Enlightenment spirit in the 21st century: hipertextual reached material of "Collected works of Dositej Obradović"

Jelena Andonovski
andonovski@unilib.rs

Ivana Gavrilović
gavrilovic@ubsm.rs

University Library
„Svetozar Marković"
Belgrade, Serbia

ABSTRACT: University library „Svetozar Marković" has been successfully cooperating for years with Foundation „Dositej Obradović" from Belgrade. The main goal of this cooperation was to create searchable digital collection of Dositej Obradović's works as well as works about Dositej Obradović, so that different target groups (students, researchers, professors), interested in Dositej Obradović's life and work, would have all necessary literature for their research gathered at one place. Foundation „Dositej Obradović" also published edition named "Collected works of Dositej Obradović" at 2007/2008 year. During implementation of the project this collection was processed using up-to-date technologies for text processing while librarians from University library created portal of Dositej Obradović's works. That portal is a part of the greater University library's digital collection.

KEYWORDS: Dositej Obradović, digital collections, annotation, XML, TEI, metadata, named entities

1 Introduction

University library „Svetozar Marković" has been successfully cooperating with Foundation „Dositej Obradović"[1] from Belgrade for years. The

---

[1] Foundation „Dositej Obradović"

main purpose of this cooperation was to create searchable digital collection of Dositej Obradović's works as well as works about Dositej Obradović so that different target groups (students, researchers, professors), interested in Dositej Obradović's life and work, would have gathered all necessary literature for their research at one place. Cooperation started in 2018 with project *Searchable collected works of Dositej Obradović: creating portal for review and unified search of digitized material published by Foundation „Dositej Obradović"*. The cooperation was continued also in 2019 with project *Collected works of Dositej Obradović for 21st century: multimedial complemented material published by Foundation „Dositej Obradović"*. Both projects were financially supported by Ministry of Culture and Information while the edition *Collected works of Dositej Obradović in six volumes* was used for their implementation. Dositej's works were popular during writer's life and they have experienced multiple editions (1870, 1902, 1932, 1953). The most famous and the almost completed among all are "Collected works" in three volumes which „Prosveta" published in 1961 within the edition „Serbian writers" (editors Đuro Gavela, Jelena Šaulić, Borivoje Marinković) (Обрадовић, 1961). Scientific edition *Collected works of Dositej Obradović in six volumes* was prepared for anniversary *200 years since Dositej's arrival in Serbia* and The Government of the Republic of Serbia proclaimed year 2007 as „Dositej Obradović's year". Foundation „Dositej Obradović" started project of publishing scientific edition *Collected works of Dositej Obradović in six volumes* and the editors were: professor PhD Mirjana D. Stefanović, PhD Miodrag Maticki and professor PhD Dušan Ivanić. In six volumes they united all Dositej's works that were officially confirmed as completed, works attributed to this writer, even in cases that it is not completely confirmed, as well as his complete correspondence known by now (Обрадовић, 2007, 5-6). In addition to the works, editors also prepared annotations and comments which better explain less known terms, people and events mentioned in the texts, dictionary of less known words, registers of proper names and geopolitical names, registers of abbreviations and references, and complete bibliography about Dositej Obradović. „Collected works of Dositej Obradović in six volumes" contained the following six books:

1. First book – *Pismo Haralampiju i Život i priključenija* (ed. Mirjana D. Stefanović);
2. Second book – *Basne i dve pripovetke, Istina i prelest, Put u jedan dan* (ed. Miodrag Maticki);
3. Third book – *Sovjeti zdravago razuma, Slovo poučitelno gospodina Georgija Joakima Colikofera, Etika* (ed. Dušan Ivanić);

4. Fourth book – *Sobranije raznih naravoučitelnih veštej, Mezimac* (ed. D. Ivanić);
5. Fifth book – *Spisi iz Dalmacije, Propovedi i besede, Zapisi i beleške, Fragmenti iz Srbije* (ed. M. Maticki);
6. Sixth book – *Pesme, Pisma, Dokumenti* (prir. M. D. Stefanović).

For the implementation of the above-mentioned projects, the Foundation "Dositej Obradovic" handed "Colected works" to University library "Svetozar Marković". As the result, the project named *Collected works of Dositej Obradović in six volumes*[2] was developed.

2 Text processing

In order to prepare a digital collection available to the users and searchable by different parameters, all material must be in appropriate digital format, processed with up-to-date technologies for text processing. If material is not in digital format, which is very common, it is first necessary to scan it (Ikonomov and Dobreva, 2008, 2), and then to convert the obtained images into machine-readable text using OCR software (Optical Character Recognition) (Тртовац, 2016, 105-107). However, if textual material is already in appropriate digital format, as it was in this case, scanning is not necessary, and the first step is text processing.[3] At the beginning of the first project in 2018, University library „Svetozar Marković" was given material in digital format, in .docx. Scanning and OCR was not necessary, which was a significant timesaving, and the first step was text processing. Text processing had a couple of phases and a final product is material transformed in appropriate XML (Extensible Markup Language)[4]format (W3C, 2020a). Material in XML format which specifies encoding methods for machine-readable texts and embedded header (TEI Header) for defining metadata about describing material (TEI Consortium, 2020) was prepared according to Text Encoding Initiative (TEI)[5] Guidelines. Creation of XML documents implies annotation of formal and structural characteristics of textual material, as well as indexing entities within text. This is important task so the material later

---

[2] Collected works of Dositej Obradović

[3] The complete versions of „Collected works of Dositej Obradović" prepared for publishing were handed to University library "Svetozar Markovic" by Foundation "Dositej Obradovic".

[4] Extensible Markup Language

[5] Text Encoding Initiative

can be searchable according to different parameters and it is also possible
to link indexed entities with appropriate resources on internet. For Dositej
Obradović's works, XML documents were prepared automatically. Textual
material in .docx format is already compressed file. It is saved in XML for-
mat, which allows extraction in appropriate XML editor. In this case the
XML editor named Oxygen XML Editor was used. Opening .docx file in
this editor, it is exported in XML format and for further processing doc-
ument.xml file is used. In Oxygen XML Editor there is an option DOCX
TEI P5 which automatically transforms document.xml file in TEI/XML file
(Figure 1). This TEI/XML files can be further processed. Librarians from
University library used Oxygen XML Editor and the complete procedure was
applied on Dositej Obradović's works which is explained in next sections.

Figure 1. File document.xml for work *Spisi iz Dalmacije, Propovedi i besede, Za-
pisi i beleške, Fragmenti iz Srbije* and DOCX TEI P5 option
.

3 Annotation of formal characteristics

Annotation of formal characteristics implies metadata creation that
"describes, explains, locates, or otherwise makes it easier to retrieve, use

or manage an information resource" (Hodž, 2004, 157) with "well de-
fined semantics and structure" (Berners-Lee, 1997). The descriptive (bib-
liographical) and administrative metadata for Dositej Obradović's works
were created in the header of TEI/XML file (teiHeader) (Тртовац, 2016,
43-45). Descriptive (bibliographical) metadata can be divided into two
groups: the first one that describes digital version of text (наслов/title,
аутор/author, одговорност/respStmt и датум/date) and the second
one that describes analog version of text (аутор/author, наслов/title,
место издавања/pubPlace, издавач/publisher, датум/date и податак о
едицији/series). Along with descriptive (bibliographical) metadata, there
is also metadata about the license (administrative metadata) which defines
permission for users regarding the usage of digital material (annotated with
tag ref in the part publisher in publicationStmt). Figure 2 represents tei-
Header example for work Pismo Haralampiju i Život i priključenija, while
Figure 3 represents visual view for metadata from this header in the portal
"Collected works of Dositej Obradović".

```xml
<teiHeader>
    <fileDesc>
        <titleStmt>
            <title>Писмо Харалампију, Живот и прикљученија: електронско издање</title>
            <author ref="viaf:76370256">Доситеј Обрадовић</author>
            <respStmt>
                <resp>Израда електронског издања</resp>
                <name>Универзитетска библиотека "Светозар Марковић"</name>
            </respStmt>
        </titleStmt>
        <editionStmt>
            <edition><date>2019-10-15</date></edition>
        </editionStmt>
        <publicationStmt>
            <publisher>
                <ref target="http://creativecommons.org/licenses/by-nc-sa/4.0/">Објавено под лиценцом Creative
                    Commons Attribution-NonCommercial-ShareAlike 4.0
                    International</ref>
                </publisher>
        </publicationStmt>
        <sourceDesc>
            <bibl type="copyText">
                <author>Доситеј Обрадовић</author>
                <title ref="COBISS-SR.ID:144066572">Писмо Харалампију, Живот и прикљученија: електронско издање</title>
                <pubPlace>Београд</pubPlace>
                <publisher>Задужбина Доситеј Обрадовић</publisher>
                <date>2007</date>
                <series n="1">Сабрана дела Доситеја Обрадовића</series>
                <idno type="wiki">https://sr.wikipedia.org/wiki/Доситеј_Обрадовић</idno>
                <idno type="viaf">https://viaf.org/viaf/76370256/#Obradović,_Dositej,_1739-1811</idno>
            </bibl>
        </sourceDesc>
    </fileDesc>
</teiHeader>
```

Figure 2. Example for teiHeader for work *Pismo Haralampiju i Život i priključenija* in TEI/XML format

.

```
<teiHeader>
    <fileDesc>
        <titleStmt>
            <title>Писмо Харалампију, Живот и прикљученија: електронско издање</title>
            <author ref="viaf:76370256">Доситеј Обрадовић</author>
            <respStmt>
                <resp>Израда електронског издања</resp>
                <name>Универзитетска библиотека "Светозар Марковић"</name>
            </respStmt>
        </titleStmt>
        <editionStmt>
            <edition><date>2019-10-15</date></edition>
        </editionStmt>
        <publicationStmt>
            <publisher>
                <ref target="http://creativecommons.org/licenses/by-nc-sa/4.0/">Објављено под лиценцом Creative
                    Commons Attribution-NonCommercial-ShareAlike 4.0
                    International</ref>
                </publisher>
        </publicationStmt>
        <sourceDesc>
            <bibl type="copyText">
                <author>Доситеј Обрадовић</author>
                <title ref="COBISS-SR.ID:144066572">Писмо Харалампију, Живот и прикљученија: електронско издање</title>
                <pubPlace>Београд</pubPlace>
                <publisher>Задужбина Доситеј Обрадовић</publisher>
                <date>2007</date>
                <series n="1">Сабрана дела Доситеја Обрадовића</series>
                <idno type="wiki">https://sr.wikipedia.org/wiki/Доситеј_Обрадовић</idno>
                <idno type="viaf">https://viaf.org/viaf/76370256/#Obradović,_Dositej,_1739-1811</idno>
            </bibl>
        </sourceDesc>
    </fileDesc>
</teiHeader>
```

Figure 3. Visual view for metadata in the portal *Collected works of Dositej Obradović* for work *Pismo Haralampiju i Život i priključenija*
.

4   Annotation of structural characteristics

In addition to the annotation of formal characteristics, the annotation of their structural characteristics was also done for every work in the Collection. Annotation of structural characteristics implies tagging of logical text structures (text body, chapters, titles, paragraphs, sentences, words etc.) with appropriate XML tags and can be done automatically or manually. The edition "Collected works of Dositej Obradović" was mainly automatically annotated. After automatic transformation in TEI/XML file (described in the section Text processing), all necessary logical structures of texts were annotated: text body, chapters, head titles and paragraphs. Text body and paragraphs were annotated with appropriate XML tags directly, tag $<body>$ for the text body and $<p>$ for the paragraphs, and corrections were not necessary. However, chapters and their titles were also annotated, but with XML tags for paragraphs. According to TEI-Schema specification, the chapters are annotated with XML tag $<div>$ while their titles are annotated with XML tag $<head>$, so the regular expressions are used for their correction.

Regular expression example:
Find:

`<p rend=Heading #[0-9]+">`
`<anchor xml:id="bookmark[0-9]+"/>`[([А-Ша-ш]+) ([а-шА-Ш]+)]+
`</p>`"
Replace:
`<head style=text-align:center;" rend="italic">`&1`</head>`"

To keep digital version identical to the analog as much as possible, with tags for heads, titles and paragraphs there are also attributes that define some content characteristics. For example:

Chapter annotation:
Example 1:
<div type="chapter">
<head style="text-align:center;">1.</head>
. . .</div>
Example 2:
<div type="chapter">
<head   style="text-align:center;"   rend="italic">Предисловије   о баснах</head>.
. . .</div>
Example 3:
<div type="chapter" n="8[9]">
<head style="text-align:center;" rend="italic">8[9]</head>
<head style="text-align:center;" rend="italic">Лав и зец</head>
. . .</div>

In examples for chapter annotation one can see that there is always an attribute *type* with value *"chapter"* (examples 1, 2 and 3) alongside the tag *<div>*. Regarding chapter titles, which are positioned central in the original document, there is always an attribute style with the value *"text-align:center;"* (examples 1 and 3) alongside the tag *<head>*. Most of chapter titles are also italic, thus alongside an attribute style and tag *<head>*, there is also an attribute rend with the value *"italic"* (examples 2 and 3).

Paragraph annotation:
Example 4:
<p>[Писмо<rs type="person">Харалампију</rs>]<note>Овај рад Доситеј Обрадовића нема свој насловни лист. Због тога приређивач, према садржају и према адресату, насловљава текст и тај наслов ставља у угласте заграде. Текст је био штампан курзивом, као што је уобичајено било за предговоре. Овде су пагиновани листови, а не странице, и то звездицама, што је, такође, био обичај обележавања за предговоре. У науци постоје различита мишљења о том да ли је овај рад Доситејев

посебна библиографска јединица (Стефановић, 2007. И ту наведену литературу; в. и Библиографија, 1990).... </p>

Example 5:

\<p style="text-align:center;">ЉУБИМАЈА ЈУНОСТЕ,</p>

Example 6:

\<p rend="italic">Здравствуј! Христос воскресе... </p>

Regarding paragraphs, tag *<p>* was almost without attributes (example 4). But there are some examples of paragraphs that are centered or italic, so it was necessary to annotate that characteristics. The central position of paragraph was annotated with attribute style and its value *"text-align:center;"* (example 5), while italic characteristic was annotated with attribute rend and its value *"italic"* (example 6), the same as with chapter titles. Tag *<pb>* with attribute *n* that annotates the number of pages was also used for pagination annotation, beside chapters and paragraphs. For example, *<pb n="15"/>* pagination was annotated in order to prepare the identical review of both TEI/XML file and theoriginal document in the portal "Collected works of Dositej Obradović". Pagination annotation was also automatically done, using regular expression. Regarding annotation of logical text structures, the parts of text in foreign language, as well as highlighted parts of text: bold, italic, underlined, etc. were annotated in each text. For parts of text in foreign language the tag *<foreign>* with attribute *xml:lang* which value was ISO alpha-2 language code was used. For example, *<foreign xml:lang="it">„Farina de diavolo va presto in semola</foreign>*. Highlighted parts of text as bold, italic, underlined were annotated with tag *<hi>* with attribute rend which value depended on highlight characteristic (bold, italic, underlined). For example, *<hi rend="bold">Кажу да се чуде</hi>*. It is important to say that highlighted parts of text were automatically annotated after transformation in TEI/XML file (explained in section Text processing).

5   Text annotation

5.1   Entity annotation with descriptive notes

Each work in *Collected works of Dositej Obradović* has a list of notes which represents very important part of this edition, at its end. These notes are very important for understanding and finding out more information about different entities in the texts: persons, geopolitical places, time periods, some titles, significant syntagmas, metaphors and verses. They help

researchers to find out more about persons that Dositej knew, the period of life he met them, their importance for his life and creativity, as well as more about certain geopolitical places he resided in during his life time. Thanks to the notes, researchers can also understand more easily the meaning of some syntagmas and metaphors that Dositej used in his writings. Notes are chronologically listed at the end of each work in edition, according to entities appearance in the text. Each item in the list contains the original entity form from the text and the explanation (note), as well as the number of pages by which the entity can be found in the text. The entity annotation with descriptive notes means that all entities from the list got their explanation in the text so that researchers, reading the text, instantly see the note and do not search for the explanation at the end of the work. Entity annotation was done manually with tag $<note>$. The note for entity "Здравствуј! Христос воскресе" in the work *Pismo Haralampiju i Život i priključenija* is represented in the Figure 4, while the annotations in TEI/XML format and visual view in the portal *Collected works of Dositej Obradović* are represented in the Figure 5 and Figure 6.

Figure 4. Note for entity "Здравствуј! Христос воскресе" at the end of the text *Pismo Haralampiju i Život i priključenija*

Figure 5. Entity annotation "Здравствуј! Христос воскресе" in TEI/XML file *Pismo Haralampiju i Život i priključenija*

5.2 Named entity annotation

Beside entity annotation with descriptive notes, the named entities such as proper names and geopolitical names were also annotated in the texts.
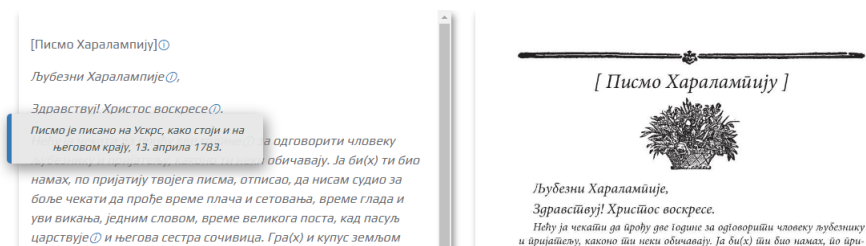
Figure 6. Visual view of entity annotation "Здравствуј! Христос воскресе" in the portal *Collected works of Dositej Obradović*

.

Named Entity Recognition (NER) is one of the technologies in information retrieval area (Chinchor and Robinson, 1997) and the leading subject in Natural Language Processing (NLP) form more than fifteen years (Vitas et al., 2007, 115-116) (Krstev et al., 2014). Named entity recognition means finding specific terms in the text and their classification in the categories (Андоновски, 2019, 124). Named Entity Recognition and their classification is important task for many scientific fields. Not only that it is one of five subtasks in information extraction in information retrieval area, but it is also important task in Natural Language Processing and machine translation within computational linguistics (Utvić, 2008, 2). Named entity annotation in edition "Collected works of Dositej Obradović" was done manually. The tag $<rs>$, attribute type with appropriate value, attribute *person* for proper names and attribute *place* for geopolitical terms were used for annotation.Figure 7 represents named entity annotation in work *Sobranije raznih naravoučitelnih veštej, Mezimac* in TEI/XML file. In this example the named entities „Сократ", „Софрониска", „Фенарете", „Атини" и „Христа"[6] are annotated. The named entity „Атини" is geopolitical place in this example, which is the city of Athens, while other named entities are proper names. Named entity annotation enables precise retrieval by search engine of the portal. It is now also possible to link them with appropriate resources

---

[6] Syntagma "469 година пре Христа" in Figure 7 is named entity classified as temporal expression. However, only proper and geopolitical names are annotated within the project described in this paper. This syntagma is not annotated as temporal expression and only name "Христ" is annotated as proper name. Temporal expressions will be annotated in the upcoming steps and this will be changed.

in Linked Open Data cloud such as authority files of proper names, authority files for geopolitical terms, proper ontology, etc. in Linked Open Data cloud (Berners-Lee, 2006; Radulovic et al., 2015; W3C, 2020b; Blaney, 2020).

```
<p><rs type="person">Сократ</rs>, синједног каменосечца именем
<rs type="person">Софрониска</rs> и једне жене званијен бабице, а
именем <rs type="person">Фенарете</rs>, родио се је у
<rs type="place">Атини</rs> 469 година пре
<rs type="person">Христа</rs>. Испочетка научи занат
својега оца и историја чини воспоминаније о трима
ста|туама, које представљаху три грације и које су предивно
41 соделате биле. Види се од приуподобљивања која је овај
преславни мук у своји разговори употребљавао, да се он
нимало није стидио заната свога родитеља и родитељнице.
„Зачудо", говораше он, ,,да један каменосечац</p>
```

Figure 7. Named entity annotation in work *Sobranije raznih naravoučitelnih veštej, Mezimac*, TEI/XML file

.

6 Linking metadata

Bibliographic metadata from the teiHeader of each document, which are author name and title, were linked with records of some appropriate resources on internet. Author name, Dositej Obradović, was linked with record about Dositej Obradović in Virtual International Authority File (VIAF)[7], which is part of LOD cloud, as well as with web page about Dositej Obradović in Serbian Wikipedia. Title of each work was linked with appropriate record in electronic catalogue COBISS.SR.[8] In mentioned databases each record has its identification number which is unique ID used for linking. ID number of records in VIAF and COBISS.SR are represented in teiHeader as value of attribute ref in tags *<author>* and *<title>*. Wikipedia link of the Dositej Obradović's web page was represented in the tag *<idno>* and its attribute *type* with value „wiki" (Figure 8). In this way users can get more information about author and titles for their research. Bases VIAF and COBISS are important especially for library and information scientific community. VIAF is international authority file of proper names, organizations and titles, joint project of several national libraries. COBISS.SR is an electronic catalogue of The Republic of Serbia. These two bases are products of library and information science and are very important in this filed.

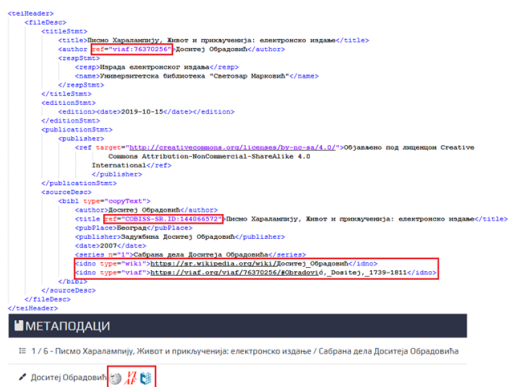[7] Virtual International Authority File

[8] COBISS.SR

Figure 8. Note for entity "Здравствуј! Христос воскресе" at the end of the text
Pismo Haralampiju i Život i priključenija

.

Authority files and electronic catalogues are among the most important resources in library and information science. Linking metadata with these resources points out their importance and impacts new technological possibilities that semantic web offers. On the other hand, they are important for libraries and librarians to see advantages of their use in information retrieval field. Richness of resources that already exists in libraries, as well as the creation of new ones are great starting point for implementing new technological possibilities of semantic web (Андоновски, 2019, 237-238). Knowledge bases, Wikipedia for example, give us interesting information about different entities and it can be starting point for some further research. For that reason, these bases are selected as the first resources to link metadata within.

7   Conclusion

After successful cooperation for years, Foundation "Dositej Obradović" and University library "Svetozar Marković" have continued their work in 2020 on new project *Bibliography about Dositej Obradović: bibliographic metadata annotation and their linking with important library and information resources in Linked Open Data cloud*. The aim of the project is to enrich digital collection in the portal *Collected works of Dositej Obradović* with a new textual corpus material, bibliographic material about Dositej Obradović and his

work. On the other hand, it is necessary to update technical characteristics of the portal so that the detailed annotation of bibliographic metadata and their linking with important library and information resources on internet can be successfully represented. In this way the creation of a big textual corpus about Dositej Obradović's life and work begins. The annotation of bibliographic metadata for each bibliography unit according to TEI Guidelines and their linking with appropriate library and information resources, authority databases and knowledge bases that are part of the Linked Open Data cloud are the special focus of the project. Having in mind that the textual corpus about Dositej Obradović and its works is huge, only a small part of it will be processed for the purposes of this project. The selected material will be also complemented with sub collection *Dositej Obradović* in the collection *Kritika (18. i 19. vek)* on the portal *Srpska književna kritika*[9], also developed by University library "Svetozar Marković".

References

Berners-Lee, Tim. "Metadata Architecture", 1997. Accessed July 31, 2020. https://www.w3.org/DesignIssues/Metadata.html

Berners-Lee, Tim. "Linked Data", 2006. Accessed July 31, 2020. http://www.w3.org/DesignIssues/LinkedData.html

Blaney, Jonathan. "Introduction to the principles of Linked Data", 2020. Accessed August 8, 2020. https://programminghistorian.org/en/lessons/intro-to-linked-data

Chinchor, Nancy and Patricia Robinson. "MUC-7 named entity task definition". In *Proceedings of the 7th Conference on Message Understanding*, Vol. 29, 1–21. 1997

TEI Consortium. *TEI P5: Guidelines for Electronic Text Encoding and Interchange*, 2020. Accessed July 31, 2020. https://tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf

Hodž, Gejl. "Metapodaci na lak način". *Glasnik Narodne biblioteke Srbije* br. 1 (2004): 157–180. Accessed July 31, 2020. http://www.nb.rs/view_file.php?file_id=860

Ikonomov, Nikola and Milena Dobreva. "The making of... digital book". *NCD Review* no. 13 (2008): 1–8. Accessed August 8, 2020. http://elib.mi.sanu.ac.rs/files/journals/ncd/13/ncd13001.pdf

---

[9] Српска књижевна критика, https://kritika.unilib.rs/. Потколекција „Доситеј Обрадовић", https://kritika.unilib.rs/en/islandora/object/kritika%3Adositej

Vitas, Duško, Cvetana Krstev and Denis Maurel. "A note on the Semantic and Morphological Properties of Proper Names in the Prolex Project". *Lingvisticae Investigationes, Special issue on Named Entities: Recognition, Classification and Use* Vol. 30, no. 1 (2007): 115–134

Krstev, Cvetana, Ivan Obradović, Miloš Utvić and Duško Vitas. "A system for named entity recognition based on local grammars". *Journal for Logic Computation* Vol. 24, no. 2 (2014): 473–489. Accessed October 19, 2020. https://academic.oup.com/logcom/article-abstract/24/2/473/1008564?redirectedFrom=fulltext

Radulovic, Filip, María Poveda-Villalón, Daniel Vila-Suero, Víctor Rodríguez Doncel, Raúl García-Castro et al.. "Guidelines for Linked Data generation and publication: An example inbuilding energy consumption". *Automation in Construction* no. 57 (2015): 178–187

Utvić, Miloš. "Konačni automati u regularnoj imenskoj derivaciji". Mag. teza, Univerzitet u Beogradu, Matematički fakultet, 2008

W3C. "Extensible Markup Language (XML)", 2020a. Accessed July 31, 2020. https://www.w3.org/XML/

W3C. "Linked Data", 2020b. Accessed July 31, 2020. https://www.w3.org/standards/semanticweb/data

Андоновски, Јелена. "Мрежа отворених података и језички ресурси у процеси+у изградње српско-немачког литерарног корпуса". Докт. дисертација, Универзитет у Београду, Филолошки факултет, 2019

Обрадовић, Доситеј. *Сабрана дела : 1811-1961. 1-3.* Београд: Просвета, 1961

Обрадовић, Доситеј. *Писмо Хparaлампију. Живот и приključenija.* Београд: Задужбина „Доситеј Обрадовић", 2007. URL https://dositej.unilib.rs/

Тртовац, Александра. "Дескриптори метаподатака и дескриптори садржаја у проналажењу информација у дигиталним библиотекама". Докт. дисертација, Универзитет у Београду, Филолошки факултет, 2016