij digitalne humanistike mogoče najti na domači strani [Cambridge Digital Humanities](#) (2020), ki ima kot del Univerze v Cambridgeu eno najdaljših raziskovalnih tradicij na tem področju. Glasi se: »Digitalna humanistika je široko področje raziskav in znanstvenih dejavnosti, ki zajema ne le uporabo digitalnih metod s strani raziskovalcev s področja humanistike in sodelovanje strokovnjakov s področja humanistike z računalniškimi in naravoslovnimi disciplinami, ampak tudi način, kako humanistika ponuja specifičen vpogled v najpomembnejša družbena in kulturna vprašanja, ki jih prinaša razvoj digitalnih tehnologij. Delo na tem področju je nujno kolaborativno, saj vključuje več veščin, disciplin in strokovnih področij«. Temu bi lahko dodali še, da je digitalna humanistika izrazito lokacijsko neodvisna, saj večina dejavnosti in raziskav poteka z uporabo infrastrukture oziroma digitalnih aplikacij in orodij, ki temeljijo na spletnih tehnologijah in rešitvah. Eden obsežnejših

---

<sup>1</sup> Dr. Sonja Svolfjšak, bibliotekarska svetnica, Narodna in univerzitetna knjižnica (NUK), Ljubljana, Slovenija, [sonja.svoljsak@nuk.uni-lj.si](mailto:sonja.svoljsak@nuk.uni-lj.si).

naborov programske opreme oziroma orodij za različne dejavnosti na področju digitalne humanistike je na voljo na domači strani [NYU Libraries](#) (Digital, 2020).

Humanistične in družboslovne znanosti, ki uporabljajo tovrstna delovna in raziskovalna okolja in orodja, so predvsem zgodovinopisje, arheologija, antropologija, etnologija, jezikoslovje, kulturne študije, literarne vede, umetnostna zgodovina, filozofija, teologija, geografija, sociologija, politične vede, pravo, muzikologija, informacijska znanost, bibliotekarstvo, pa tudi različne interdisciplinarne in multidisciplinarne vede.

Med letoma 2016 in 2019 je v okviru Delovne skupine za digitalno humanistiko in digitalno kulturno dediščino [LIBER](#) (2020a) potekalo več aktivnosti, katerih cilj je bil ugotoviti, kakšna je vloga knjižnic in drugih dediščinskih ustanov na področju digitalne humanistike, kaj lahko knjižnice (tudi kot upravljavke z metapodatki in digitalnimi objekti) prispevajo k razvoju področja, na kakšne načine se lahko bibliotekarska stroka s svojimi specifičnimi znanji kar najbolj (pro)aktivno vključuje v raziskovalno skupnost »digitalnih humanistov« in kako je mogoče vzpostaviti povratne zanke pri uporabi raziskav v knjižnicah in stroki. Izsledki delovne skupine so bili objavljeni v obliki strnjenih povzetkov, ki vključujejo tudi sezname izbrane oziroma temeljne literature in virov za posamezno temo (LIBER, 2020b). Eden od namenov omenjenih aktivnosti je bil vnesti nekoliko strukture v poplavo (dostikrat fragmentarnih) informacij s področja digitalne humanistike in knjižnic.

Najpomembnejši dejavniki, ki knjižnicam in drugim dediščinskim ustanovam omogočajo vlogo (pro)aktivnih deležnikov na področju digitalne humanistike, so obsežne izkušnje pri upravljanju z objekti kulturne dediščine, (bibliografskimi) metapodatki in digitalnimi objekti ter njihova tradicija zagotavljanja širokega spektra podpore raziskovalcem in raziskovalnim skupnostim. Številne nacionalne in univerzitetne knjižnice so v preteklih letih prek vzpostavljanja tako imenovanih digitalnih laboratorijev in delavnic postale nepogrešljivi partnerji pri razvoju področja, raziskovalcem pa poleg znanj in veščin zagotavljajo tudi prostore in infrastrukturo za delo. Leta 2019 je na temo vzpostavljanja digitalnih laboratorijev v dediščinskih ustanovah na primer izšel obsežnejši priročnik (Mahey, M. et al., 2020). Tudi v Sloveniji deluje skupina [DARIAH-SI](#) (2020), ki je del panevropske digitalne raziskovalne infrastrukture za umetnost in humanistiko [DARIAH-EU](#) (2020). Njena osnovna dejavnost zajema oblikovanje, implementacijo in promocijo raziskovanja, projektov, orodij in najboljših praks na področju digitalne humanistike. V povezavi z DARIAH-SI, pa tudi neodvisno, potekajo posamezne raziskave in projekti, kot je na primer projekt *Neznani rokopisi slovenskega slovstva 17. in 18. stoletja* (2011), vendar med aktivnimi partnerskimi ustanovami vsaj trenutno ni nobene knjižnice.

Ne glede na obilico entuziazma, ki spremlja digitalno humanistiko, pa nekateri vidnejši akterji na razmah področja gledajo z določeno mero skepse. A. Degkwitz (2018), ki je bil med drugim tudi podpredsednik delovne skupine LIBER za digitalno humanistiko in digitalno kulturno dediščino, je tako denimo zapisal: »Analogno nas vabi – ali celo sili – k doživljanju in premoru in razmišljanju. V nasprotju z digitalnim analogno ni zgolj kopija ali reprezentacija. Če bi sprejeli digitalno humanistiko in njene posledice kot namen sam po sebi, bi to razblinilo vse izkušnje v korist neskončnih zank nesmiselnega medsebojnega delovanja. To je cena, ki jo plačujemo za digitalne medije. In to je cena, ki je redko vključena v analize stroškov in koristi. Analogno lahko prekine neskončno zanko digitalnega, vendar le, če analogni pojavi ostanejo naš cilj.« Temu razmisleku v prid govori dejstvo, da je kljub številnim potencialom aplikacij in orodij velik del

digitalne humanistike po svoji naravi v bistvu izrazito analogen, saj večina problemskih izhodišč in raziskovalnih izsledkov še vedno temelji na povsem »analognem« vložku, ki nenazadnje denimo vključuje tudi programiranje oziroma »učenje« programov za besedilno analizo (digitalna lingvistika) in podatkovno rudarjenje. Na obstoječih epistemoloških in metodoloških izhodiščih posameznih znanosti in ved je zasnovana tudi praktično celotna raziskovalna infrastruktura.

## 2 Digitalna humanistika in knjižnice

Knjižnice s(m)o v digitalno humanistiko lahko vpete aktivno, denimo prek sodelovanja v različnih projektih ter kot gostiteljice digitalnih laboratorijev in delavnic, pa tudi pasivno. Metapodatke in digitalne objekte vseh vrst, ki jih prispevamo v različne bibliografske baze in digitalne zbirke, uporabljajo tudi zelo oddaljeni raziskovalci in raziskovalne skupnosti. Lahko jih agregirajo, rudarijo, indeksirajo, mapirajo, povezujejo, analizirajo, rekonstruirajo, vizualizirajo ipd. S tem ti metapodatki dobivajo nove dimenzije in vrednosti, iz njih pa nastajajo tudi povsem novi raziskovalni viri.

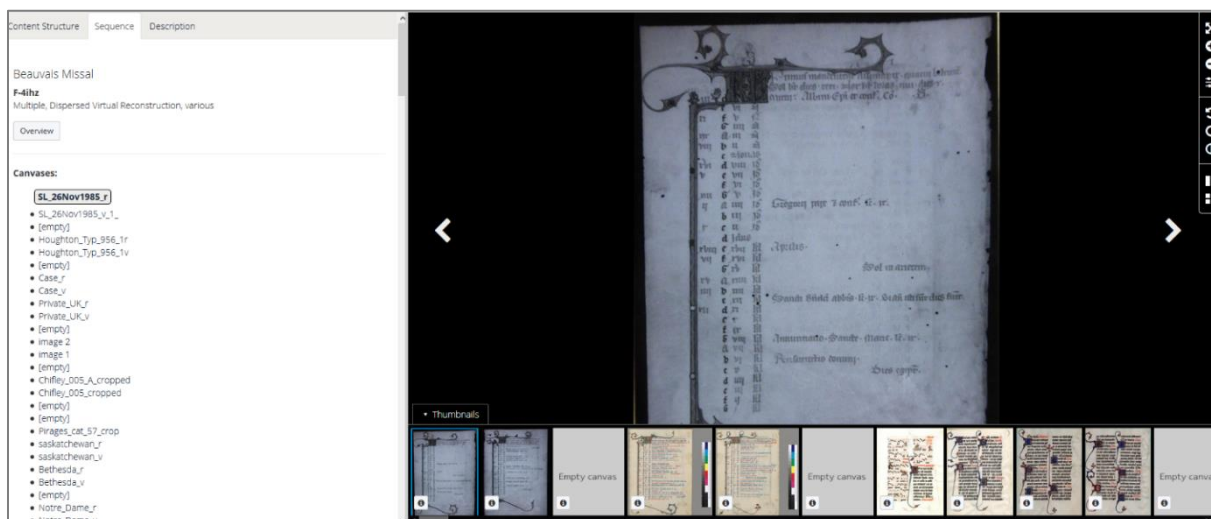
Za nazornejšo predstavo o tem, na kakšne načine se bibliografski metapodatki, digitalni objekti ter bibliotekarska in bibliografska strokovna znanja in veščine (lahko) umeščajo v svet digitalne humanistike, bo v nadaljevanju prispevka predstavljenih nekaj najvidnejših mednarodnih projektov oziroma raziskovalnih infrastruktur, v katere so aktivno vključene knjižnice in ki na osnovi kolaborativnega, lokacijsko neodvisnega raziskovalnega pristopa ter spletnih aplikacij in orodij nadgrajujejo obstoječa spoznanja o splošnih in specifičnih družbenih in kulturnozgodovinskih fenomenih v različnih obdobjih ter kulturni dediščini nasploh. Poudarek bo predvsem na tistih projektih, ki se ukvarjajo z analizami, rekonstrukcijami in vizualizacijami razpršenih historičnih korpusov in so pri tem neposredno odvisni od bibliografskih (meta)podatkov ter digitalnih reprodukcij objektov kulturne dediščine, po svoji naravi in z raziskovalnimi funkcionalnostmi pa so mnogo več kot zgolj zbirke metapodatkov in digitalnih objektov, saj omogočajo aktivno in lokacijsko povsem neodvisno sodelovanje. Posreden namen izbranih predstavitev je med drugim vpogled v preplet analognosti in digitalnosti na področju digitalne humanistike in v možnosti, ki jih ta ponuja knjižnicam in raziskovalcem v knjižnicah in drugih dediščinskih ustanovah – tudi kot potencialnim pobudnikom, gostiteljem in aktivnim partnerjem v tovrstnih projektih in raznih raziskovalnih skupnostih.

## 3 Fragmentologija in Fragmentarium: raziskovanje fragmentov in rekonstrukcija srednjeveških kodeksov

Čeprav uradna ocena ne obstaja, je povsem verjetno, da število ohranjenih fragmentov rokopisov iz latinskega (zahodnoevropskega) srednjega veka presega petnajst milijonov. Nekatero težavo in oviro, povezane z njihovim raziskovanjem, v zadnjih letih premošča mednarodni spletni raziskovalni laboratorij [Fragmentarium](#) (2020). Zasnovan je kot prosto dostopno, vendar strokovno kurirano raziskovalno omrežje, temelji pa na tehnologijah, ki so bile razvite za potrebe baze [E-codices – Virtual Manuscript Library of Switzerland](#) (2020). Knjižnicam, drugim dediščinskim ustanovam, zasebnim zbirateljem in raziskovalcem omogoča objavljanje digitalnih reprodukcij fragmentov, njihovo katalogizacijo oziroma vnos metapodatkov po enotnih standardih, iskanje, primerjalno analizo oziroma vizualno vzporejanje fragmentov iz različnih hranilišč ter njihovo opisovanje in transkripcijo. Na podlagi metapodatkov, digitalnih kopij ter opisov in transkripcij je tako mogoča rekonstrukcija danes

razpršenih srednjeveških kodeksov ne glede na njihovo trenutno hranilišče ter ne glede na lokacijo ustanov in posameznikov, ki jih raziskujejo. V okviru Fragmentariuma potekajo tudi usposabljanja in tematske delavnice.

Spletna aplikacija Fragmentarium temelji na [IIIF](#) (International, 2020) in programskem jeziku MySQL, združuje pa več orodij: orodje za katalogizacijo oziroma vnos metapodatkov, orodje za upravljanje s spletnimi vsebinami, orodje za vnos kontroliranih označevalcev, ključnih besed in faset, orodje, ki omogoča iskanje in raziskovanje digitalnih reprodukcij in metapodatkov, njihovo primerjavo, navzkrižno primerjavo in (vizualno) vzporejanje, in orodje, ki omogoča povezovanje in združevanje metapodatkov ter digitalnih reprodukcij, s tem pa tudi rekonstrukcijo posameznih kodeksov. Orodja so sicer zasnovana na tradicionalnih raziskovalnih pristopih in metodah. Močan poudarek je na vizualno in besedilno primerjalni analitiki; omogočajo jo standardizirani bibliografski metapodatki, po katerih je mogoče iskati in brskati, v kombinaciji s pripadajočimi digitalnimi reprodukcijami fragmentov (slika 1). Kombinacija strokovnega znanja in tehničnih možnosti, ki jih ponujajo orodja, pravzaprav omogoča uresničitev projektov in iniciativ posameznih raziskovalcev in raziskovalnih skupin za dopolnitev obstoječih spoznanj o vsebinskih ter umetnostnih in drugih vidikih *membra disiecta* srednjeveških kodeksov na enem mestu. Ti podatki in iz njih izhajajoča spoznanja vodijo k celostnejšemu poznavanju in razumevanju srednjeveške kulture, znanosti in umetnosti, kot se odražajo prek pisne dediščine. S stališča knjižnic, ki hranijo največ fragmentov in ki tudi prevladujejo med projektnimi partnerji, je Fragmentarium zanimiv s številnih vidikov; na prvem mestu je gotovo dejstvo, da je to trenutno največja tovrstna zbirka podatkov, ki temelji na enotnih metapodatkovnih standardih.



Slika 1: Misal iz Beauvaisa, ki je bil (deloma) rekonstruiran na podlagi metapodatkov in digitalnih kopij fragmentov iz več kot 50 različnih javnih ustanov in zasebnih zbirk (Vir: Fragmentarium, <https://fragmentarium.ms/searchresult/overview/F-4ihz>, 2020)

#### 4 Material Evidence in Incunabula (MEI): globalna bibliografska baza in analitično orodje za področje založništva, knjigotrštva, knjižne umetnosti ter vsebinske strukture in recepcije prvotiskov

Ocenjeno je, da je na svetu ohranjenih približno 30.000 izdaj prvotiskov v približno 450.000 izvodih. Ti kot celota predstavljajo specifičen korpus, ki priča o različnih vidikih knjižne kulture v najzgodnejšem obdobju tiskarstva. Podatki, ki se nanašajo na posebnosti posameznih izvodov prvotiskov, so sicer (pretežno v opisni obliki) že prisotni v knjižničnih katalogih. Digitalna humanistika knjižnicam omogoča, da strokovno podprejo raziskovalno skupnost, zastavlja pa se še vrsta vprašanj o drugih potencialih in vlogi knjižnic na tem področju. Precej večjih zbirk je že na voljo tudi v digitalni obliki, vendar pa posamezni izvodi ali posamezne zbirke metapodatkov in digitalne zbirke prvotiskov ne morejo ponuditi zanesljivih odgovorov na širša vprašanja, kot so na primer:

- Katere knjige so bile tiskane v Benetkah in so bile v 15. stoletju kupljene in uporabljane v Nemčiji ali na Poljskem? Kdo jih je uporabljal in kako? So jih brale tudi ženske?
- Kakšne knjige so v 15. stoletju v Italiji in Franciji kupovali študentje, pravniki ali duhovniki?
- Katere izdaje matematičnih del so kupovale verske ustanove?
- Kdo je v renesansi kupoval Ptolemejeve Kozmografije? Kdo jih je zbiral v 18. stoletju?
- Ali lahko določimo število tiskanih knjig iz 15. stoletja, ki so prišle v evropske in ameriške zbirke kot neposredna posledica razpusta verskih institucij in njihovih knjižnic v 18. in 19. stoletju?
- Kolikokrat so bila v 15. stoletju natisnjena dela Petrarce ali Boccaccia?
- Kolikšen je bil delež klasičnih, srednjeveških ali sodobnih besedil, objavljenih v prvih petdesetih letih tiskarstva?
- Katera dela, natisnjena v 15. stoletju, so danes ohranjena v največjem številu izvodov? In katera v najmanjšem? Kdo je izbiral, urejal in prispeval besedila v izdajah iz 15. stoletja? Koliko del je bilo v 15. stoletju natisnjenih v eni, koliko pa v več izdajah? Koliko izdaj vsebuje ilustracije?
- Ali lahko danes ocenimo kakovost in število tiskov iz 15. stoletja, ki smo jih izgubili v 500-letnem obdobju?

Da bi lahko odgovorili na ta in na druga vprašanja, povezana z založništvom, knjigotrštvom, knjižno umetnostjo, knjižno kulturo ter vsebinsko strukturo in recepcijo posameznih izdaj in izvodov prvotiskov, s tem pa tudi vseh ohranjenih prvotiskov kot specifičnega historičnega korpusa, je bila leta 2016 v okviru projekta [15cBooktrade](#) (2020) zasnovana podatkovna baza [MEI – Material Evidence in Incunabula](#) (2020). Namenjena je beleženju in iskanju podatkov o značilnostih in zgodovini posameznih izvodov danes ohranjenih prvotiskov, saj le-ti (predvsem s svojo sekundarno vsebino in formo) lahko ponudijo natančnejši vpogled v zgoraj omenjene fenomene na globalni ravni. Standardizirani metapodatki v bazi, ki se nanašajo na posamezno izdajo, izhajajo iz baze [ISTC](#) (Incunabula, 2016), metapodatke, ki se nanašajo na posebnosti posameznih izvodov, pa je treba ročno vnesti v obrazec. Vnos (možna je izbira s spustnih seznamov ali vnos besedila) vključuje podatke o vezavi izvoda, vodnih znakih, rokopisnih dodatkih, znamenjih nekdanjega lastništva in kronologiji nekdanjih lastnikov (provenience), ceni, trenutnem hranilišču in nekatere druge podatke, ki so značilni za vsak posamezni izvod. Vsakemu opisovanemu izvodu oziroma zapisu je po končani avtopsiji in opisu dodeljen



edinstven identifikator, lokacije provenienc so povezane tudi z bazo podatkov Geographic Regions, s čimer je omogočena vizualizacija kroženja prvotiskov skozi stoletja, od kraja izdelave pa vse do njihovih sedanjih lokacij. MEI dopolnjujejo naslednje tri pridružene oziroma povezane baze: baza nekdanjih lastnikov, baza parabesedil in baza ilustracij, ki je namenjena iskanju in primerjavi likovnih elementov oziroma ilustracij v okviru istega izvoda, v okviru iste izdaje določenega dela, v okviru različnih izdaj določenega dela in v okviru različnih izdaj različnih del (15cBooktrade, 2020).

Ljubljana, Narodna in univerzitetna knjižnica (SI): Ti II 2844		All Copies	
ISTC No.	ij00213000	Bagnacavallo, Biblioteca comunale 'Giuseppe Taroni' (IT): Inc. 25233	
Author	Jacobus Philippus de Bergamo	Bergamo, Biblioteca Civica Angelo Mai (IT): Inc. 4-158	
Title	Supplementum chronicarum [Italian] Cronicha de tuto el monde vulgare. Tr: C. Franciscus	Ljubljana, Narodna in univerzitetna knjižnica (SI): Ti II 2844	
Imprint	Venice : Bernardinus Rizus, Novariensis, 8 Oct. 1491	London, Wellcome Collection (GB): 1.e.8 (SR)	
Format	f°	Milano, Biblioteca Trivulziana (IT): Triv. Inc. A 88	
Language	Italian	Oxford, Taylor Institution (GB): Arch. Fol. It. 1491 (2).	
GW No.	GW M10976	Pordenone, Biblioteca del Seminario (IT): INC/D/FOR	
Subject	History	Udine, Biblioteca civica Vincenzo Joppi (IT): Thes. II. 136	
Keywords	Literature-social; translation	Venezia, Biblioteca Nazionale Marciana (IT): Inc. V. 0280-281	
Period	contemporary	Wien, Österreichische Nationalbibliothek (AT): Ink 32-36	
<b>Description of Copy</b>			
Copy Id	02013830		
Holding Institution	Ljubljana, Narodna in univerzitetna knjižnica (SI)		
Shelfmark	Ti II 2844		
Copy History	Created by Sonja Svoljšak.		
Copy Features	Complete		
Support material (book)	Paper		
Support material (plates)	Paper		
Watermark	1		
Source	Book in hand.		
<b>Provenance 1739 - 1782</b>			
Place	Amsterdam (Geonames Id: 2759794)		
Area	Netherlands		
Time period	1739 - 1782		
Provenance name	Sepp, Jan Christiaan , 1739-1811 [Person; Former Owner] (Male, Booktrade, No characterisation/layout)		
Provenance Type	Inscription		
Date of evidence	1782		
Method of acquisition	unknown		
Binding date	18th cent.		
Binding type	Boards		
Board material	Paper		
Cover material	Leather		
Binding status	Rebound		
Binding Dimensions	303 × 208 × 50 mm		
Titling	Tooled		
Tooling	Gold		
Edges	Coloured		
Writing on edges	None		
Gaufered	No		
Certainty	The recording of this evidence is considered certain		

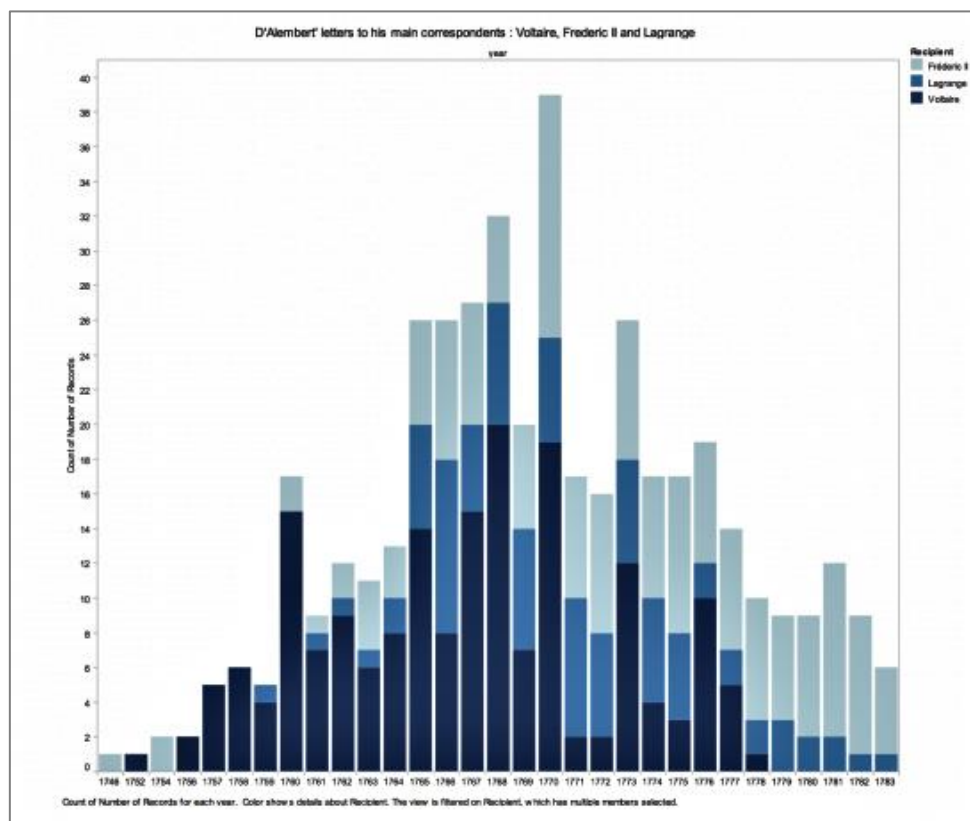
Slika 2: Del zapisa za izvod *Supplementum chronicarum* Jakoba Filipa iz Bergama iz Zbirke Žige Zoisa, ki ga hrani NUK (Vir: MEI, <https://data.cerl.org/mei/02013830>, 2020)

Pri dopolnjevanju baze, ki se je leta 2014 pod vodstvom Cristine Dondi s fakultete Lincoln College Univerze v Oxfordu in Data Conversion Group z Univerze v Göttingenu začela kot evropski projekt, od leta 2019 dalje pa poteka pod okriljem [CERL](https://www.cerl.org/) (Consortium, 2020), sodeluje svetovna raziskovalna skupnost na čelu z raziskovalci – bibliografi iz nacionalnih in univerzitetnih knjižnic, ki hranijo večje zbirke prvotiskov. Problematiko manjših in nejavnih zbirk na primer v Italiji rešujejo tudi ob pomoči potujočih katalogizatorjev/popisovalcev, ki obiskujejo samostane, razne zasebne ustanove in zbiratelje. Vsako leto je za namen dopolnjevanja baze razpisanih tudi več raziskovalnih štipendij. Za pristop k sodelovanju je potrebna predhodna udeležba na delavnici, v času nastanka tega prispevka pa je bilo v MEI popisanih 51.681 posameznih izvodov prvotiskov s celega sveta, predvsem pa iz evropskih in ameriških zbirk. Čeprav se opisom posameznih izvodov pogosto pridružujejo tudi digitalne

kopije, je ročen vnos (meta)podatkov, ki so na podlagi avtopsije specifični za posamezen izvod, bistvenega pomena za analize, ki jih MEI predvideva oziroma omogoča (slika 2).

## 5 Mapping the Republic of Letters: analiza in vizualizacija korespondenčnih in drugih medosebnih stikov od renesanse do razsvetljenstva

[Mapping the Republic of Letters](#) (2020) je projekt skupine Humanities+Design in Univerze Stanford, med partnerskimi ustanovami pa so tudi posamezne evropske univerze in raziskovalne ustanove. Osredotoča se na raziskovanje omrežij, ki so jih prek literarne republike, osebnih stikov, objav in potovanj ustvarili posamezni znanstveniki in literati od renesanse do razsvetljenstva. Ta omrežja so bila od 16. do 19. stoletja ključna za diseminacijo znanstvenih in drugih informacij, s tem pa tudi za razvoj znanosti, književnosti, ekonomije, kulture, politike in drugih področij družbe. Namen posameznih študij primerov, med katerimi so denimo Condorcetova, D'Alembertova, Voltairova, Franklinova, Galilejeva, Lockova in Kircherjeva korespondenca, vizualne in relacijske rekonstrukcije evropskih plemiških zrelostnih potovanj (angl. *grand tours*), literarnih salonov, znanstvenega korespondenčnega omrežja v okviru Španskega imperija in krajev, v katerih so bila izdana Voltairova dela, je ugotoviti, kako obsežne so bile te mreže, kakšne so bile njihove značilnosti in kako so se razvijale.

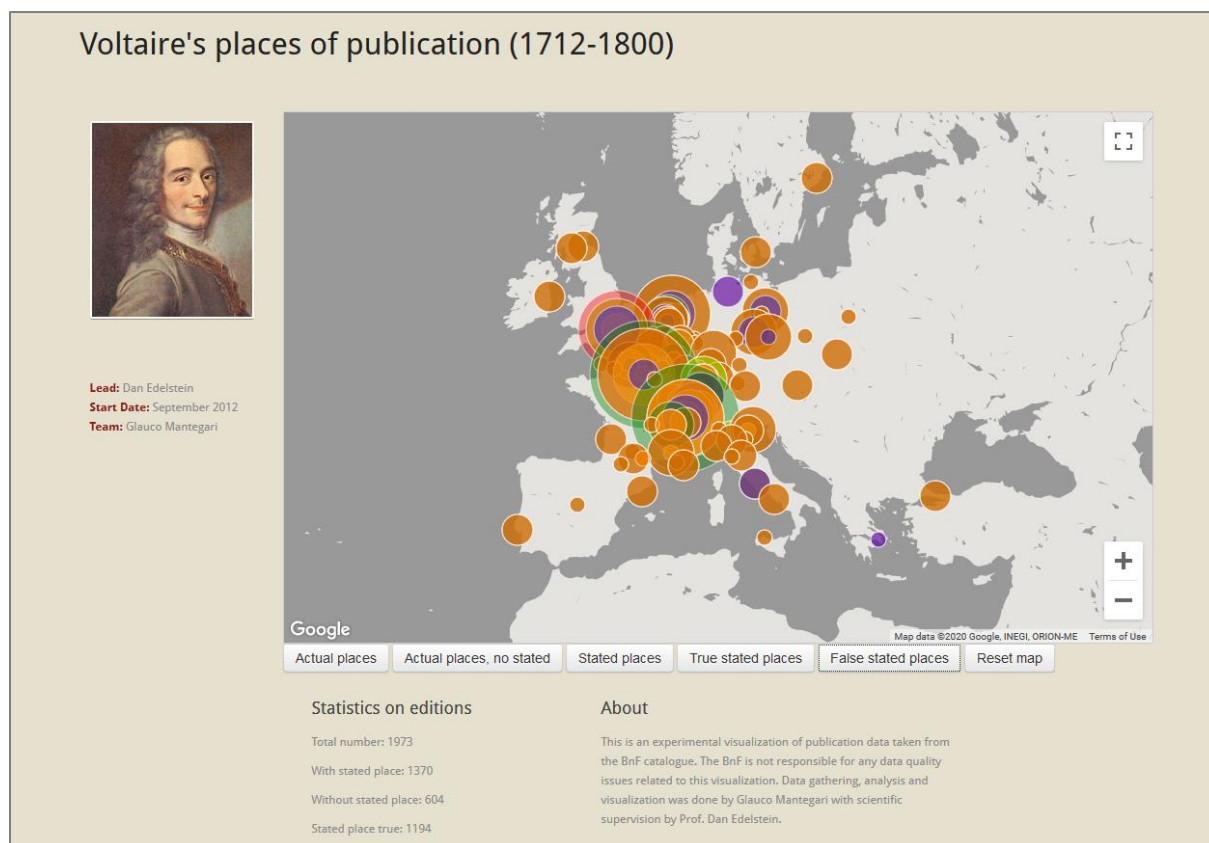


Slika 3: Grafični prikaz D'Alembertove korespondence z Voltairom, Lagrangeem in Friderikom II. med letoma 1748 in 1873 (Vir: [Mapping the Republic of Letters](http://republicofletters.stanford.edu/casestudies/dalembert.html), <http://republicofletters.stanford.edu/casestudies/dalembert.html>, 2020)

Mrežo stikov med posamezniki pa tudi ustanovami, kakršne so bile recimo znanstvene akademije, je mogoče rekonstruirati na podlagi posameznih študij in baz (meta)podatkov,



digitalnih zbirk ohranjenih pisem in drugih zapiskov ter objavljenih del, ki jih hranijo različne dediščinske ustanove na čelu s knjižnicami. Podatki o osebnih in drugih stikih posameznih oseb, zbranih z uporabo aplikacij in orodij za vizualizacijo različnih družbenih fenomenov, omogočajo tudi njihove statistične, kronološke in prostorske analize ter prikaze (sliki 3 in 4).



*Slika 4: Vizualizacija krajev izida Voltairovih del, ki jih hrani Francoska nacionalna knjižnica, v obdobju med 1712 in 1800 s filtri za: dejanske kraje izida, kraje izida, ki niso navedeni, pa so bili ugotovljeni iz drugih virov, kraje izida, ki so navedeni, kraje izida, ki so navedeni in so tudi dejanski kraji izida, ter namenoma napačno navedene kraje izida (Vir: [Mapping the Republic of Letters](http://republicofletters.stanford.edu/casestudies/voltairepub.html), 2020)*

## 6 Nekaj sklepnih misli

Digitalna humanistika knjižnicam omogoča, da strokovno podprejo raziskovalno skupnost, zastavlja pa se še vrsta vprašanj o drugih potencialih in vlogi knjižnic na tem področju. Prvo je gotovo vprašanje o strategiji upravljanja z metapodatki ter gradnji bibliografskih baz in digitalnih zbirk na tak način, da bi bile te uporabne za čim več raziskovalnih področij, vključno z lokacijsko neodvisno in avtomatizirano analitiko večjih količin (meta)podatkov in digitalnih objektov. Bolj(e) kot so ti strukturirani, medsebojno povezani in kontekstualizirani, bolj(e) se lahko posvečamo vsebinskim vidikom bibliografskega in drugega strokovnega dela, raziskovanju in razvoju, spoznavanju in uporabi sodobnih standardov, metodologij in (digitalnih) orodij ter (kritičnemu) razumevanju svojega strokovnega dela v širšem znanstvenoraziskovalnem kontekstu. Drugo je pogosto prezrto vprašanje lastnih iniciativ knjižnic na področju sistematičnega raziskovanja in kontekstualizacije večjih, predvsem globalno relevantnih korpusov pisne kulturne dediščine tudi v digitalnem okolju oziroma z

uporabo (kolaborativno zasnovanih) spletnih aplikacij in analitičnih orodij. Le z lastnimi pobudami in (pro)aktivnim pristopom lahko knjižnice in druge ustanove, ki hranijo kulturno dediščino, postanejo aktivni deležniki in sodelavci v projektih, tudi tistih, ki po svoji naravi in zasnovi deloma ali v celoti spadajo na področje digitalne humanistike. Potenciali bibliografskih baz in digitalnih zbirk so – v kombinaciji s preišljenimi vprašanji, ki si jih zastavimo v zvezi z njihovo naravo, vlogo, pomenom in kar najširšo možno uporabnostjo – praktično neomejeni. Kljub temu pa se, kot pravi A. Degkwitz (2018) in kot je razvidno tudi iz izbranih projektov in iniciativ, predstavljenih v tem članku, digitalna humanistika začne s temeljitim poznavanjem ter strokovno bibliografsko obravnavo primarnih virov in korpusov, ki jih hranimo. Aplikacije in orodja, ki spadajo v območje »digitalnega«, so pri tem pravzaprav večinoma le preišljeno zasnovani pripomočki za hitrejšo in učinkovitejšo analizo in predstavitev velikih količin pogosto zelo razpršenih podatkov, ki presegajo človeške, torej »analogne« kapacitete.

## Reference

*15cBooktrade*, 2020. Dostopno na: <http://15cbooktrade.ox.ac.uk> [25. 5. 2020].

*Cambridge Digital Humanities*, 2020. Dostopno na: <http://www.cdh.cam.ac.uk/cdh/what-is-dh> [25. 5. 2020].

*Consortium of European Research Libraries*, 2020. Dostopno na: <http://www.cerl.org> [25. 5. 2020].

*DARIAH-SI*, 2020. Dostopno na: <http://www.dariah.si> [25. 5. 2020].

*DARIAH-EU*. Dostopno na: <http://www.dariah.eu> [25. 5. 2020].

Degkwitz, A., 2018. The revenge of the analogue. *De Gruyter Conversations*, 22. 8. 2018. Dostopno na: <https://blog.degruyter.com/the-revenge-of-the-analogue> [25. 5. 2020].

*Digital humanities: tools & software: an introductory guide to the wide world of the digital humanities*, 15. 4. 2020. Dostopno na: <http://guides.nyu.edu/dighum/tools> [25. 5. 2020].

*E-codices – Virtual Manuscript Library of Switzerland*, 2020. Dostopno na: <http://www.e-codices.unifr.ch/en> [25. 5. 2020].

*Fragmentarium: Digital Research Laboratory for Medieval Manuscript Fragments*, 2020. Dostopno na: <http://fragmentarium.ms/> [25. 5. 2020].

*International Image Interoperability Framework (IIIF)*, 2020. Dostopno na: <http://iiif.io> [25. 5. 2020].

*Incunabula Short Title Catalogue (ISTC)*, 2016. Dostopno na: <http://data.cerl.org/istc> [25. 5. 2020].

*LIBER. Digital Humanities & Digital Cultural Heritage Working Group*, 2020a. Dostopno na: <http://libereurope.eu/strategy/digital-skills-services/digitalhumanities> [25. 5. 2020].

*LIBER. A Digital Humanities Reading List: Part 1*, 2020b. Dostopno na: <http://libereurope.eu/blog/2018/01/29/digital-humanities-reading-list-part-1> [25. 5. 2020].

Mahey, M., Al-Abdulla, A., Ames, S., Bray, P., Candela, G., Chambers, S., et al., 2020. *Open a GLAM Lab*. Doha: Digital Cultural Heritage Innovation Labs, Book Sprint. Dostopno na: <http://glamlabs.io/books/open-a-glam-lab> [25. 5. 2020].

*Mapping the Republic of Letters*, 2020. Dostopno na: <http://republicofletters.stanford.edu> [25. 5. 2020].

*Material Evidence in Incunabula (MEI)*, 2020. Dostopno na: <http://data.cerl.org/mei> [25. 5. 2020].

*Neznani rokopisi slovenskega slovstva 17. in 18. stoletja (NRSS)*, 2011. Dostopno na: <http://ezb.ijs.si/fedora/get/nrss:nrss/VIEW/> [25. 5. 2020].

# Zgodovinska vloga operacijskega sistema VMS pri razvoju COBISS-a

## The historical role of the VMS operating system in the development of COBISS

Davor Šoštarič<sup>1</sup>

**IZVLEČEK:** V prispevku se z nekaj zgodovinske distance spomnimo na razmere, ki so pripeljale do tega, da je operacijski sistem VMS postal podlaga za potrebe razvoja in implementacije nacionalnega vzajemnega bibliografskega sistema COBISS. Najprej spoznamo nekatere bistvene značilnosti operacijskega sistema VMS, nato pa spremljamo pot njegove transformacije v tej podporni vlogi za COBISS, vključno s kasnejšo opustitvijo.

**KLJUČNE BESEDE:** operacijski sistem, VMS, OpenVMS, VAX/VMS, COBISS

**ABSTRACT:** The paper reviews, with some historical distance, the circumstances leading to the VMS operating system becoming the basis for developing and implementing the COBISS national shared bibliographic system. Following a description of the basic characteristics of VMS, the paper shows the path of its transformation in its supportive role to COBISS, as well as its omission later on.

**KEYWORDS:** operating system, VMS, OpenVMS, VAX/VMS, COBISS

---

## 1 Uvod

Skupno sistemsko programsko opremo, ki je odgovorna za delovanje celotnega računalnika (strojna, programska, aplikativna in komunikacijska oprema), obenem pa skrbi za skupne vire in servise za računalniške programe in interakcijo z okolico, označujemo z izrazom **operacijski sistem**. Praviloma je imela vsaka vrsta računalnikov svoj operacijski sistem; le zelo sorodni računalniki iz iste ali sorodne družine posameznega proizvajalca so uporabljali istega. Kasneje je ta izključna navezanost med računalnikom in operacijskim sistemom sicer izginila, saj so nekateri proizvajalci začeli izdelovati računalnike, na katere so nameščali operacijske sisteme neodvisnih proizvajalcev. Vendar pa v času, ko se je v Mariboru vizionarska ekipa entuziastov začela ukvarjati s knjižničnimi informacijskimi sistemi, iz katerih je kasneje zrasel danes vsem znani COBISS, ni bilo možno razmišljati o ločenih, samostojnih poteh računalnika in operacijskega sistema.

V začetku osemdesetih let prejšnjega stoletja se je v Sloveniji začel krhati na videz nedotakljivi imperij svetovnega računalniškega giganta, ki je praktično do zadnje pore obvladoval slovensko računalniško sceno. Vsaj dva slovenska predstavnika sta začela oglaševati »svojo« proizvodno linijo, čeprav je šlo le za licenčno kopiranje in lepljenje slovenskih imen na sicer znane tuje blagovne znamke. Zelo ugodna tržna pravila za akademska okolja so pripeljala računalnike z operacijskim sistemom VMS tudi na Univerzo v Mariboru in v njen novoustanovljeni Računalniški center Univerze v Mariboru (RCUM). Iz njega izhajajoči kasnejši samostojni javni zavod IZUM je tej informacijskotehnoški usmeritvi v začetku devetdesetih sledil.

---

<sup>1</sup> Davor Šoštarič, Institut informacijskih znanosti (IZUM), Maribor, Slovenija, [davor.sostaric@izum.si](mailto:davor.sostaric@izum.si).

RCUM in za njim IZUM sta tako na svojih računalniških sistemih praktično od začetka za vse svoje potrebe uporabljala operacijski sistem VMS. V letu 2020 IZUM ugaša še zadnje produkcijske aplikacije, ki so tekle na tem sistemu. Trenutno je operacijski sistem VMS v IZUM-u začasno v uporabi le še za nekatere neizogibne specialne aplikacije, nadzor podatkovne integritete, začasne podporne postopke, kontinuiteto sistemov COBISS zunaj Slovenije in statistične izračune; z zaključkom leta 2020 bo po vsej verjetnosti tudi tega konec. Obdobje skoraj štiridesetih let zvestobe svetovni uspešnici bo s tem zaokroženo. Vsekakor je bilo to obdobje zelo uspešno.

## 2 Značilnosti operacijskega sistema VMS

Družina 32-bitnih računalnikov VAX ameriškega proizvajalca Digital Equipment Corporation (DEC) je bila zasnovana na posebni arhitekturi z novimi neodvisnimi operatorji, podatkovnimi tipi, mehanizmi naslavljanja in nabori ukazov, ki jih drugi njihovi računalniki niso poznali. Na omenjenih računalnikih VAX je tekel operacijski sistem VMS. V skladu z različnimi poimenovanji procesorjev in novimi generacijami se je tudi ime tega operacijskega sistema pojavljalo v več različicah: VAX/VMS, OpenVMS pa tudi samo VMS. V tem prispevku bomo uporabljali izraz operacijski sistem VMS.

Pri obeh nazivih – VAX za računalnik in VMS za operacijski sistem – je ključnega pomena črka **V**, ki je kratica za besedo *virtual* (navidezni). Celotno ime VMS pomeni *Virtual Management System*, torej sistem navideznega upravljanja. Upravljanje ni v resnici čisto nič navidezno; mišljeno je *upravljanje navideznega naslovnega prostora*. Operacijski sistem VMS namreč zna za vsakega uporabnika sproti naslavljanju (in s tem uporabljati) zelo velik pomnilniški naslovni prostor; več kot štiri milijarde (od 0 do  $2^{32} - 1$ ) naslovov obsega *virtualni naslovni prostor*, računalniška arhitektura pa je omogočala hitre in zanesljive preslikave med fizičnim pomnilnikom in diskovnim prostorom za shranjevanje pomnilniških vsebin, ki začasno nimajo prostora v fizičnem pomnilniku. Vsakemu uporabniku je na voljo njegov celotni navidezni (*virtualni*) prostor. Uporabnik se ne ukvarja z delitvijo na realni in navidezni del. Zanj je vse en sam enotni pomnilniški prostor. Operacijski sistem VMS zna poskrbeti, da so pomnilniške vsebine najprej ustrezno razporejene na večje enote – strani, v postopku njihovih obdelav pa uporabniku neopazno prehajajo iz fizičnega pomnilnika v prostor na disku in obratno.

Domorodni nabor ukazov (angl. *native instruction set*) v splošnem ni karakteristika operacijskega sistema, ampak je vezan na procesor (v našem primeru prvotno VAX, kasneje Alpha in Itanium). Tu ga omenjamo le zaradi popolnejše slike. Pri procesorjih VAX je šlo za zelo obogaten nabor ukazov, kasnejša Alpha in Itanium pa že uporabljata omejen nabor ukazov (angl. *RISC – Reduced Instruction Set Computing*), ki je prišel v uporabo pri skoraj vseh kasnejših procesorjih na tržišču. V svoji osnovni različici pozna šest primarnih podatkovnih tipov operandov, skupaj z dodatnimi variacijami obstaja skupno 14 (z upoštevanjem predznačenosti pa celo 19) podatkovnih tipov različnih velikosti, natančnosti in obsegov. Osnovna računalniška beseda je velika 32 bitov na procesorjih VAX ter 64 bitov na procesorjih Alpha in Itanium. Na voljo je 16 registrov (enake velikosti kot osnovna beseda). Nekateri med njimi imajo posebno vlogo (programski števec, kazalec sklada, kazalec seznama argumentov, status procesorja ...) in jih programer lahko uporabi neposredno za konkretne namene. Ukazi so praviloma zelo intuitivno razumljivi, med njimi pa so tudi dokaj nenavadni ukazi v primerjavi z ukazi v drugih sistemih (npr. ukaz za neposredno izračunavanje vrednosti polinoma, selektivni stavki, ukazi za delo z velikimi kosi pomnilnika naenkrat ...). Osnovne operacije nad

posameznimi operandi so praviloma odvisne od podatkovnega tipa, ukazi za vejanje, pogojne in brezpogojne skoke ter za operacije nad biti in bitnimi polji pa so na pogled zelo podobni tistim, ki jih srečamo v drugih sistemih.

Tukaj srečamo praktično vse svetovno razširjene programske jezike tistega časa: fortran, cobol, PL/I, RPG, pascal, algol, basic, C, lisp ... Operacijski sistem VMS je imel tudi nekatere bolj ali manj lastne programske jezike, ki so bili zunaj njegovega okolja oziroma zunaj okolja sorodnih računalnikov istega proizvajalca zelo redki, npr. zbirnik in bliss. Vsem je bilo skupno to, da so se lahko med seboj poljubno klicali in uporabljali enotne mehanizme prenosa parametrov oziroma argumentov. Ne glede na jezik izvorne kode so prevajalniki programerjeve izdelke predelali v *objektni modul*, pri katerem se je praviloma izgubila sled izvornega programskega jezika. Načeloma bi torej algoritem, sprogramiran v dveh različnih programskih jezikih, dal skoraj identična objektna modula. Več objektnih modulov med seboj in z vnaprej pripravljenimi moduli sistemskih knjižnic je poseben program (angl. *linker*) povezal v izvršljivo kodo (angl. *executable image*).

Operacijski sistem VMS je tudi večopravilen, paketno in transakcijsko orientiran ter omogoča delitev istega računalniškega vira med več uporabnikov (angl. *time-sharing*).

### 3 Prednosti za COBISS

Med prvimi informacijskimi podporami, ki jih je RCUM začel razvijati za potrebe mariborske univerze, je bila tudi izgradnja knjižničnega informacijskega sistema. Glede na vse že akumulirano znanje, jasno načrtano podporno pot jedrnega razvoja informacijske tehnologije na obeh takratnih slovenskih univerzah in ekonomske diktate računalniške nove dobe je bila odločitev o podlagi operacijskega sistema VMS za predhodnika nacionalnega vzajemnega bibliografskega sistema COBISS popolnoma naravna. Ni šlo torej za nikakršno z analizami podkrepljeno zavestno odločitev, ampak se je tako zgodilo le zaradi takratnega stanja.

Preden si ogledamo pogloblitve prednosti operacijskega sistema VMS, ki smo jih prepoznali kot posebej pomembne pri razvoju sistema COBISS, omenimo še dve navidezno oteževalni okoliščini za odločitev za uporabo prav tega operacijskega sistema. Operacijski sistem VMS ni imel lastnega sistema za upravljanje z bazami podatkov. Iz zgodovinskih razlogov je bil v slovenskem okolju in tudi pri nas razširjen mrežni sistem baz podatkov TOTAL ameriškega proizvajalca Cincom Systems (v kasnejših različicah ULTRA in SUPRA – sistem SUPRA se je že spogledoval z relacijskim modelom), tako da ni čudno, da so tudi začetki sistema COBISS delovali na tej platformi. Slovenski distributer računalnikov, na katerih je teklen operacijski sistem VMS, je razvil lastni produkt za upravljanje baz podatkov, vendar so testiranja pokazala, da za naše namene ne zadošča. Kmalu se je pokazalo, da nobeden od nam dostopnih sistemov za upravljanje baz podatkov ne ustreza. Vsi sistemi so namreč izhajali iz klasične poslovne informatike, mi pa smo se srečevali z zakonitostmi, ki so bile s temi orodji zelo težko obvladljive. Tu mislimo predvsem na številskemu svetu podrejene podatkovne strukture, pri svojem delu pa smo imeli opravka z nizi znakov zelo spremenljivih dolžin. Seveda so obstajale svetovne ustrezne komercialne rešitve, vendar so nam bile cenovno nedostopne. Zaradi tega in tudi zaradi odličnega internega sistema za delo z datotečnimi zapisi smo se odločili, da razvoja sistema COBISS ne vežemo na noben uveljavljeni sistem za upravljanje baz podatkov, ampak da naredimo svojega, ki bo temeljil na standardiziranem datotečnem sistemu. Temu načelu smo ostali zvesti precej dolgo.

Podobna zgodba se je odvijala na področju programskih modulov za vnos podatkov (angl. *screen management services*). Uveljavljena, a za naše namene že zastarela orodja za vnos in manipulacijo podatkov oziroma interakcijo med človekom in programskim okoljem smo nadomestili s povsem novimi sistemi, ki so bili plod lastnega znanja in razvoja.

Katere so torej pglavitne prednosti operacijskega sistema VMS, ki so se nam zdele posebej pomembne pri razvoju sistema COBISS? To so:

### **3.1 Učinkovita razpršenost delovnih mest (cenovno zelo ugodno terminalsko omrežje)**

Poslovna informatika se je v zgodnjih osemdesetih letih prejšnjega stoletja ravno začela oddaljevati od masovnega zajemanja podatkov iz obrazcev in se približevati k računalniku prijaznejšemu mediju (kartice in trakovi) v specializiranih oddelkih za zajemanje podatkov (pog. *luknjačice*), tako da so za to odgovorni sodelavci neposredno vnašali podatke v računalniški sistem prek namenskih enot (terminali). Hkrati se je zgodil premik iz tipičnih *paketnih obdelav* (naknadno in praviloma zunaj običajnega delovnega časa) v sprotno (*interaktivno*) obdelavo podatkov. Zato je bilo potrebno dvoje: drugačna zasnova programskih algoritmov in programov ter ustrezno opremljeno delovno mesto (terminali, ponekod tudi majhni tiskalniki) neposredno pri strokovnem sodelavcu. Pri računalnikih, na katerih je tekel operacijski sistem VMS, so bili terminali razmeroma preprosti (ekran s tipkovnico brez lastnih procesorskih zmogljivosti), njihova povezava do centralnih računalniških enot pa tehnično in cenovno dokaj nezahtevna (tudi na nekoliko večje razdalje), še posebej po uveljavitvi posrednih poti (npr. pasivno omrežje s protokolom *ethernet*). Nekateri konkurenčni sistemi so temu sicer sledili, nekateri pa še dolgo ne. K ekonomskemu učinku je pripomogel tudi znaten prihranek pri človeških virih.

### **3.2 Sobivanje različnih programskih jezikov z enotnim principom delovanja**

Večina takratnih računalniških sistemov pri nas je preferirala enega ali kvečjemu dva bolj razširjena splošna programska jezika (npr. fortran ali pascal za naravoslovne in tehniške probleme ter cobol ali PL/I za poslovne probleme). Vsak računalniški sistem je imel tudi ekskluzivni sebi lasten zbirni jezik (angl. *assembler*), ki je zahteval zelo specifično izobraževanje in poznavanje jedrnih internih struktur operacijskega sistema. Na računalnikih z operacijskim sistemom VMS so drug ob drugem brez posebnosti tekli skoraj vsi svetovno znani programski jeziki, ki so bili med seboj popolnoma kompatibilni, zamenljivi, medsebojno klicljivi in so delovali na identičnih zasnovah. Tako so bile prve verzije programske opreme za posamezne segmente sistema COBISS napisane v cobolu z dodanimi posebnimi moduli v fortranu in zbirniku, kasneje pa se je zbirnik obdržal samo še za najbolj specialne systemske module in orodja, glavno vlogo programja pa je skoraj v celoti prevzel basic.

### **3.3 Bogata dokumentacija (tudi iz virov, neodvisnih od proizvajalca)**

Poznavanje višjih programskih jezikov (vsi razen zbirnikov) je praviloma zadoščalo za klasično programiranje (kodiranje algoritmičnih zasnov). Za optimiziranje izvršljive programske kode pa je bilo koristno poznati notranji ustroj računalniške arhitekture in operacijskega sistema VMS. Govorimo o *internih gradnikih*; le-ti se tukaj pojavljajo v natančno oblikovanih *podatkovnih strukturah*, ki so specifična operacijskega sistema VMS. Za razliko od večine takratnih računalniških sistemov v naši okolici to znanje ni bila izključna domena posvečenih. Že proizvajalec je dal na voljo ogromno odličnih priročnikov, osnovna uradna priročniška dokumentacija pa je v svojih dveh najobsežnejših izdajah (»siva« in »oranžna«) pred



prehodom na izključno e-obliko presejala dolžino štirih metrov (tekoča postavitve knjige ob knjigo). Tudi slovenski distributer operacijskega sistema VMS je prirejal odlične tečaje in seminarje z izvrstnim in obsežnim spremljajočim gradivom.

Zaradi proizvajalčeve politike odprtosti in neekskluzivnosti so številni založniki (predvsem v ZDA) izdajali kopico zelo pomembnih priročnikov in vrhunskih knjig s področja razkrivanja internih skrivnosti operacijskega sistema VMS. V našem okolju (RCUM in IZUM) smo poleg uradne literature pridno posegali po predvsem naslednjih knjigah (slika 1): *VAX Architecture Handbook* (Digital Equipment Corporation, 1981), *VAX-11 Assembly Language Programming* (Baase, 1983) in *VAX/VMS Internals and Data Structures* (Goldenberg in Kenah, 1991). Prva knjiga prihaja sicer iz založniške produkcije proizvajalca računalnikov, na katerih je tekel operacijski sistem VMS, drugi dve pa iz neodvisnih krogov. Vse tri so imele v RCUM-u in IZUM-u skoraj ikonični status in še danes zavzemajo posebno mesto v avtorjevi osebni in IZUM-ovi knjižnici.



Slika 1: Najpomembnejši viri znanja za VMS (Vir: osebni arhiv, 1982–2020)

### 3.4 Zаметki omrežnih aplikacij

Zamislil o povezovanju računalnikov so le malo mlajše od računalnikov samih, vendar se je fokus uporabnosti teh idej močno spremenil. V začetku je šlo za težnjo po večji procesorski moči in izmenjavi podatkov ter boljšo izkoriščenost podatkovnih kapacitet, verjetno pa tudi še za rezervni načrt za okrevanje po hujšem izpadu. Vsekakor je bil pogled na omrežje povsem drugačen kot danes (tudi za izraz »omrežje« se je uporabljala danes nezaželena »mreža«).



Zasnova operacijskega sistema VMS je že od začetka presegala te poglede. Omogočala je neposredno interakcijo med programi v realnem izvajalnem času. Enotni sistem naslavljanja podatkovnih virov je poleg diskovne enote, mape in imena datoteke v ožjem smislu predvideval tudi ime omrežnega gostitelja. Nekaj podobnega so sicer poznali konkurenčni računalniški in operacijski sistemi, in tudi VMS ni bil izjema v tem, da je omrežje moralo biti homogeno. Možnost povezovanja virov na različnih sistemih je bila še vedno iluzija. Lepota lastnega koncepta računalniškega omrežja (prvotna blagovna znamka je nosila ime *DECnet*) je bila tudi v tem, da je močno ustrezala teoretičnemu sedemnivojskemu referenčnemu modelu (*OSI*), ki je kasneje postal standard in so mu sledile praktično vse komercialno uspešne različice omrežnih konceptov (ne nazadnje tudi različice omrežij po protokolih TCP/IP, vključno z internetom). Ne glede na naravno delitev sedemnivojskega referenčnega modela na tri skupine plasti (medijska, gostiteljska in aplikativna raven) je bila moč VMS-ovske implementacije prav v najvišji (aplikativni) ravni. Čeprav terminološko nesistematičen in zmeden je operacijski sistem VMS omogočal elegantno komunikacijo med programi najprej znotraj domorodnega omrežja DECnet, kasneje pa tudi po sodelujočih omrežjih, ki so upoštevale skupne de facto standarde.

### 3.5 Prefinjen datotečni sistem in delo z datotečnimi zapisi

Operacijski sistem VMS je uporabljal datotečni sistem z nazivom *Files-11*. Že iz dela imena datotek (tip) se je dalo sklepati na vsebino in spekter predvidenih tipov datotek je bil zelo širok. Struktura osnovnih komponent datotečnega opisa (angl. *file header*) je bila enotna ne glede na predvideno uporabo, medij ali karakteristike strojne opreme.

Pravo bogastvo pa je predstavljal sistem za delo z datotečnimi zapisi (*RMS – Record Management System*), ki je s svojimi bogatimi knjižnicami klicljivih modulov omogočal elegantno delo neposredno tudi iz višjih programskih jezikov. Široka paleta kombinacij datotečne organizacije s sistemom datotečnih zapisov in sistemom zaklepanja zapisov (angl. *lock management*) je načrtovalcem in programerjem dajala skoraj neomejene možnosti.

## 4 Razlogi za opustitev

Po približno četrto stoletja sožitja je operacijski sistem VMS kljub novim verzijam začel ovirati nadaljnji razvoj sistema COBISS. Večino tega, kar je v prejšnjem poglavju omenjeno kot izrazita prednost, so začeli ponujati tudi drugi operacijski sistemi. Proizvajalci računalnikov, na katerih je deloval operacijski sistem VMS, so se združevali v korporacije in razdruževali, njihov tržni delež v svetovnem merilu je doživljal pretrese, konkurenca ni počivala, pa tudi cenovna politika se je spremenila, tako da nismo bili več deležni nekdanjih velikih ugodnosti. Največji udarec pa je predstavljal upad zagona pri razvoju operacijskega sistema VMS. Sledenje razvojnim trendom, novostim in strokovnim dognanjem je začelo močno zaostajati. Nove funkcionalnosti, ki so bile pri drugih operacijskih sistemih naravne, so pri operacijskem sistemu VMS delovale le z nesistematičnim dopolnjevanjem systemske programske opreme. Zelo težko je bilo slediti eksploziji novih komponent strojne opreme, ustrezno podpirati vedno nove večpredstavne vmesnike in uporabljati nove programerske pristope, vključno s programskimi jeziki. Tudi drugje po Sloveniji so se začele migracije informacijskih sistemov na druge platforme, predvsem na različice Unix in na Microsoftov Windows. Posledično je bilo tudi ustreznih strokovnjakov vse manj v naši bližini, programerske veščine za obvladovanje okolij,

povezanih z operacijskim sistemom VMS, pa se v slovenskih izobraževalnih procesih niso več negovale.

Prodor osebnih računalnikov je na smetišče tehnološke zgodovine odplaval alfanumerične terminale, ki niso več zadoščali niti za najosnovnejše posodobitve informacijskega okolja na delovnih mestih v knjižnicah. IZUM se je soočal z vedno pogostejšimi zahtevami po spremembah nekaterih osnovnih konceptov pri avtomatizaciji knjižničnih procesov in strokovnih izzivih. Pasivni odjemalci (terminali) niso omogočali aktivnejše vloge pri lokalnih obdelavah in načelo »vse je sprogramirano vnaprej« je tudi ob obilni parametrizaciji onemogočalo propulzivnejši razvoj. Nova generacija programske opreme COBISS (popularno imenovana COBISS3) je predvidevala take funkcionalnosti, ki jih je bilo v tistem trenutku z že zastarelim pristopom možno implementirati le z obilo napora.

Naprej ni šlo več. Posamezne komponente sistema COBISS so bile predelane na način, ki ni determiniral informacijskega okolja ne s sistemsko programsko opremo ne s strojno opremo. Na srečo je bila programska oprema sistema COBISS že od samega začetka zasnovana zelo modularno, tako da je IZUM zadnjih nekaj let lahko postopoma v posameznih segmentih opuščal operacijski sistem VMS in ga polagoma nadomeščal z drugimi koncepti. Lasten sistem za upravljanje baz podatkov je bil v določenih segmentih nadomeščen s komercialnim, čeprav smo se zavedali, da gre le za začasno premostitev najbolj perečih težav. Pri programiranju smo se vse bolj naslanjali na nove programerske principe in jezike, prilagodili smo interne standarde in vpeljali nove metodologije pri organizaciji, načrtovanju in testiranju. Rezultati niso izostali, vendar je operacijski sistem VMS še vedno kraljeval v ozadju.

IZUM se je zavedal, da bo prehod na novo informacijsko okolje nujen. Hkrati je bilo jasno, da se je nanj treba zelo dobro pripraviti. Najbolj kritično vprašanje je bilo, kakšen naj bo novi sistem, na katerem bo COBISS tekel vsaj enako dobro, obenem pa bodo ostala vrata dovolj odprta za več kot le kratkoročni nadaljnji razvoj. Ob nekajletnem obdobju analiziranja in načrtovanja, včasih stopicanja na mestu ali celo vrtenja v krogu, pa tudi vrnitve s poti, ki so se izkazale za slepe, in ponovnega poskusa preboja, je IZUM še vedno redno nadgrajeval obstoječi sistem COBISS3 in ga dopolnjeval. Slika, kaj hočemo, je bila dokaj jasna, pot do uspešne izvedbe pa zapletena.

## 5 Proces prehoda v sedanje okolje

Prehod iz operacijskega sistema VMS ni bil premočrten. Najprej so se novosti pojavile na strežniški strani, in sicer z uporabo strežnikov z operacijskima sistemoma Windows in Linux, nato so sledili aplikacijski strežniki platforme Java EE.

Za shranjevanje podatkov se je v zaledju, po neposrečeni<sup>2</sup> fazi implementacije posebnega modela organizacije podatkov, ki je nekje vmes med datotečnim sistemom in relacijskimi sistemi za upravljanje baz podatkov začela na strežnikih z operacijskim sistemom Linux uporabljati relacijsko bazo podatkov Oracle. Tudi na strani odjemalca so se z uporabo aplikacije na osnovi programskega jezika java lahko za delovna mesta v knjižnicah uporabljali osebni ali prenosni računalniki z različnimi operacijskimi sistemi (večinoma Windows), nekaj pa tudi Linux in MacOS.

---

<sup>2</sup> Osebno avtorjevo mnenje.

Prva segmenta programske opreme »po novem« sta bila COBISS3/Medknjižnična izposoja in COBISS3/Upravljanje aplikacij. Pri tem se je hitro pokazala nujnost gradnje ogrodja za razvoj aplikacij, ki bi omogočalo bistveno učinkovitejšo razvojno pot številnih razmeroma kompleksnih modulov med seboj zelo povezanih lokalnih aplikacij COBISS. Ogrodje je bilo razvito že v novem programskem okolju (java) in se je izkazalo za ključno prebojno orodje pri nadaljnjih programskih modulih in segmentih.

Pri načrtovanju prehoda programske opreme na drugačne osnove so prednost imeli segmenti, ki v obstoječem okolju še niso bili podprti. Najprej so bili na vrsti COBISS3/Nabava, COBISS3/Serijske publikacije in COBISS3/Zaloga s shranjevanjem zaloge, vzporedno z njimi pa COBISS3/Izpisi, ki so knjižničarjem omogočili predvsem parametrizirano pripravo obrazcev, njihovo povezovanje s podatki v bazah podatkov in pripravo izpisov v različnih izhodnih formatih.

Sledila sta najzahtevnejša segmenta: COBISS3/Katalogizacija in COBISS3/Izposoja. Razvoj nove generacije katalogizacije je bil zastavljen tako, da ni bilo postopnega prehoda, zato ga je bilo možno razmeroma hitro vključiti v produkcijo oziroma preklopiti iz starega načina dela v novi. Po uspešni implementaciji v IZUM-ovi knjižnici je leta 2012 nova katalogizacija uspešno stekla v Narodni in univerzitetni knjižnici. Posamezne knjižnice so jo osvajale postopoma po natančnem načrtu in ob polovici leta 2016 so vse knjižnice v Sloveniji svoje gradivo katalogizirale izključno v segmentu COBISS3/Katalogizacija.

Pri izposoji smo ubrali drugačno pot. Najprej smo podprli knjižnice, ki potrebujejo najmanj funkcionalnosti v postopku avtomatizirane izposoje, v kasnejših fazah pa smo vključevali vedno zahtevnejše funkcionalnosti za različne tipe knjižnic in za različno raven njihovih zahtev. Tako so šele na koncu prišle na vrsto splošnoizobraževalne knjižnice, ki so imele s stališča izposoje največje zahteve in so bile tudi najzahtevnejše za operativno izvedbo prehoda iz starega v novo okolje.

Potrebovali smo nove komponente za prikaz bibliografskih podatkov. Njihov razvoj je omogočil tudi realizacijo novih verzij bibliografij.

Vpeljevanje sofisticiranega sodobnega iskalnika je bil posebno trd oreh. Novi COBISS+, ki je razvit na platformi Java, ni zgolj nadomestil nekdanje uspešne in široko uporabljane spletne aplikacije COBISS/OPAC. Z mislijo na kasnejšo integracijo s ponudniki integralnih indeksov elektronskih virov smo za že obstoječe vsebine iskali načine za obogatitev zmožnosti z novimi paradigmi, kot so fasetno iskanje, urejanje po relevantnosti, iskanje po več virih hkrati, krnilnik, upoštevanje funkcionalnih zahtev za bibliografske zapise itd. Po intenzivnih testiranjih in lokalnih prototipnih različicah smo se odločili za nabavo odprtokodnega iskalnika Lucene/SOLR, ki je izpolnil vsa naša pričakovanja.

Pri spoštovanju vredni dolžini življenjskega cikla programske opreme COBISS, ki je tekla na operacijskem sistemu VMS, je nastala cela vrsta pomožnih modulov in orodij (različni izpisi, poročila, statistike ...), za katere je bilo treba najti ustrezno nadomestilo. Šele ko je bilo to urejeno, je bilo možno »pritisniti na gumb« in ugasniti operacijski sistem VMS.

Ostala sta še vprašanje sistema za upravljanje baz podatkov in generalna zamenjava operacijskega sistema. Pri bazah podatkov smo se odločili za zamenjavo relacijske baze Oracle

z odprtokodnim sistemom PostgreSQL. Produkcija za končne uporabnike se seli na operacijski sistem iz družine Unix; konkretno gre za CentOS (ena od implementacij Linuxa).

## 6 Sklepna misel

Operacijski sistem VMS je bil mnoga leta ena od temeljnih podlag računalniškega razvoja v Sloveniji. Generacijam študentov je bil eden od najpopolnejših vpogledov v principe delovanja računalnikov (izven ravni elektronike). Deloval je na računalniških sistemih univerz, raziskovalnih inštitutov in raziskovalnih organizacij, pa tudi v gospodarstvu ni bil redek. V predinternetni dobi je bil glavni povezovalni faktor neformalnega omreženja slovenskih računalnikov.

*Tabela 1: Razvojna pot osnovnih elementov podlag, na katerih stoji in deluje COBISS*

Komponenta	Na začetku	Potem	Danes
operacijski sistem	VMS	VMS; MS Windows	CentOS
baza podatkov	Total/Ultra/Supra	lasten sistem; Oracle	PostgreSQL
programski jezik	cobol; basic; zbirnik		java

Več kot 35 let je bil operacijski sistem VMS podlaga za razvoj in delovanje nacionalnega vzajemnega bibliografskega sistema COBISS. V zadnjih letih pa ni več zagotavljal zadostne podpore za nove generacije vsebinskih in računalniških zahtev in je postal resna ovira na nadaljnji razvojni poti. Te ovire sicer niso nepremostljive, zahtevajo pa ogromna finančna vlaganja, ki si jih ne IZUM ne knjižnična skupnost ne moreta privoščiti. Rešitev za spremembo se je iskala dolgo in odgovorno konservativno, pot do sprememb ni bila premočrna. Tabela 1 povzema v prispevku opisano razvojno pot komponent z začetnim, vmesnim in končnim stanjem.

## Zahvala

Na tem mestu bi se rad zahvalil nekdanjim in sedanjim sodelavcem in pionirjem dela v operacijskem sistemu VMS, ki so prebrali posamezne dele tega prispevka in me s svojimi pripombami in predlogi spodbujali, pa tudi obvarovali kakšne nedoslednosti. Če se nam je kljub vsemu kakšna izmuznila, gre na moj račun. Navajam po abecednem redu: Saša Bibič, Slavko Breznik, Domen Šetar, Bojan Štok, Robert Vehovec.

## Reference

Digital Equipment Corporation, 1981. *VAX Architecture Handbook*. Maynard, ZDA: Digital Equipment Corporation.

Baase, S., 1983. *VAX-11 Assembly Language Programming*. Englewood Cliffs, ZDA: Prentice-Hall Inc.

Goldenberg, R. in Kenah, L., 1991. *VMA/VMS Internals and Data Structures: version 5.2*, Woods Drive, ZDA: Digital Press.

Osebni arhiv avtorja, 1982–2020.

# COBIB.SI: prirast bibliografskih zapisov v letu 2019

## COBIB.SI: increment of bibliographic records in 2019

**Emma Dornik<sup>1</sup>**

**IZVLEČEK:** V Sloveniji je v vzajemno katalogizacijo vključena 301 knjižnica, sodeluje pa 689 oseb z dovoljenjem za vzajemno katalogizacijo monografskih publikacij. S člankom prikažemo prirast bibliografskih in normativnih zapisov v COBIB.SI za leto 2019. Podatke o prirastu zbiramo iz vzajemne baze čez opazovano leto. V analizo podatkov smo vključili vse dneve leta in vse zapise, ki jih je prispevalo 575 katalogizatorjev iz 254 knjižnic. Podatki o prirastu so predstavljeni na letni, mesečni in dnevni ravni. V letu 2019 so katalogizatorji v COBIB.SI prispevali 176.021 bibliografskih zapisov ter 60.068 normativnih zapisov za osebna imena in imena korporacij v bazi CONOR.SI. V zadnjem kvartalu leta je bilo kreiranih največ (28,7 %) bibliografskih zapisov. Največ bibliografskih zapisov v enem dnevu pa je bilo kreiranih v torek, 26. 11. 2019, in sicer 1.022. Torek je na letni ravni najproduktivnejši dan v tednu. Najmanj zapisov je bilo kreiranih junija in avgusta. Po tipu zapisov se nadaljuje večletni trend: največ zapisov je bilo kreiranih za monografske publikacije (51,2 %) in sestavne dele (38,1 %). Predstavljeni so tudi podatki o prevzetih zapisih iz drugih baz podatkov. V Sloveniji je 645 oseb, ki so v letu 2019 aktivno sodelovale pri kreiranju bibliografskih in normativnih zapisov. Glede na 2,3-odstotni delež letnega prirasta zunaj delovnih dni sklepamo, da obstajajo vse večje potrebe po katalogizatorjih. Dve ali več uporabniških imen ima 30,9 % katalogizatorjev, torej delujejo v več organizacijah. Predstavljeni rezultati analize prirasta COBIB.SI kažejo na pomembnost dela katalogizatorjev v sistemu COBISS.

**KLJUČNE BESEDE:** bibliografske baze podatkov, COBIB.SI, katalogizacija, letni prirast, 2019

**ABSTRACT:** In Slovenia, 301 libraries and 689 cataloguers with permits for shared cataloguing of monographs participate in shared cataloguing. The article presents the annual increment of bibliographic and authority records in COBIB.SI in 2019. The increment data is collected from the shared database during the entire observed year. All the days of the year and all the records contributed by 575 cataloguers from 254 libraries are included in the analysis. The annual, monthly and daily increment data is presented. In 2019, 176,021 bibliographic records were contributed to COBIB.SI, and 60,068 authority records for personal names and for corporate bodies were contributed to CONOR.SI. The largest number of bibliographic records (28.7%) was created in the last quarter of the year. On a daily basis, the number of created bibliographic records was largest on Tuesday, 26 November 2019, when 1,022 bibliographic records were created. At the annual level, Tuesdays are the most productive days. In June and August, the numbers of created records were the smallest. In view of the record type, the trend of the past several years continues: most records were created for monographs (51.2%) and component parts (38.1%). The data on bibliographic and authority records downloaded from other databases is presented, too. In 2019, 645 persons actively participated in the creation of bibliographic and authority records in Slovenia. As 2.3% of bibliographic records were created on work-free days, it can be concluded that cataloguers are more and more in demand; 30.9% of cataloguers have two or more usernames, meaning that they work in more than one organisation. The analysis results of COBIB.SI increment presented indicate the importance of cataloguing in the COBISS system.

**KEYWORDS:** bibliographic databases, COBIB.SI, cataloguing, annual increment, 2019

---

<sup>1</sup> Doc. dr. Emma Dornik, prof. zdr. vzg., bibliotekarska višja svetnica, Institut informacijskih znanosti, Maribor, Slovenija, [ema.dornik@izum.si](mailto:ema.dornik@izum.si).

## 1 Uvod

V mreži COBISS.net (Kooperativni online bibliografski sistem in servisi) sodeluje 1.397 knjižnic (podatek na dan 3. 8. 2020) iz osmih držav (v abecednem vrstnem redu: Albanija, Bolgarija, Bosna in Hercegovina, Črna gora, Kosovo, Makedonija, Slovenija in Srbija) (COBISS.net, 2020; COBISS.net v številkah, 2020). V Sloveniji je na dan 31. 12. 2019 sodelovalo 919 polnopravnih članic, ki jih delimo po tipih knjižnic: Narodna in univerzitetna knjižnica, 83 univerzitetnih in visokošolskih knjižnic, 134 specialnih, 62 splošnih in 639 šolskih knjižnic (Bibliografske baze podatkov in podatki o zalogi, 2020; IZUM, 2020). Gre za nacionalni knjižnični informacijski sistem z vzajemno katalogizacijo, vzajemno bibliografsko-kataložno bazo podatkov COBIB.SI in lokalnimi bazami podatkov sodelujočih knjižnic, bazo podatkov o knjižnicah COLIB.SI, normativno bazo podatkov CONOR.SI in s številnimi drugimi funkcijami (Organizacijski model sistema COBISS, 2020). Vzajemna katalogizacija omogoča racionalno delitev dela in prihrank pri zahtevnem postopku obdelave knjižničnega gradiva in vodenju katalogov (Vzajemna katalogizacija (COBISS3/Katalogizacija), 2020). V Sloveniji je v vzajemno katalogizacijo vključena 301 knjižnica (COBISS3 – Knjižnice, 2020). Glede na tip zapisa je bilo v letu 2019 (skupaj) v vzajemni katalog dodanih 51,2 % zapisov za monografske publikacije, 38,1 % za sestavne dele, 8,2 % za izvedena dela (dogodke), 2,2 % za serijske publikacije, 0,2 % za zbirne zapise in 0,2 % zapisov za integrirne vire. Iz COBIB.SI je bilo skupaj izbrisanih 1.117 bibliografskih zapisov (Bibliografske baze podatkov in podatki o zalogi, 2020). V Sloveniji imamo 689 oseb z dovoljenjem za vzajemno katalogizacijo monografskih publikacij, v državah, vključenih v COBISS.net, pa še 625.

S predstavitvijo prirasta bibliografskih zapisov v COBIB.SI želimo prikazati obseg in poudariti pomembnost katalogizatorjevega dela v sistemu COBISS. Letni prirast bibliografskih zapisov analiziramo že od leta 2014. Sistematično spremljanje prirasta nam daje vpogled v dinamičnost bibliografske baze in obremenjenost katalogizatorjev čez celo leto. Ugotovljene značilnosti omogočajo načrtovanje aktivnosti, ki so povezane s katalogizacijo.

## 2 Metodologija

Vir podatkov je vzajemna baza COBIB.SI, v kateri s pregledom pojmov za iskalno predpono za datum kreiranja/prezema (dm=) poiščemo podatke za pretekli mesec. Podatke zbiramo v začetku vsakega meseca za pretekli mesec. Vključeni so vsi dnevi v letu, saj se COBIB.SI ves čas spreminja in dopolnjuje. Za delovne dni smo šteli vse dni od ponedeljka do petka, za dela proste dni smo upoštevali sobote, nedelje in praznike. Vzajemna baza se dnevno spreminja, zato številke, vključene v analizo, ne predstavljajo trenutnega stanja podatkov (npr. števila zapisov) v bazi. Iz zajema podatkov so izvzeti bibliografski zapisi, ki so bili v vzajemni katalog shranjeni programsko (npr. ELINKS.SI,  $n = 2.522$ ), in lokalni zapisi, ki so bili shranjeni v vzajemno bazo. Vključeni so vsi kreirani in prevzeti bibliografski zapisi, ki so jih v vzajemni katalog prispevali katalogizatorji.

Zbrani podatki so predstavljeni na letni ali mesečni ravni. V analizo podatkov za leto 2019 smo vključili 176.021 bibliografskih zapisov. Na slikah so podatki predstavljeni po datumih, v tabelah so prazniki prikazani ločeno od preostalih dni. Ob slikah so zabeleženi dnevi zajema podatkov, veljajo pa tudi za podatke, predstavljene v tabelah mesečnih prirastov. Ker prirast sistematično spremljamo od leta 2014, smo vključili trend prirasta za obdobje 6 let (Dornik, 2015; 2016; 2017; 2018; 2019).

### 3 Prirast bibliografskih zapisov

Konec leta 2019 (podatki na dan 31. 12.) je bilo v vzajemni bazi podatkov COBIB.SI skupaj 5.431.373 bibliografskih zapisov (od tega 1.858.181 z vsaj enim avtorjem s šifro raziskovalca). V letu 2019 je aktivno sodelovalo 575 katalogizatorjev iz 254 knjižnic.

Katalogizatorjem je pri delu v pomoč možnost prevzemanja bibliografskih zapisov iz drugih baz podatkov. Iz mreže COBISS.net je bilo prevzetih 2.515 zapisov, iz baze WorldCat (OCLC) 9.527, iz Registra ISSN 3.236 zapisov (od tega 2.522 ob ažuriranju baze podatkov ELINKS.SI), iz Online kataloga Kongresne knjižnice (LC) pa 2.033 zapisov. Največja dodana vrednost zasnove vzajemnega kataloga je kooperativno sodelovanje pri kreiranju bibliografskih zapisov (Bibliografske baze podatkov in podatki o zalogi, 2020).

#### 3.1 Letni prirast

V letu 2019 je bilo največ zapisov kreiranih v zadnji četrtini leta ( $Q4 = 28,7\%$ ). V prvi četrtini leta ( $Q1$ ) je bilo kreiranih  $24,7\%$ , v drugi ( $Q2$ )  $24\%$ , v tretji ( $Q3$ ) pa  $22,5\%$  letnega prirasta.

Pregled prirasta bibliografskih zapisov po mesecih leta 2019 prikazuje tabela 1.

*Tabela 1: Mesečni prirast bibliografskih zapisov v COBIB.SI (padajoče glede na število zapisov)*

Mesec	Kreirani zapisi (n)	Kreirani zapisi (%)
november	17.601	10,0
oktober	17.399	9,9
januar	16.417	9,3
december	15.491	8,8
september	15.261	8,7
maj	15.128	8,6
julij	14.285	8,1
april	14.270	8,1
marec	13.982	7,9
februar	13.107	7,4
junij	12.935	7,3
avgust	10.145	5,8
Skupaj	176.021	100

*Legenda: n – število, % – odstotek*

Prirast bibliografskih zapisov na letni ravni (upoštevani so vsi dnevi v letu) kaže, da je bilo v letu 2019, enako kot v letu prej, največ zapisov kreiranih ob torkih. Podrobnosti so predstavljene v tabeli 2.

Tabela 2: Letni prirast bibliografskih zapisov v COBIB.SI (padajoče po dnevih v tednu glede na število kreiranih zapisov)

Dan	Štev. dni v letu (n)	Zapisi (n)	Zapisi ( $\bar{x}$ )	Zapisi (s)	Zapisi (%)	Zapisi (maks)	Zapisi (min)
ponedeljek	51	35.750	701,0	150,4	20,31	986	329
torek	51	37.365	732,6	156,7	21,23	1022	234
sreda	49	34.952	713,3	118,5	19,86	910	360
četrtek	48	33.930	706,9	102,6	19,28	904	424
petek	50	30.003	600,1	149,2	17,05	878	128
sobota	52	1.915	36,8	23,8	1,09	124	4
nedelja	52	1.794	34,5	21,9	1,02	87	3
praznik	12	312	26,0	11,9	0,18	46	6
<b>Skupaj</b>	<b>365</b>	<b>176.021</b>	<b>482,2</b>	<b>328,5</b>	<b>100</b>	<b>1.022</b>	<b>3</b>

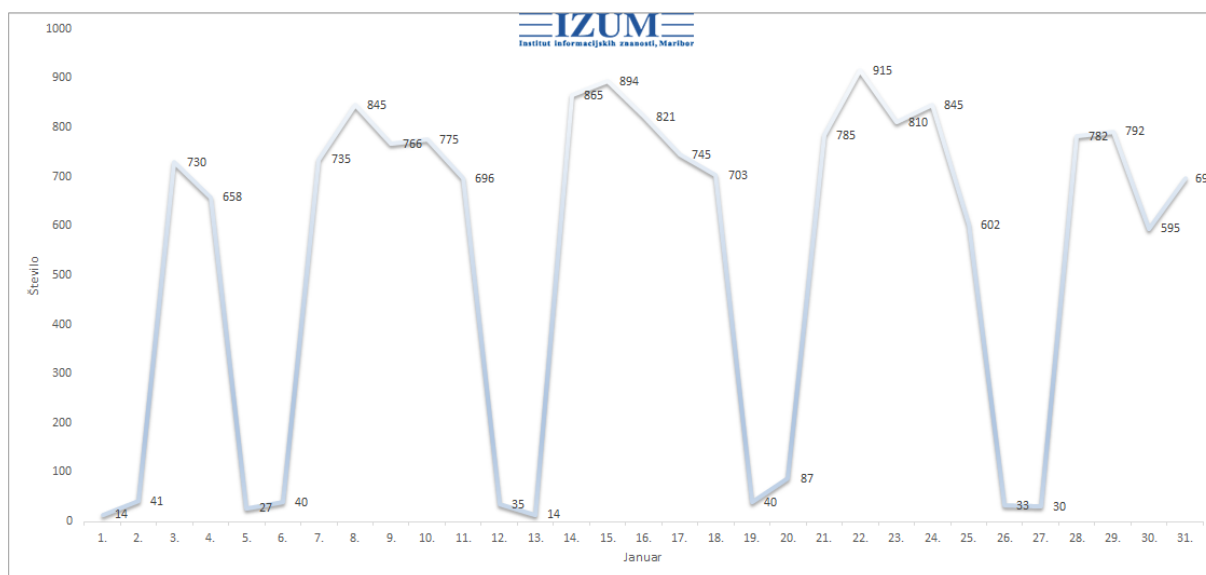
Legenda: n – število,  $\bar{x}$  – povprečje, s – standardni odklon, % – delež, maks – maksimum, min – minimum

### 3.2 Mesečni prirast

V nadaljevanju so predstavljeni podatki o kreiranih bibliografskih zapisih po posameznih mesecih leta 2019.

#### 3.2.1 Januar

V analizo smo vključili 16.417 bibliografskih zapisov. Januarja je bilo največ bibliografskih zapisov kreiranih v torek, 22. 1. 2019, in sicer 915 (slika 1). Glede na dan v tednu je bilo januarja največ zapisov kreiranih v 5 četrtkih ( $n = 3.792$  ali 23,1 %; maks = 845; s = 56). Povprečni dnevni prirast je v 21 delovnih dneh ( $\bar{x}$ ) znašal 764,6 zapisa, standardni odklon (s) pa 87. V delovnih dneh januarja je bilo na dan kreiranih najmanj 595 bibliografskih zapisov.



Slika 1: Prirast bibliografskih zapisov – januar 2019 (podatki na dan 1. 2. 2019)

Povprečni dnevni prirast števila bibliografskih zapisov v 10 dela prostih in prazničnih dneh ( $\bar{x}$ ) je znašal 36,1 zapisa, odklon (s) pa 20,4 zapisa ( $n = 361$ ). Na mesečni ravni to predstavlja 2,1 % prirasta. V dela prostih dneh je bilo kreiranih najmanj 14 in največ 78 zapisov na dan. Zapisi,



kreirani januarja, predstavljajo 9,3 % celoletnega prirasta. Januar po številu kreiranih zapisov na letni ravni zaseda 3. mesto. Podrobnosti prirasta v januarju so predstavljene v tabeli 3 (podatki so razvrščeni glede na rang).

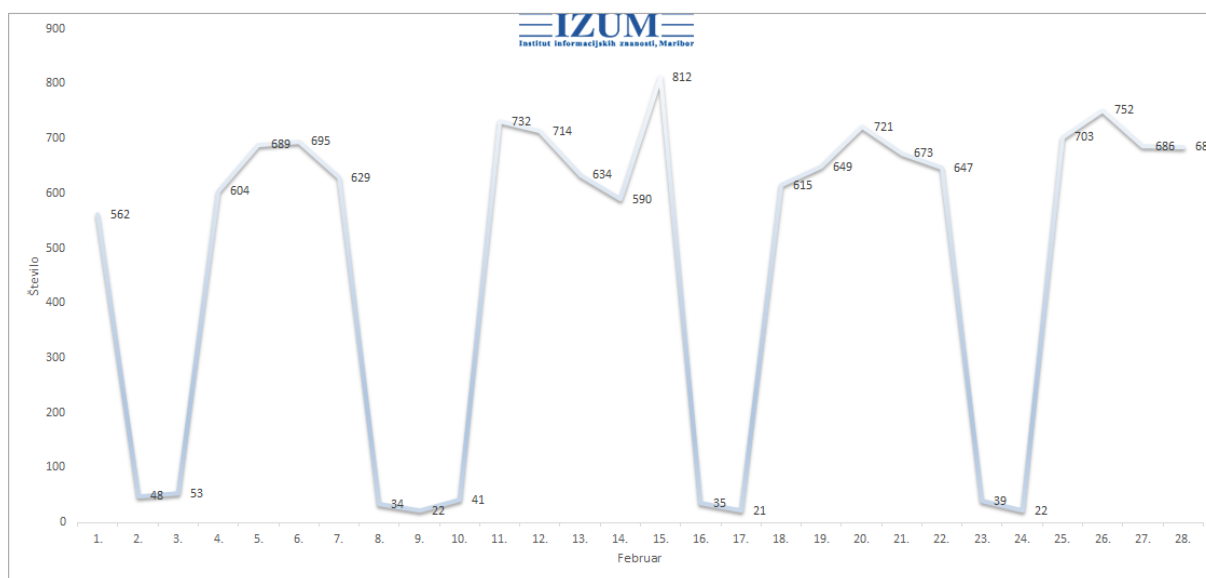
Tabela 3: Bibliografski zapisi, kreirani januarja 2019

Dan	Štev. dni v mescu ( <i>n</i> )	Zapisi ( <i>n</i> )	Zapisi ( $\bar{x}$ )	Zapisi ( <i>s</i> )	Zapisi (%)	Zapisi ( <i>maks</i> )	Zapisi ( <i>min</i> )	<i>R</i>
četrtek	5	3.792	758,4	56,0	23,1	845	697	8
torek	4	3.446	861,5	54,8	21,0	915	792	7
ponedeljek	4	3.167	791,8	53,9	19,3	865	735	6
sreda	4	2.992	748,0	104,7	18,2	821	595	5
petek	4	2.659	664,8	46,3	16,2	703	602	4
nedelja	4	171	42,8	31,4	1,0	87	14	3
sobota	4	135	33,8	5,4	0,8	40	27	2
praznik	2	55	27,5	19,1	0,3	41	14	1
<b>Skupaj</b>	<b>31</b>	<b>16.417</b>	<b>529,6</b>	<b>353,5</b>	<b>100</b>	<b>915</b>	<b>14</b>	<b>/</b>

Legenda: *n* – število,  $\bar{x}$  – povprečje, *s* – standardni odklon, % – delež, *maks* – maksimum, *min* – minimum, *R* – absolutni rang

### 3.2.2 Februar

V analizo smo vključili 13.107 zapisov, kreiranih v mesecu februarju (slika 2). V 19 delovnih dneh je bilo kreiranih povprečno ( $\bar{x}$ ) 673 zapisov na dan, odklon (*s*) pa je znašal 61. V 9 dela prostih dnevih in prazniku je bilo kreiranih 315 bibliografskih zapisov, kar na mesečni ravni predstavlja 2,4-odstotni prirast.



Slika 2: Prirast bibliografskih zapisov – februar 2019 (podatki na dan 1. 3. 2019)

Zapisi, kreirani v februarju, predstavljajo 7,4 % letnega prirasta. Februar je na desetem mestu glede na delež letnega prirasta. Največ zapisov na posamezen dan je bilo kreiranih v petek, 15. 2. 2019, in to 812, sicer pa so bili torki najbolj produktivni dnevi v delovnem tednu. Štirje

delovni torki v mesecu so prinesli 21,4 % mesečnega prirasta. Celoten mesečni prirast je predstavljen v tabeli 4 (podatki so razvrščeni glede na rang).

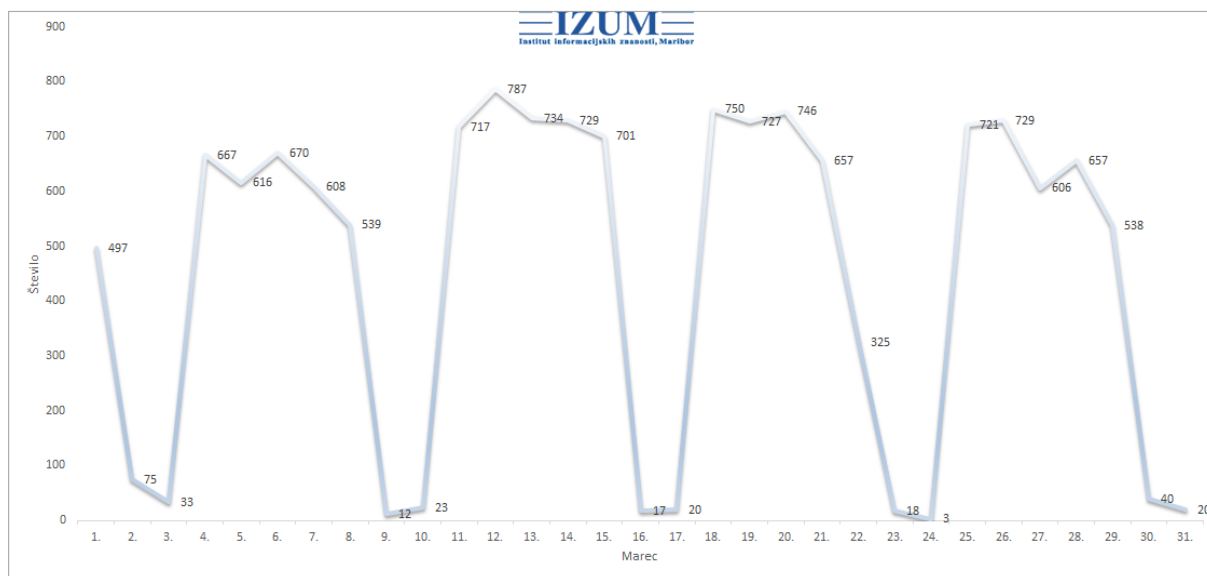
Tabela 4: Bibliografski zapisi, kreirani februarja 2019

Dan	Štev. dni v mescu (n)	Zapisi (n)	Zapisi ( $\bar{x}$ )	Zapisi (s)	Zapisi (%)	Zapisi (maks)	Zapisi (min)	R
torek	4	2.804	701,0	43,3	21,4	752	649	8
sreda	4	2.736	684,0	36,5	20,9	721	634	7
ponedeljek	4	2.654	663,5	63,6	20,2	732	604	6
četrtek	4	2.577	644,3	43,4	19,7	685	590	5
petek	3	2.021	673,7	127,1	15,4	812	562	4
sobota	4	144	36,0	10,8	1,1	48	22	3
nedelja	4	137	34,3	15,5	1,0	53	21	2
praznik	1	34	34,0		0,3	34	34	1
<b>Skupaj</b>	<b>28</b>	<b>13.107</b>	<b>468,1</b>	<b>307,7</b>	<b>100</b>	<b>812</b>	<b>21</b>	<b>/</b>

Legenda: n – število,  $\bar{x}$  – povprečje, s – standardni odklon, % – delež, maks – maksimum, min – minimum, R – absolutni rang

### 3.2.3 Marec

Marca je bilo kreiranih 13.982 bibliografskih zapisov, kar predstavlja 7,9-odstotni letni prirast, tako da je marec na 9. mestu glede na letoletni prirast. V 21 delovnih dneh so katalogizatorji prispevali 13.721 bibliografskih zapisov ( $\bar{x} = 653,4$ ;  $s = 108,4$ ;  $maks = 787$ ;  $min = 325$ ), v 10 dela prostih dneh pa še dodatnih 261 bibliografskih zapisov ( $\bar{x} = 26,1$ ;  $s = 20$ ;  $maks = 75$ ;  $min = 4$ ) (slika 3).



Slika 3: Prirast bibliografskih zapisov – marec 2019 (podatki na dan 1. 4. 2019)

Glede na delovne dneve v tednu so bili v marcu najbolj produktivni torki, v štirih torkih je bilo kreiranih 2.859 bibliografskih zapisov. Največ zapisov na posamezen dan meseca je bilo prav tako kreiranih v torek, in sicer 12. 3. 2019, ko je bilo kreiranih 787 zapisov. Podrobnosti mesečnega prirasta so predstavljene v tabeli 5 (podatki so razvrščeni padajoče glede na število

zapisov po dnevih). Čeprav je bilo marca med delovnimi dnevi največ petkov, in sicer 5, pa je petek dan, ko je bilo skupaj kreiranih najmanj zapisov ( $n = 2.600$ ).

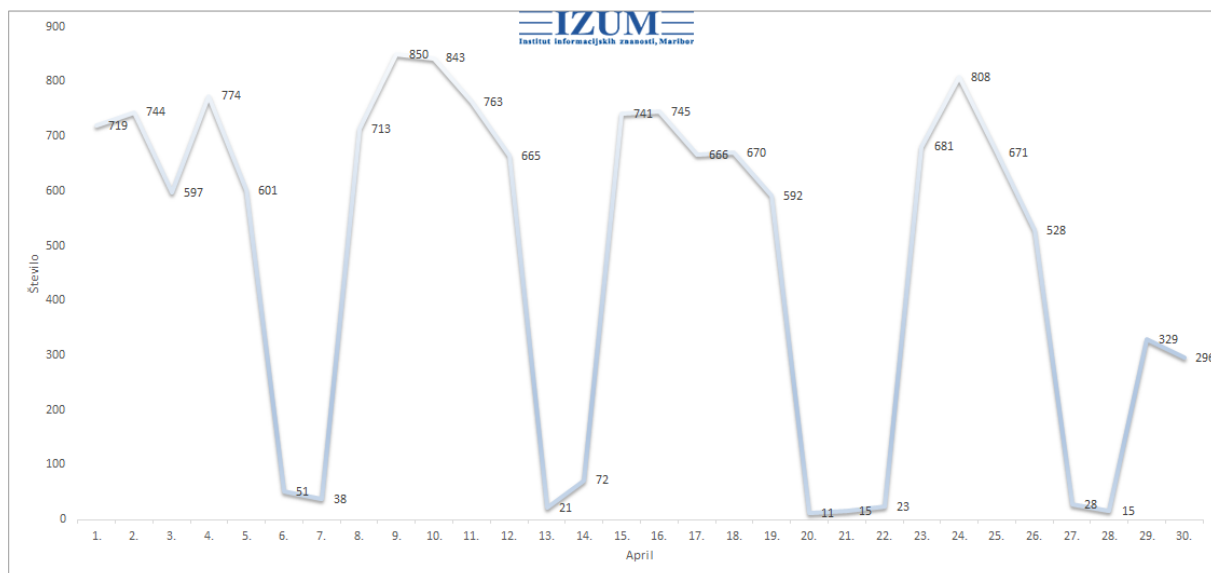
Tabela 5: Bibliografski zapisi, kreirani marca 2019

Dan	Štev. dni v mescu ( $n$ )	Zapisi ( $n$ )	Zapisi ( $\bar{x}$ )	Zapisi ( $s$ )	Zapisi (%)	Zapisi ( $maks$ )	Zapisi ( $min$ )	$R$
torek	4	2.859	714,8	71,5	20,4	787	616	7
ponedeljek	4	2.855	713,8	34,5	20,4	750	667	6
sreda	4	2.756	689,0	64,6	19,7	746	606	5
četrtek	4	2.651	662,8	49,8	19,0	729	608	4
petek	5	2.600	520,0	134,2	18,6	701	325	3
sobota	5	162	32,4	26,1	1,2	75	12	2
nedelja	5	99	19,8	10,8	0,7	33	3	1
Skupaj	31	13.982	451,0	311,1	100	787	3	/

Legenda:  $n$  – število,  $\bar{x}$  – povprečje,  $s$  – standardni odklon, % – delež, maks – maksimum, min – minimum,  $R$  – absolutni rang

### 3.2.4 April

Aprila se je baza COBIB.SI povečala za 14.270 bibliografskih zapisov (slika 4). Največ bibliografskih zapisov, in sicer 850, je bilo kreiranih v torek, 9. 4. 2019. Prav tako ob torkih beležimo največji skupni prirast bibliografskih zapisov v delovnih dneh. V 21 delovnih dnevih je povprečni dnevni prirast števila bibliografskih zapisov ( $\bar{x}$ ) znašal 666,5 zapisa, odklon ( $s$ ) je bil 143,8.



Slika 4: Prirast bibliografskih zapisov – april 2019 (podatki na dan 6. 5. 2019)

Podrobnosti prirasta v mesecu aprilu so predstavljene v tabeli 6 (podatki so razvrščeni glede na rang).

Tabela 6: Bibliografski zapisi, kreirani aprila 2019

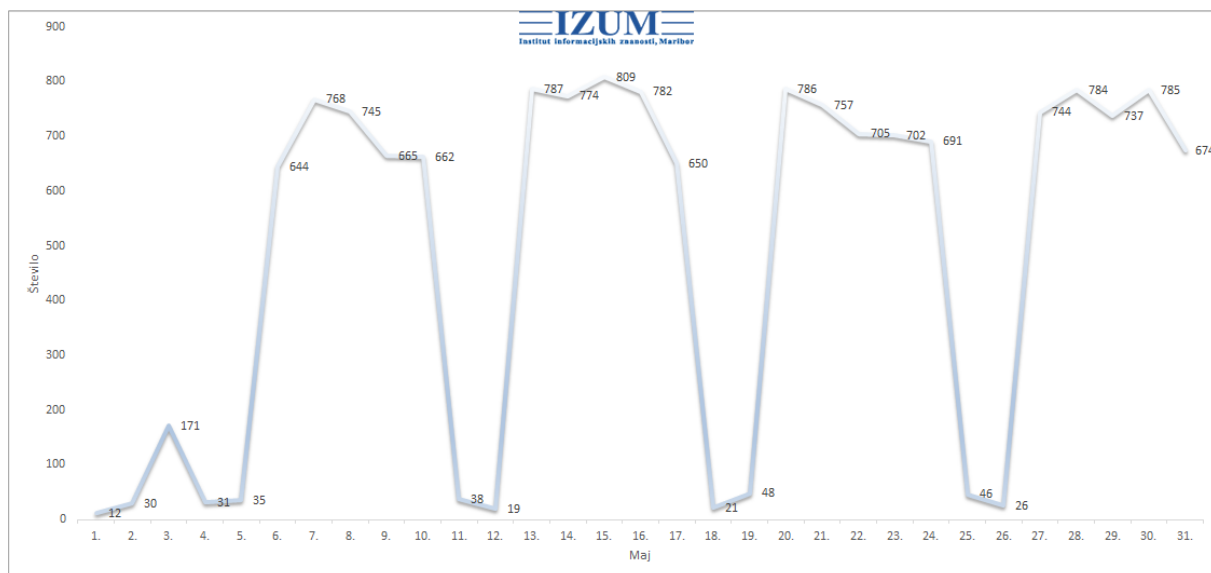
Dan	Štev. dni v mescu (n)	Zapisi (n)	Zapisi ( $\bar{x}$ )	Zapisi (s)	Zapisi (%)	Zapisi (maks)	Zapisi (min)	R
torek	5	3.316	663,2	214,0	23,2	850	296	8
sreda	4	2.914	728,5	116,4	20,4	843	597	7
četrtek	4	2.878	719,5	56,8	20,2	774	670	6
ponedeljek	4	2.502	625,5	198,0	17,5	741	329	5
petek	4	2.386	596,5	56,1	16,7	665	528	4
nedelja	4	140	35,0	26,9	1,0	72	15	3
sobota	4	111	27,8	17,0	0,8	51	11	2
praznik	1	23	23,0	/	0,2	23	23	1
<b>Skupaj</b>	<b>30</b>	<b>14.270</b>	<b>475,7</b>	<b>319,8</b>	<b>100</b>	<b>850</b>	<b>11</b>	<b>/</b>

Legenda: n – število,  $\bar{x}$  – povprečje, s – standardni odklon, % – delež, maks – maksimum, min – minimum, R – absolutni rang

Aprila je bilo 9 dela prostih dni, za katere beležimo 274 kreiranih zapisov ( $\bar{x}$  = 30,4; s = 20). Na letni ravni predstavljajo bibliografski zapisi, kreirani aprila, 8,1 %, tako da je med meseci v letu 2019 april na 8. mestu po številu kreiranih zapisov.

### 3.2.5 Maj

Maja 2019 je bilo kreiranih 15.128 zapisov, kar predstavlja 8,6-odstotni letni prirast bibliografskih zapisov. Maj je tako na 6. mestu glede na letni prirast. Najvišji prirast tega meseca beležimo v sredo, 15. 5., in sicer 809, sledi ponedeljek, 13. 5., s 787 kreiranimi bibliografskimi zapisi. Glede na dneve v tednu so za skupni prirast najbolj zaslužni zapisi, kreirani ob torkih; v štirih torkih v maju je bilo kreiranih 3.083 zapisov (slika 5).



Slika 5: Prirast bibliografskih zapisov – maj 2019 (podatki na dan 3. 6. 2019)

V 21 delovnih dneh je bilo kreiranih 14.822 bibliografskih zapisov. Glede na delovni dan je bilo največ zapisov kreiranih ob torkih (20,8 %), sledijo štiri srede (20,2 %) in štirje ponedeljki (20 %). V 10 dela prostih dneh so katalogizatorji v COBIB.SI dodali 306 bibliografskih zapisov

( $\bar{x}$  = 30,6;  $s$  = 11,6). Podrobnosti za vse dni v mesecu so predstavljene v tabeli 7, dnevi so razvrščeni glede na rang.

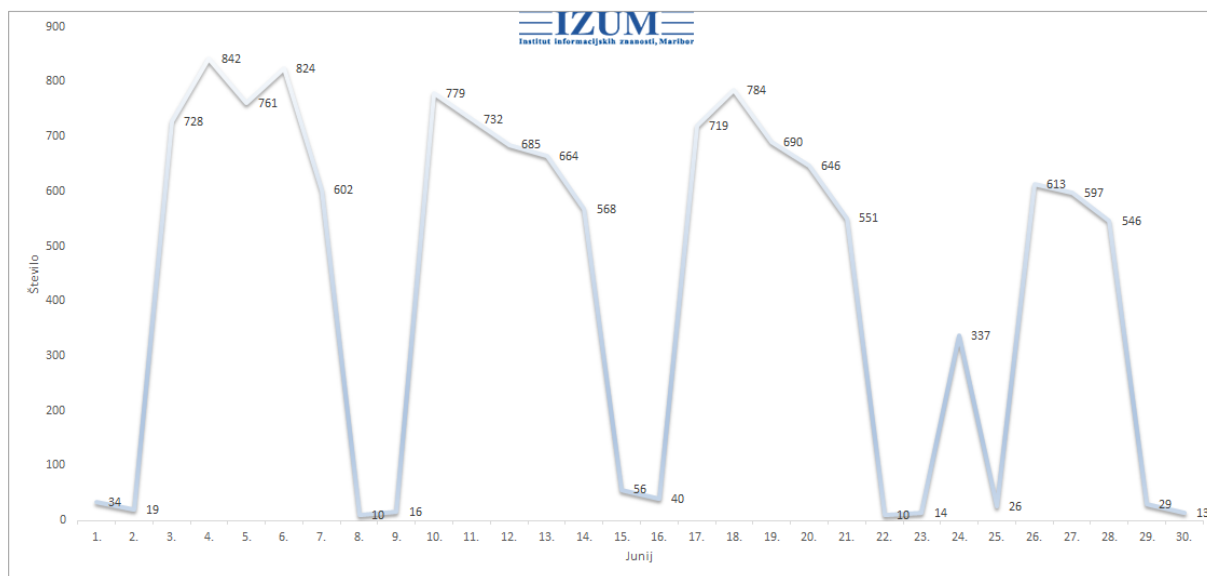
Tabela 7: Bibliografski zapisi, kreirani maja 2019

Dan	Štev. dni v mescu ( $n$ )	Zapisi ( $n$ )	Zapisi ( $\bar{x}$ )	Zapisi ( $s$ )	Zapisi (%)	Zapisi ( $maks$ )	Zapisi ( $min$ )	$R$
torek	4	3.083	770,8	11,3	20,4	784	757	8
sreda	4	2.996	749,0	43,6	19,8	809	705	7
ponedeljek	4	2.961	740,3	67,2	19,6	787	644	6
četrtek	4	2.934	733,5	59,7	19,4	785	665	5
petek	5	2.848	569,6	223,3	18,8	691	171	4
sobota	4	136	34,0	10,6	0,9	46	21	3
nedelja	4	128	32,0	12,5	0,8	48	19	2
praznik	2	42	21,0	12,7	0,3	30	12	1
<b>Skupaj</b>	<b>31</b>	<b>15.128</b>	<b>488,0</b>	<b>338,8</b>	<b>100</b>	<b>809</b>	<b>12</b>	<b>/</b>

Legenda:  $n$  – število,  $\bar{x}$  – povprečje,  $s$  – standardni odklon, % – delež, maks – maksimum, min – minimum,  $R$  – absolutni rang

### 3.2.6 Junij

Mesečni prirast junija predstavlja 7,3 %, tako da je junij na 11. mestu po številu prirasta bibliografskih zapisov na letni ravni. V 30 dneh junija je bilo kreiranih 12.935 bibliografskih zapisov. Glede na dneve v tednu je bilo junija (kumulativno po dnevih) največ zapisov kreiranih ob sredah, in sicer 2.749. Glede na posamezen dan v mesecu pa je bilo največ zapisov kreiranih v torek, 4. 6. 2019, in sicer 842 (slika 6).



Slika 6: Prirast bibliografskih zapisov – junij 2019 (podatki na dan 1. 7. 2019)

V 19 delovnih dneh je bilo kreiranih 12.668 bibliografskih zapisov, na delovni dan je povprečni prirast ( $\bar{x}$ ) znašal 666,7, odklon ( $s$ ) pa 120,4 zapisa. Zapisi, kreirani v 11 dela prostih dneh (petih sobotah, petih nedeljah in enem prazničnem dnevu), predstavljajo 2,1-odstotni delež prirasta

v juniju ( $n = 267$ ;  $\bar{x} = 24,3$ ;  $s = 14,5$ ). Kumulativni podatki za vse dni v juniju so predstavljeni v tabeli 8 (podatki so razvrščeni po dneh).

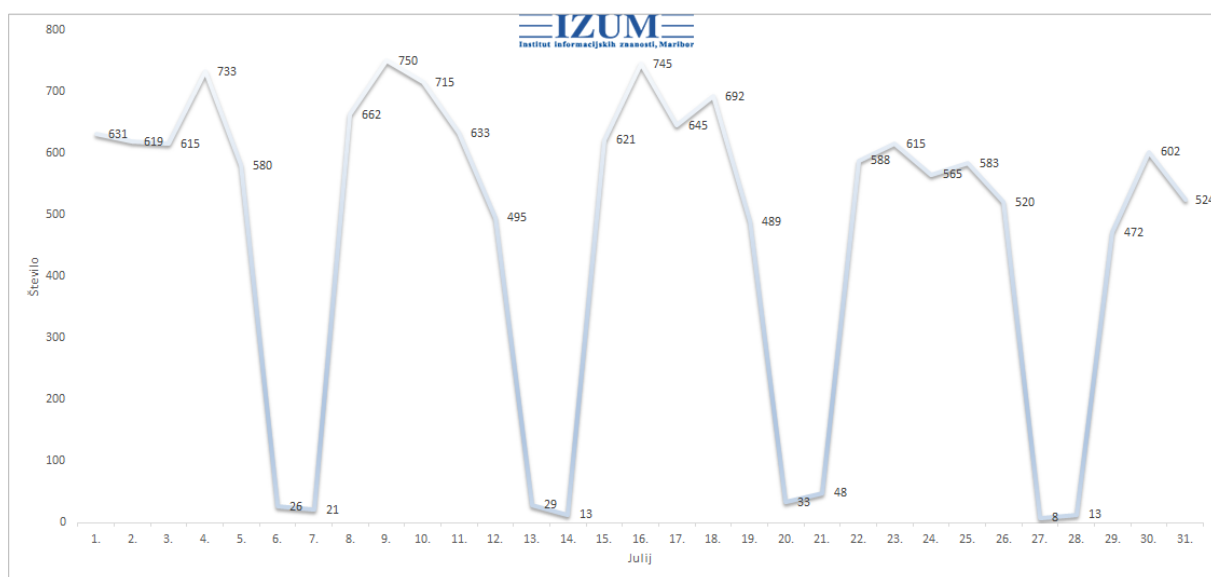
Tabela 8: Bibliografski zapisi, kreirani junija 2019

Dan	Štev. dni v mescu ( $n$ )	Zapisi ( $n$ )	Zapisi ( $\bar{x}$ )	Zapisi ( $s$ )	Zapisi (%)	Zapisi ( $maks$ )	Zapisi ( $min$ )	$R$
ponedeljek	4	2.563	640,8	204,2	19,8	779	337	6
torek	3	2.358	786,0	55,0	18,2	842	732	5
sreda	4	2.749	687,3	60,5	21,3	761	613	8
četrtek	4	2.731	682,8	98,3	21,1	824	597	7
petek	4	2.267	566,8	25,3	17,5	602	546	4
sobota	5	139	27,8	19,2	1,1	56	10	3
nedelja	5	102	20,4	11,2	0,8	40	13	2
praznik	1	26	26,0	/	0,2	26	26	1
<b>Skupaj</b>	<b>30</b>	<b>12.935</b>	<b>431,2</b>	<b>329,0</b>	<b>100</b>	<b>842</b>	<b>10</b>	<b>/</b>

Legenda:  $n$  – število,  $\bar{x}$  – povprečje,  $s$  – standardni odklon, % – delež, maks – maksimum, min – minimum,  $R$  – absolutni rang

### 3.2.7 Julij

Prirast bibliografskih zapisov v COBIB.SI obsega 14.285 zapisov (slika 7). Največ zapisov v tem mesecu je bilo kreiranih v torek, 9. 7. 2019, in sicer 750.



Slika 7: Prirast bibliografskih zapisov – julij 2019 (podatki na dan 1. 8. 2019)

Na mesečni ravni (zajeti so vsi dnevi) so katalogizatorji kreirali največ zapisov ob torkih, in sicer 23,3 % ( $n = 3.331$ ;  $\bar{x} = 666,2$ ;  $s = 74,5$ ). Torkom sledijo srede (tabela 9, podatki so razvrščeni glede na rang).

Tabela 9: Bibliografski zapisi, kreirani julija 2019

Dan	Štev. dni v mescu ( <i>n</i> )	Zapisi ( <i>n</i> )	Zapisi ( $\bar{x}$ )	Zapisi ( <i>s</i> )	Zapisi (%)	Zapisi ( <i>maks</i> )	Zapisi ( <i>min</i> )	<i>R</i>
torek	5	3.331	666,2	74,5	23,3	750	602	7
sreda	5	3.064	612,8	73,6	21,4	715	524	6
ponedeljek	5	2.974	594,8	73,6	20,8	662	472	5
četrtek	4	2.641	660,3	65,9	18,5	733	583	4
petek	4	2.084	521,0	41,6	14,6	580	489	3
sobota	4	96	24,0	11,0	0,7	33	8	2
nedelja	4	95	23,8	16,6	0,7	48	13	1
<b>Skupaj</b>	<b>31</b>	<b>14.285</b>	<b>460,8</b>	<b>270,9</b>	<b>100</b>	<b>750</b>	<b>8</b>	<b>/</b>

Legenda: *n* – število,  $\bar{x}$  – povprečje, *s* – standardni odklon, % – delež, *maks* – maksimum, *min* – minimum, *R* – absolutni rang

Bibliografski zapisi, kreirani v tem mesecu, predstavljajo 8,1 % letnega prirasta (enako kot leto prej), tako je julij, čeprav je poletni mesec, ko je čas letnih dopustov, na 7. mestu glede na letni prirast. Povprečni dnevni prirast števila bibliografskih zapisov ( $\bar{x}$ ) v 23 delovnih dneh je bil 612,8, odklon (*s*) pa 80,4 (*n* = 14.094) zapisa. V 8 dela prostih dneh je bilo kreiranih 191 zapisov (1,4 %), povprečni dnevni prirast ( $\bar{x}$ ) v dela prostih dneh je znašal 23,9, standardni odklon (*s*) pa 13,1.

### 3.2.8 Avgust

Avgust je, pričakovano, po prirastu bibliografskih zapisov na letni ravni na zadnjem mestu z 10.145 kreiranimi zapisi ali 5,8-odstotnim letnim prirastom. V 21 delovnih dneh je bil povprečni dnevni prirast ( $\bar{x}$ ) 477 zapisov, odklon (*s*) je bil 93,7; skupno število kreiranih zapisov v delovnih dneh je znašalo 10.026. V 10 dela prostih dneh je bilo kreiranih 119 bibliografskih zapisov, povprečni dnevni prirast gradiva ( $\bar{x}$ ) je znašal 11,9, odklon (*s*) pa 6,4 zapisa. V prikaz podatkov v tabeli 10 so vključeni vsi dnevi meseca, podatki so razvrščeni po dnevih.

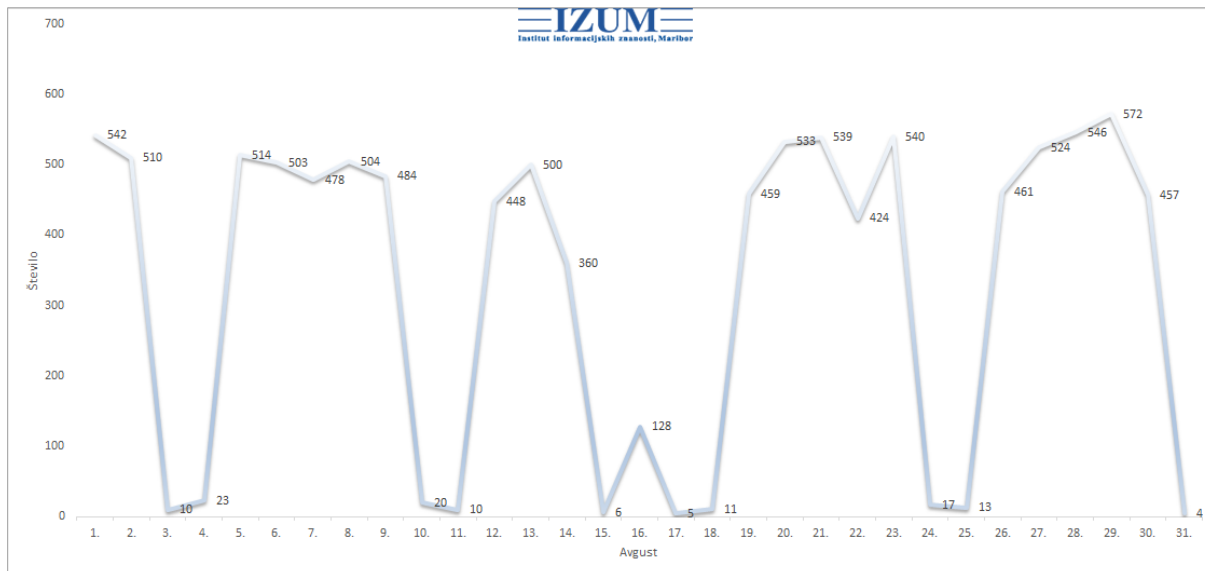
Tabela 10: Bibliografski zapisi, kreirani v mesecu avgustu 2019

Dan	Štev. dni v mescu ( <i>n</i> )	Zapisi ( <i>n</i> )	Zapisi ( $\bar{x}$ )	Zapisi ( <i>s</i> )	Zapisi (%)	Zapisi ( <i>maks</i> )	Zapisi ( <i>min</i> )	<i>R</i>
ponedeljek	4	1.882	470,5	29,6	18,6	514	448	4
torek	4	2.060	515,0	16,1	20,3	533	500	7
sreda	4	1.923	480,8	86,1	19,0	546	360	5
četrtek	4	2.042	510,5	64,0	20,1	572	424	6
petek	5	2.119	423,8	168,2	20,9	540	128	8
sobota	5	56	11,2	7,1	0,6	20	4	2
nedelja	4	57	14,3	6,0	0,6	23	10	3
praznik	1	6	6,0	/	0,1	6	6	1
<b>Skupaj</b>	<b>31</b>	<b>10.145</b>	<b>327,3</b>	<b>234,1</b>	<b>100</b>	<b>572</b>	<b>4</b>	<b>/</b>

Legenda: *n* – število,  $\bar{x}$  – povprečje, *s* – standardni odklon, % – delež, *maks* – maksimum, *min* – minimum, *R* – absolutni rang

Poleg tega, da je avgust na zadnjem mestu glede na število mesečnega prirasta, je zanimivo, da so bili najbolj produktivni delovni dnevi petki. Pet petkov je zaslužnih za 2.119 novo

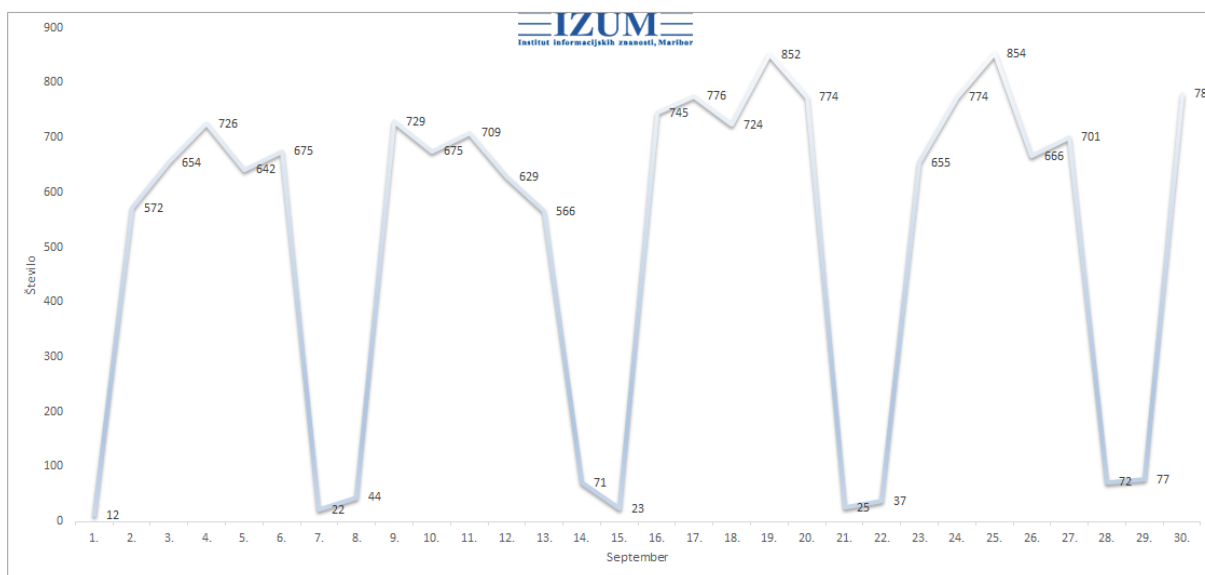
kreiranih zapisov, kar na mesečni ravni predstavlja 20,9 % ( $\bar{x} = 423,8$ ;  $s = 168,2$ ). Gledano na posamezen dan v mesecu, so katalogizatorji največ zapisov v COBIB.SI prispevali v četrtek, 29. 8. 2019, in sicer 572, sledita sreda, 28. 8. 2019, s 546 in spet četrtek, 1. 8. 2019, s 542 zapisi (slika 8).



Slika 8: Prirast bibliografskih zapisov – avgust 2019 (podatki na dan 2. 9. 2019)

### 3.2.9 September

September je na 5. mestu po številu kreiranih zapisov na letni ravni (8,7 %;  $n = 15.261$ ). Največ zapisov v tem mesecu je bilo kreiranih v sredo, 25. 9. 2019, in sicer 854, sledita četrtek, 19. 9. 2019, z 852 in ponedeljek, 30. 9. 2019, s 780 kreiranimi bibliografskimi zapisi (slika 9).



Slika 9: Prirast bibliografskih zapisov – september 2019 (podatki na dan 1. 10. 2019)

Podrobni podatki o mesečnem prirastu za vse dni septembra so predstavljeni v tabeli 11. Podatki so razvrščeni po dneh.



Tabela 11: Bibliografski zapisi, kreirani v mesecu septembru 2019

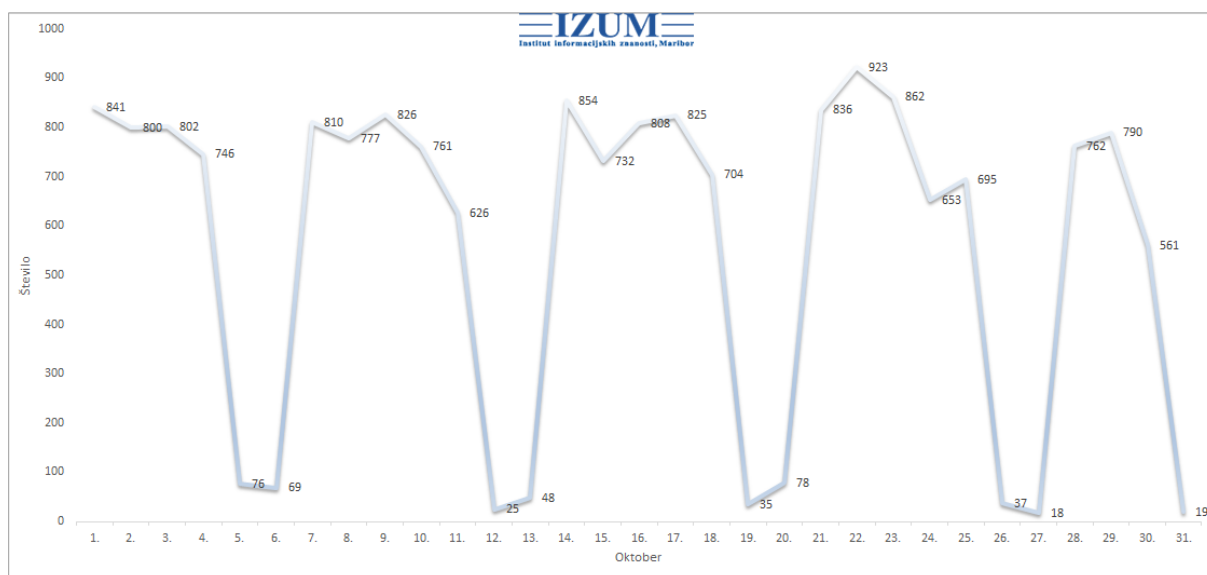
Dan	Štev. dni v mescu (n)	Zapisi (n)	Zapisi ( $\bar{x}$ )	Zapisi (s)	Zapisi (%)	Zapisi (maks)	Zapisi (min)	R
ponedeljek	5	3.481	696,2	83,1	22,8	780	572	7
torek	4	2.879	719,8	64,4	18,9	776	654	5
sreda	4	3.013	753,3	67,6	19,7	854	709	6
četrtek	4	2.789	697,3	104,3	18,3	852	629	4
petek	4	2.716	679,0	86,2	17,8	774	566	3
sobota	4	190	47,5	27,7	1,2	72	22	1
nedelja	5	193	38,6	24,8	1,3	77	12	2
<b>Skupaj</b>	<b>30</b>	<b>15.261</b>	<b>508,7</b>	<b>317,3</b>	<b>100</b>	<b>854</b>	<b>12</b>	<b>/</b>

Legenda: n – število,  $\bar{x}$  – povprečje, s – standardni odklon, % – delež, maks – maksimum, min – minimum, R – absolutni rang

V 21 delovnih dneh je bilo pet ponedeljkov, ko je bilo kreiranih 3.481 zapisov (23,4 %), povprečno 696,2 zapisa ( $s = 83,1$ ). Povprečje mesečnega prirasta v COBIB.SI v delovnih dneh septembra je 708,5 zapisa ( $s = 78,1$ ). Sobote in nedelje v septembru ( $n = 9$ ) so prirastu dodale 383 zapisov ali 2,5 % na mesečni ravni.

### 3.2.10 Oktober

Oktober 2019 je s 17.399 kreiranimi bibliografskimi zapisi ali 9,9-odstotnim letnim prirastom na drugem mestu po številu kreiranih zapisov. Katalogizatorji so v 22 delovnih dnevih kreirali 16.994 zapisov ( $\bar{x} = 772,5$ ;  $s = 84,5$ ). V oktobru je bilo 5 torkov in 5 sred, zapisi, kreirani v teh dneh, skupaj predstavljajo 46,6 % bibliografskih zapisov, kreiranih v oktobru. Največ zapisov na posamezen dan je bilo kreiranih v torek, 22. 10. 2019, in sicer 923, sledi sredi, 23. 10. 2019, z 862 bibliografskimi zapisi. V 9 dela prostih dnevih je bilo kreiranih 405 zapisov ( $\bar{x} = 45$ ;  $s = 24$ ), kar znaša 2,3 % mesečnega prirasta (slika 10).



Slika 10: Prirast bibliografskih zapisov – oktober 2019 (podatki na dan 4. 11. 2019)

Podrobnosti prirasta v mesecu (zajeti so vsi dnevi) so predstavljene v tabeli 12, podatki so razvrščeni po rangi.

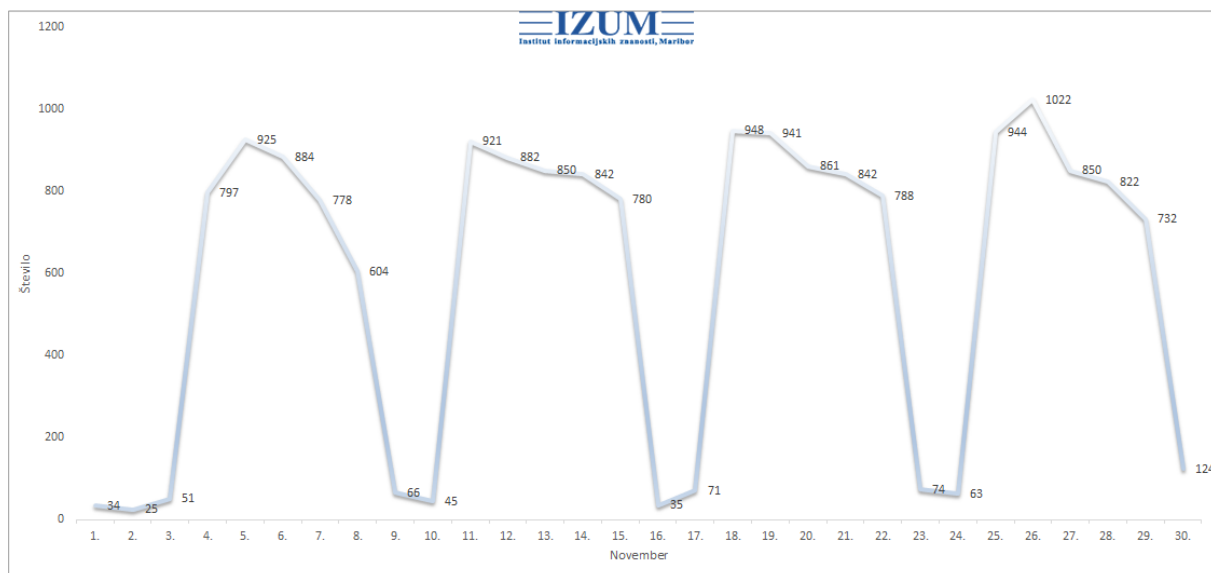
Tabela 12: Bibliografski zapisi, kreirani oktobra 2019

Dan	Štev. dni v mesecu (n)	Zapisi (n)	Zapisi ( $\bar{x}$ )	Zapisi (s)	Zapisi (%)	Zapisi (maks)	Zapisi (min)	R
torek	5	4.063	812,6	72,9	23,4	923	732	8
sreda	5	3.857	771,4	120,0	22,2	862	561	7
ponedeljek	4	3.262	815,5	40,0	18,7	854	762	6
četrtek	4	3.041	760,3	76,2	17,5	825	653	5
petek	4	2.771	692,8	49,7	15,9	746	626	4
nedelja	4	213	53,3	26,7	1,2	78	18	3
sobota	4	173	43,3	22,5	1,0	76	25	2
praznik	1	19	19,0	/	0,1	19	19	1
<b>Skupaj</b>	<b>31</b>	<b>17.399</b>	<b>561,3</b>	<b>343,2</b>	<b>100</b>	<b>923</b>	<b>18</b>	<b>/</b>

Legenda: n – število,  $\bar{x}$  – povprečje, s – standardni odklon, % – delež, maks – maksimum, min – minimum, R – absolutni rang

### 3.2.11 November

Novembra je bila prvič v letu presežena meja 1.000 zapisov, kreiranih na posamezen dan, saj je bilo v torek, 26. 11. 2019, kreiranih več kot 1.000 zapisov na posamezen dan, in sicer 1.022; sledi ponedeljek, 18. 11. 2019, z 948 zapisi. Glede na dan v tednu so bili torki (skupno z 21,4 %) najbolj produktivni dnevi za kreiranje zapisov na mesečni ravni (slika 11). November je s 17.601 (10 %) kreiranim bibliografskim zapisom na 1. mestu po številu kreiranih zapisov na letni ravni.



Slika 11: Prirast bibliografskih zapisov – november 2019 (podatki na dan 2. 12. 2019)

Katalogizatorji so v štirih delovnih torkih kreirali 3.770 zapisov, kar predstavlja 21,4 % mesečnega prirasta; sledijo 4 delovni ponedeljki z 20,5 % ali 3.610 zapisi in 4 delovne srede z 19,6 % ali 3.445 zapisi. V delovnih dneh se je COBIB.SI povečal za 17.013 novih zapisov,

povprečno za 850,7 na dan ( $s = 92,2$ ). V 10 dela prostih dneh se je COBIB.SI povečal za 588 zapisov ( $\bar{x} = 58,8$ ;  $s = 28,4$ ). Zapisi, kreirani ob vikendih in na praznični dan, tako predstavljajo 3,3 % mesečnega prirasta.

Mesečni prirast s podrobnimi podatki je predstavljen v tabeli 13.

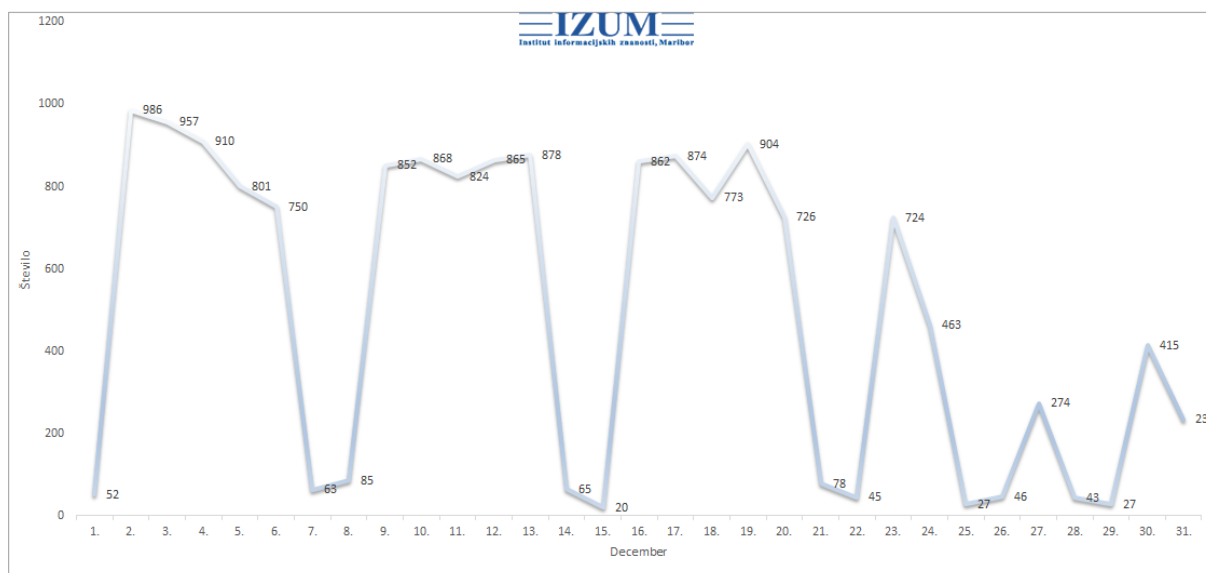
Tabela 13: Bibliografski zapisi, kreirani novembra 2019

Dan	Štev. dni v mescu (n)	Zapisi (n)	Zapisi ( $\bar{x}$ )	Zapisi (s)	Zapisi (%)	Zapisi (maks)	Zapisi (min)	R
ponedeljek	4	3.610	902,5	71,3	20,5	948	797	7
torek	4	3.770	942,5	58,6	21,4	1022	882	8
sreda	4	3.445	861,3	16,0	19,6	884	850	6
četrtek	4	3.284	821,0	30,2	18,7	842	778	5
petek	4	2.904	726,0	85,0	16,5	788	604	4
sobota	5	324	64,8	38,9	1,8	124	25	3
nedelja	4	230	57,5	11,7	1,3	71	45	2
praznik	1	34	34,0	/	0,2	34	34	1
<b>Skupaj</b>	<b>30</b>	<b>17.601</b>	<b>586,7</b>	<b>387,2</b>	<b>100</b>	<b>1022</b>	<b>25</b>	<b>/</b>

Legenda: n – število,  $\bar{x}$  – povprečje, s – standardni odklon, % – delež, maks – maksimum, min – minimum, R – absolutni rang

### 3.2.12 December

December 2019 je z 8,8-odstotnim letnim prirastom uvrščen na četrto mesto po številu kreiranih zapisov na letni ravni ( $n = 15.491$ ). Na sliki 12 je vidno, da je bilo največ zapisov kreiranih v ponedeljek, 2. 12. 2019, in sicer 986, sledita torek, 3. 12. 2019, z 957 in sredo, 4. 12. 2019, z 910 kreiranimi bibliografskimi zapisi.



Slika 12: Prirast bibliografskih zapisov – december 2019 (podatki na dan 6. 1. 2020)

Decembra je bilo 5 delovnih ponedeljkov, ki so bili tudi najbolj produktivni dnevi glede na skupno število kreiranih zapisov ( $n = 3.839$  ali 24,8 %), sledi 5 torkov z 21,9 % ali 3.396 kreiranimi zapisi. Podrobnosti mesečnega prirasta so predstavljene v tabeli 14.

Tabela 14: Bibliografski zapisi, kreirani decembra 2019

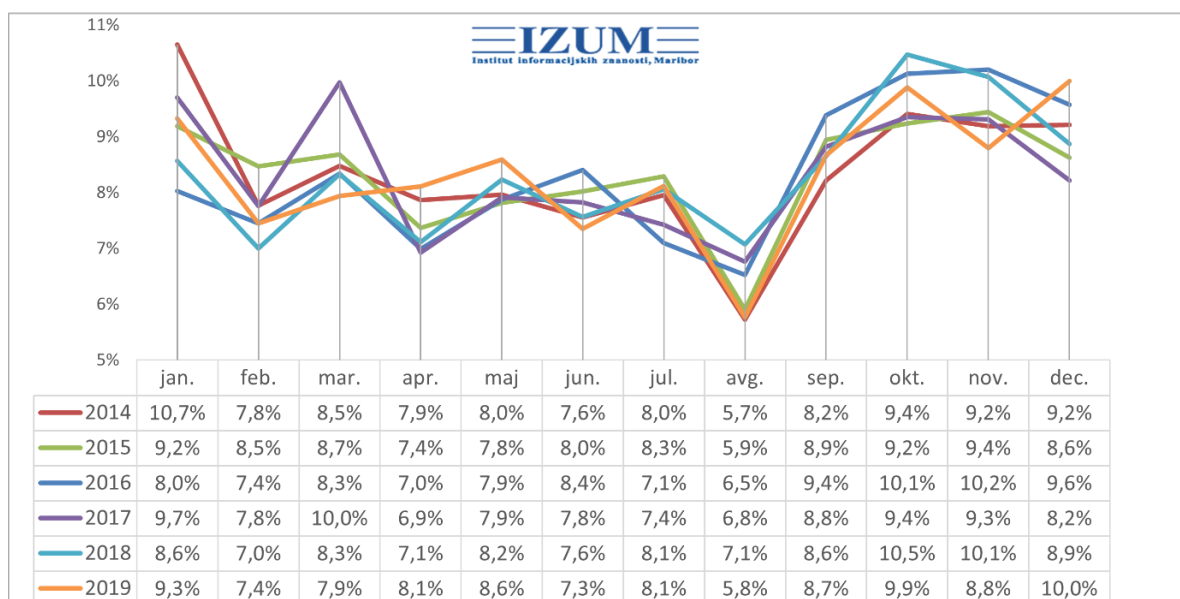
Dan	Štev. dni v mescu ( <i>n</i> )	Zapisi ( <i>n</i> )	Zapisi ( $\bar{x}$ )	Zapisi ( <i>s</i> )	Zapisi (%)	Zapisi ( <i>maks</i> )	Zapisi ( <i>min</i> )	<i>R</i>
ponedeljek	5	3.839	767,8	217,9	24,8	986	415	8
torek	5	3.396	679,2	314,5	21,9	957	234	7
sreda	3	2.507	835,7	69,2	16,2	910	773	4
četrtek	3	2.570	856,7	52,0	16,6	904	801	5
petek	4	2.628	657,0	263,9	17,0	878	274	6
sobota	4	249	62,3	14,5	1,6	78	43	3
nedelja	5	229	45,8	25,5	1,5	85	20	2
praznik	2	73	36,5	13,4	0,5	46	27	1
<b>Skupaj</b>	31	15.491	499,7	382,0	100	986	20	/

Legenda: *n* – število,  $\bar{x}$  – povprečje, *s* – standardni odklon, % – delež, *maks* – maksimum, *min* – minimum, *R* – absolutni rang

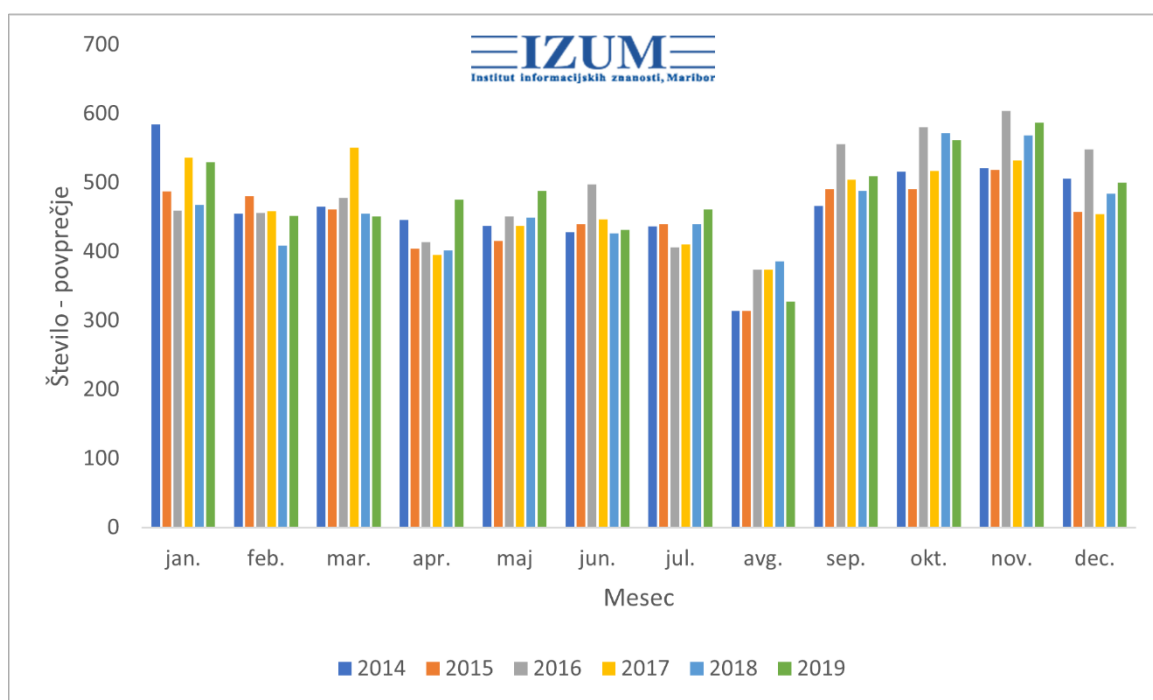
### 3.3 Trend prirasta (2014–2019)

Večletni trend prirasta bibliografskih zapisov se nadaljuje v leto 2019. Vidne so sezonske variacije prirasta (slika 13). Izkazalo se je, da sta torek in sreda najbolj produktivna dneva za kreiranje zapisov. V opazovanem obdobju se je zgodilo nekaj pomembnih dogodkov, npr. 30. 6. 2016 se je končalo obdobje uporabe COBISS2/Katalogizacije. Zagotovo je tudi prehod na COBISS3/Katalogizacijo v letih izvajanja projekta (2012–2016) vplival na prirast. Katalogizatorji so morali usvojiti nov način dela s tretjo generacijo programske opreme. Prav tako je leto 2016 postreglo z novim mejnikom, kar v 4 dneh je bilo kreiranih več kot 1.000 zapisov na dan. Zasluga gre predvsem vnosu zapisov za zaključna dela, saj se je 30. 9. 2016 zaključilo obdobje, do katerega je bilo še možno zaključiti izobraževanje po predbolonjskih študijskih programih. V letu 2018 so se spremenili pogoji za vzdrževanje aktivnega dovoljenja za vzajemno katalogizacijo v sistemu COBISS.SI (2017), ki prav tako vplivajo na delo katalogizatorjev. Poudarek ni več na kvantitativnem delu katalogizatorjev, temveč na izobraževanju kot pogoju za ohranitev aktivnega dovoljenja za vzajemno katalogizacijo.

Povprečni mesečni prirast za opazovano obdobje za posamezno leto je prikazan na sliki 14 (vključeni vsi dnevi leta).



Slika 13: COBIB.SI, trend prirasta, 2014–2019 (%)

Slika 14: COBIB.SI, povprečni mesečni prirast, 2014–2019 ( $\bar{x}$ )

#### 4 Prirast normativnih zapisov

Katalogizatorji pri katalogizaciji bibliografske zapise povezujejo z normativnimi zapisi oz. pravimo, da vzajemna katalogizacija poteka z normativno kontrolo. Če npr. normativni zapis za osebno ime avtorja v bazi CONOR.SI že obstaja, ga povežejo; če ne, kreirajo nov normativni zapis. Za katalogizatorje to pomeni dodatno in zahtevno nalogo, saj morajo biti podatki o avtorjih pravilni in natančni, ob tem pa morajo ločevati med morebitnimi soimenjaki in preprečevati pojav podvojenih zapisov za istega avtorja. Leta 2018 je bila dodana tudi

normativna kontrola imen korporacij, ki je zaživela z letom 2019 (Kos, 2020). Nove normativne zapise za imena korporacij lahko kreirajo katalogizatorji, ki imajo opravljeno dodatno izobraževanje (Kurnjek, 2018). V normativni bazi podatkov CONOR.SI je bilo v letu 2019 kreiranih 60.068 novih normativnih zapisov, od tega 58.015 za osebna imena in 2.053 za imena korporacij (tabela 15). Iz baze LCNAF (Library of Congress Name Authority File) je bilo prevzetih 8,3 % ali 4.976 zapisov. Pri kreiranju zapisov v normativni bazi je sodelovalo 529 katalogizatorjev iz 245 knjižnic (IZUM, 2020; Normativna baza podatkov CONOR.SI, 2020).

Tabela 15: Normativni zapisi v bazi CONOR.SI v letu 2019

Mesec	Zapisi, kreirani v CONOR.SI (n)	Zapisi, prevzeti iz LCNAF (n)
januar	5.511	523
februar	4.848	372
marec	5.274	493
april	5.028	457
maj	4.974	462
junij	3.994	336
julij	4.776	346
avgust	3.638	339
september	5.256	327
oktober	5.944	443
november	5.911	420
december	4.914	458
<b>Skupaj</b>	<b>60.068</b>	<b>4.976</b>

Legenda: n – število; CONOR.SI – normativna datoteka osebnih in korporativnih imen; LCNAF – Library of Congress Name Authority File

## 5 Diskusija

Bibliografsko-kataložna baza COBIB.SI je živa baza, saj se število kreiranih zapisov ves čas spreminja. Predstavljeni so podatki, zbrani čez leto, zato se stanje bibliografsko-kataložne baze in absolutne številke kreiranih zapisov (lahko) razlikujejo od trenutnega števila kreiranih zapisov po posameznem dnevu in/ali mesecu (npr. izbrisani podvojeni zapisi, iz prikaza smo izvzeli tudi bibliografske zapise, ki so bili vključeni v COBIB.SI zaradi ažuriranja baze ELINKS.SI, in zapise iz lokalnih baz podatkov, ki se shranjujejo na vzajemni nivo).

Letni prirast, razdeljen na kvartale, kaže, da je bil skupni prirast najvišji v zadnjem kvartalu (28,7 %), sledijo prvi (24,7 %), drugi (24 %) in tretji (22,5 %). V prvem, drugem in četrtem kvartalu je bil najbolj produktiven dan torek, v tretjem kvartalu pa ponedeljek. Največ zapisov glede na posamezen dan v tednu in na letni ravni je bilo kreiranih ob torkih (21,2 %), sledijo: ponedeljki (20,3 %), srede (19,9 %), četrtki (19,3 %) in petki (17 %). Torek je najbolj produktiven dan tudi na ravni posameznih mesecev, in sicer: v februarju, marcu, aprilu, maju, juliju, oktobru in novembru. Ponedeljek je bil (na mesečni ravni) najbolj produktiven v septembru in decembru, sreda v juniju, četrtek v januarju, avgusta pa je najbolj produktiven dan petek.

Po podatkih v Centralnem registru uporabniških imen (21. 9. 2020) ima v Sloveniji 645 oseb aktivno uporabniško ime in Dovoljenje za vzajemno katalogizacijo monografskih publikacij (DVKA). Seznam vseh je dosegljiv na spletni strani portala za izobraževanje IZUM (Dovoljenja za vzajemno katalogizacijo, 2020). Med njimi jih ima 69,1 % ( $n = 446$ ) eno uporabniško ime, 17,2 % ( $n = 111$ ) dve uporabniški imeni, 8,8 % ( $n = 57$ ) jih ima tri uporabniška imena, 4,8 % ( $n = 31$ ) pa štiri ali več. Več uporabniških imen kaže na to, da knjižničarji sodelujejo v več organizacijah. To sklepamo tudi zaradi 2,3-odstotnega deleža letnega prirasta bibliografskih zapisov, ki so bili kreirani ob praznikih in dela prostih dneh. Vse to tudi nakazuje na večje potrebe po usposobljenih katalogizatorjih. Prav tako v letu 2019 ni enega dneva, ko ne bi bil kreiran vsaj en bibliografski zapis. V vzajemni katalogizaciji je zapise kreiralo 575 katalogizatorjev (Bibliografske baze ..., 2020).

Pričakovano je bilo najmanj zapisov kreiranih avgusta (5,8 %), kar je tudi sicer večletni trend in ga pripisujemo sezonski komponenti, saj je to mesec letnih dopustov. Pričakovali smo, da bo tudi februar pod vplivom sezonske komponente, saj so takrat šolske počitnice, in predvidevamo, da je posledično v delovnem okolju prisotnih manj katalogizatorjev, a v letu 2019 ni bilo tako.

Poudariti moramo kakovost bibliografskih in normativnih zapisov. V vseh pogledih ima kakovostno delo katalogizatorjev zelo pomembno vlogo v sistemu COBISS in vpliva na več storitev, ki temeljijo na bibliografskih zapisih (npr. izposoja po knjižnicah, vodenje osebnih bibliografij, evalvacija znanstvenoraziskovalnega dela, knjižnično nadomestilo itd.). Enako velja za normativne zapise (npr. normativni zapisi za osebna imena avtorjev baze podatkov CONOR.SI so preneseni v mednarodni sistem Virtual International Authority File (VIAF) (Krajnc Vobovnik, 2020) in so dostopni širši svetovni javnosti). Ob na novo kreiranih zapisih pa ne smemo pozabiti na številne redakcije bibliografskih in normativnih zapisov, ki jih katalogizatorji opravijo. Po podatkih je bilo v opazovanem letu popravljenih 83.589 bibliografskih zapisov<sup>2</sup> (Bibliografske baze podatkov ..., 2020). Pri redakciji normativnih zapisov je poudarek na nenehnem vzdrževanju in nadgrajevanju podatkov v zapisu. To daje podatkom dodano vrednost in kaže na njihovo kakovost. Glede na pomembnost katalogizatorjevega dela in ob zavedanju pomena kakovostnih podatkov v COBIB.SI, v IZUM-u izvajamo aktivnosti, ki so del sistema zagotavljanja kakovosti bibliografskih in normativnih zapisov (Badovinac, 2016; 2020). S tem želimo prispevati k čim višji kakovosti podatkov v COBIB.SI.

## 6 Zaključek

V vseh lokalnih bazah podatkov v mreži COBISS.net je v letu 2019 vnesenih že več kot 43 milijonov enot gradiva. Februarja 2019 smo v lokalnih bazah COBISS.SI zabeležili kreacijo 16-milijontega zapisa. Prav tako je v nekaterih lokalnih bazah še 442.217 zapisov, ki niso povezani<sup>3</sup> (Bibliografske baze podatkov ..., 2020). Ti zapisi so lahko dragocen (zgodovinski) podatek o obstoju določenega gradiva, dogodka ipd. Potrebna sta dodaten trud in čas, da se taki zapisi ustrezno uredijo (popravijo/dopolnijo). Seveda gre lahko tudi za podvajanje podatkov z že obstoječimi zapisi v vzajemni bazi, v tem primeru se morajo lokalni zapisi brisati, a tudi to zahteva katalogizatorjev vložek v redakcijo baze. Informacije in podatki, ki so zabeleženi v bazah, so pomembni podporniki razvoja novega znanja in razvoja družbe kot

<sup>2</sup> Število zapisov, ki jih je posamezna knjižnica uredila v vzajemni bazi podatkov COBIB.SI v letu 2019, vendar ni njihov kreator.

<sup>3</sup> Število lokalnih zapisov na dan 31. 12. 2019.

celote. Glede na zapisano ni dvoma, kako pomembna baza podatkov je COBIB.SI za slovenski prostor, baze COBISS.net pa za jugovzhodno Evropo.

## Zahvala

Iskrena hvala vsem in vsakemu katalogizatorju posebej, saj s svojim strokovnim in vestnim delom prispevajo h kakovostnim podatkom v vseh bazah COBISS.net.

## Reference

Badovinac, B., 2016. Spremljanje kakovosti bibliografskih in normativnih zapisov dnevne produkcije. *Blog COBISS*, 25. 7. 2016. Dostopno na:

<https://blog.cobiss.si/2016/07/25/spremljanje-kakovosti-bibliografskih-in-normativnih-zapisov-dnevne-produkcije/> [22. 9. 2020].

Badovinac, B., 2020. Več podatkov zagotavlja boljšo fasetno navigacijo v COBISS+. *Blog COBISS*, 11. 6. 2020. Dostopno na: <https://blog.cobiss.si/2020/06/11/vec-podatkov-zagotavlja-boljso-fasetno-navigacijo-v-cobiss/> [22. 9. 2020].

*Bibliografske baze podatkov in podatki o zalogi: polnopravni člani sistema COBISS.SI in njihovo sodelovanje v sistemu vzajemne katalogizacije*, 2020. Dostopno na:

[https://www.cobiss.si/kazalci/doc/STAT\\_PRIRAST\\_2019\\_SI\\_2.pdf](https://www.cobiss.si/kazalci/doc/STAT_PRIRAST_2019_SI_2.pdf) [2. 6. 2020].

*COBISS.net v številkah*, 2020. Dostopno na:

[https://www.cobiss.net/doc/Statistike\\_2019\\_slo.pdf](https://www.cobiss.net/doc/Statistike_2019_slo.pdf) [11. 9. 2020].

*COBISS.net: mreža knjižničnih informacijskih sistemov in informacijskih sistemov o raziskovalni dejavnosti*, 2020. Dostopno na: <https://www.cobiss.net/si/> [3. 8. 2020].

*COBISS3 – Knjižnice*, 2020. Dostopno na: <http://home.izum.si/cobiss/cobiss3/> [20. 7. 2020].

Dornik, E., 2015. COBIB.SI: prirast bibliografskih zapisov v letu 2014. *Organizacija znanja*, 20(1), str. 21–26. Dostopno na: <https://doi.org/10.3359/oz1501021> [2. 10. 2020].

Dornik, E., 2016. COBIB.SI: prirast bibliografskih zapisov v letu 2015. *Organizacija znanja*, 21(2), str. 77–84. Dostopno na: <https://doi.org/10.3359/oz1602077> [2. 10. 2020].

Dornik, E., 2017. COBIB.SI: prirast bibliografskih zapisov v letu 2016. *Organizacija znanja*, 22(1–2), str. 27–33. Dostopno na: <https://doi.org/10.3359/oz1712027> [2. 10. 2020].

Dornik, E., 2018. COBIB.SI: prirast bibliografskih zapisov v letu 2017. *Organizacija znanja*, 23(1–2), str. 21–28. Dostopno na: <https://doi.org/10.3359/oz1812021> [2. 10. 2020].

Dornik, E., 2019. COBIB.SI: prirast bibliografskih zapisov v letu 2018. *Organizacija znanja*, 24(1–2), 1924006, str. 1–19. Dostopno na: <https://doi.org/10.3359/oz1924006> [2. 10. 2020].

*Dovoljenja za vzajemno katalogizacijo*, 2019. Dostopno na:

<https://izobrazevanje.izum.si/EntryFormDesktopDefault.aspx?tabid=360&type=licence> [25. 6. 2020].

IZUM, 2020. *Letno poročilo o delu IZUM za leto 2019*. Maribor: Institut informacijskih znanosti. Dostopno na: [https://www.izum.si/doc/Letno\\_porocilo\\_2019.pdf](https://www.izum.si/doc/Letno_porocilo_2019.pdf) [3. 8. 2020].



Kos, J., 2020. Tehnike iskanja v CONOR.SI pri kreiranju normativnih zapisov za imena korporacij. *Blog COBISS*, 14. 9. 2020. Dostopno na: <https://blog.cobiss.si/2020/09/14/tehnike-iskanja-v-conor-si-pri-kreiranju-normativnih-zapisov-za-imena-korporacij/> [22. 9. 2020].

Krajnc Vobovnik, A., 2020. Ažuriranje zapisov v VIAF in priporočila za kakovostnejši prikaz slovenskih normativnih zapisov v VIAF. *Organizacija znanja*, 25(1–2), 2025005, str. 1–18. Dostopno na: <https://doi.org/10.3359/oz2025005> [22. 9. 2020].

Kurnjek, B., 2018. Projekt Normativna kontrola za imena korporacij v sistemu COBISS.SI. *Blog COBISS*, 28. 9. 2018. Dostopno na: <https://blog.cobiss.si/2018/09/28/normativna-korporacij/> [18. 9. 2020].

*Normativna baza podatkov CONOR.SI*, 2020. Dostopno na: [https://www.cobiss.si/kazalci/doc/STAT\\_PRIRAST\\_CONOR\\_2019\\_SI.pdf](https://www.cobiss.si/kazalci/doc/STAT_PRIRAST_CONOR_2019_SI.pdf) [18. 9. 2020].

*Organizacijski model sistema COBISS*, 2020. Dostopno na: <https://www.cobiss.net/si/platforma-cobiss.htm#c2> [31. 7. 2020].

*Vzajemna katalogizacija (COBISS3/Katalogizacija)*, 2020. Dostopno na: <https://www.cobiss.net/si/platforma-cobiss.htm#c4> [25. 6. 2020].

*Vzdrževanje aktivnega dovoljenja za vzajemno katalogizacijo v sistemu COBISS.SI*, 2017. Dostopno na:

[http://home.izum.si/cobiss/obvestila\\_novosti/dokumenti/Priloga\\_Vzdr%C5%BEevanje\\_dovoljenja\\_za\\_vzajemno\\_katalogizacijo.pdf](http://home.izum.si/cobiss/obvestila_novosti/dokumenti/Priloga_Vzdr%C5%BEevanje_dovoljenja_za_vzajemno_katalogizacijo.pdf) [1. 10. 2020].