



Multilingual terminological resources and services used for supporting OER multilinguality

DEV 3.3



BAEKTEL

Blending academic and entrepreneurial knowledge
in technology enhanced learning

*Blending **Academic** and **Entrepreneurial** Knowledge
in Technology enhanced learning – BAEKTEL*

www.baektel.eu

Project no.:

544482-TEMPUS-1-2013-1-IT-TEMPUS-JPHES

Project acronym:

BAEKTEL

Project full title:

**Blending Academic and Entrepreneurial Knowledge
in Technology Enhanced Learning**

Start date of project: 2013-12-01

Duration:

3 years

DEV 3.3

**Multilingual terminological resources and services
used for supporting OER multilinguality**

Due delivery date: 2016-10-31

Actual delivery date: 2016-10-31

Organisation name of lead participant for this deliverable:

University of Belgrade (UB)

Project funded by the TEMPUS IV programme		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Deliverable number:	3.3
Deliverable name:	Multilingual terminological resources and services used for supporting OER multilinguality
Work package:	WP3– Development of BAEKTEL infrastructure
Lead participant:	University of Belgrade

Author (s) – in alphabetical order		
Name	Organisation	e-mail
Ranka Stanković	University of Belgrade	ranka.stankovic@baektel.eu
Ivan Obradović	University of Belgrade	ivan.obradovic@baektel.eu
Olivera Kitanović	University of Belgrade	olivera.kitanovic@baektel.eu
Bojan Zlatic	University of Belgrade	bojan.zlatic@baektel.eu
Dalibor Vorkapić	University of Belgrade	dalibor.vorkapic@baektel.eu

Contents

1	Introduction	5
2	The language support system overview	5
3	Multilingual resources.....	10
3.1	Computer-lexicographical process for online dictionaries	10
3.2	Terminological standards and guidelines.....	11
3.2.1	TBX as standard for data exchange	12
3.2.2	TMX as standard for data exchange.....	15
3.3	Translation tools.....	16
3.3.1	Tools overview	16
3.3.2	OmegaT CAT tool	18
3.3.3	MemoQ	21
3.3.4	SDL Trados.....	22
4	Services for multilingual support.....	23
4.1	Termi.....	23
4.1.1	Browse/ Prelistavanje.....	24
4.1.2	Search/ Pretraga	26
4.1.3	Get account for Termi/ Dobijanje naloga za Termi.....	27
4.1.4	Writing/ Ažuriranje.....	28
4.1.5	Bibliographic sources/ Bibliografski izvori.....	34
4.1.6	User profiles/ Korisnički profili	35
4.1.7	Tips and tricks.....	37
5	OER tagging.....	39
5.1	Elements for BMP: Description and cataloguing guidelines.....	39
5.2	Annotation graph.....	42
5.3	Annotated text.....	43
6	Conclusions.....	44
7	References.....	45

1 Introduction

Due to the combined effect of globalization and European integration the need for support of multilingual business and academic communication is growing. The language barrier is the biggest obstacle that stands in the way of broader usability of OERs. Its immediate effect on the approach adopted in BAEKTEL was to provide OER materials in various languages. It is an efficient answer to the main criticism of OER that they are biased towards the so called "big" languages. In addition to that, in order to offer support in terminology, the BAEKTEL platform provides electronic terminological resources, parallel (multilingual) corpora of lessons and texts in written form, and functionalities for searching and browsing of terminological resources and using them for text annotation.

Multilingual terminological resources and services used for supporting OER multilinguality are an important element of the BAEKTEL OER software solution.

This document describes the linguistic component of the system, i.e. the resources and tools to be used to expand the usability of BAEKTEL OERs and improve their visibility on the Internet.

Zbog kombinovanog dejstva globalizacije i evropskih integracija potreba za podrškom višejezičnoj poslovnoj i akademskoj komunikaciji raste. Jezičke barijere su najveća prepreka koja stoji na putu šire upotrebljivosti otvorenih obrazovnih resursa. Njihov neposredni efekat na pristup usvojen u BAEKTEL-u je da se obezbede OER materijali na različitim jezicima. To je efikasan odgovor na glavnu kritiku OER da su pristrasni prema takozvanim "velikim" jezicima. Pored toga, u cilju pružanja podrške u terminologiji, BAEKTEL platforma pruža elektronske terminološke izvore, paralelne (višejezične) korpuse lekcija i tekstova u pisanoj formi, i funkcionalnosti za pretraživanje i pregledanje terminoloških resursa i njihovo korišćenje za označavanje teksta.

Višejezični terminološki resursi i usluge koje se koriste za podršku višejezičnosti OER su važan element softverskog rešenja za BAEKTEL OER.

U ovom dokumentu je opisana lingvistička komponenta sistema, odnosno resursi i alati koji se koriste da se proširi upotrebljivost BAEKTEL OER i poboljša njihova vidljivost na Internetu.

2 The language support system overview

Learning materials within BAEKTEL network are available in different languages. Besides WB languages, English and Russian are also envisaged, as well as other languages, for which electronic language resources are available, particularly terminology. This multilingual approach is a combined effect of globalization and European integration, and represents a holistic approach that takes into account all the languages a learner may use,

Obrazovni materijali u BAEKTEL mreži su dostupni na različitim jezicima. Pored jezika sa prostora Balkana, takođe su predviđeni engleski i ruski, kao i drugi jezici za koje postoje dostupni elektronski jezički resursi, prvenstveno terminološki. Ovaj višejezični pristup je kombinovani efekat globalizacije i evropskih integracija, i predstavlja holistički pristup koji uzima u obzir sve jezike koje učenik može koristiti, za razliku od

as opposed to the more traditional approach looking at one language at a time (Cenoz & Gorter, 2011).

Specifically, providing language and education resources in WB languages efficiently answers one of the major criticisms of OER, namely that they are biased towards the so called "big" languages. In order to cope successfully with multilinguality within the network, especially when terminology is concerned, a language support system is developed within the BAEKTEL metadata portal. Besides sustaining expert terminology in a multilingual environment this language support system will also improve the search and browse functions of BMP.

The BMP languages support system, whose structure is outlined in Figure 1, is based on electronic language resources, namely, lexical resources, textual resources and grammars. The simplest multilingual lexical resources in general are bilingual dictionaries in electronic form. However, for their full functionality in languages with complex morphology, such as Serbian, Bosnian and Montenegrin language, they need to be coupled with language specific morphological dictionaries.

Morphological dictionaries in the so-called LADL format (Krstev, Stanković, Obradović, Vitas, & Utvić, 2010) are thus a necessary part of the lexical resources used by the BMP language support system. Besides Serbian, such resources exist for many other languages, including English and Russian, which are also envisaged as OER languages within our network. Besides morphological dictionaries, for full functionality of the language support system grammars are needed as well, and they are implemented by the so called finite state automata, finite state transducers and compound inflection rules

tradicionalnijeg pristupa gde se svaki jezik posmatra ponaosob (Cenoz & Gorter, 2011).

Posebno, publikovanje jezičkih i obrazovnih resursa na jezicima zapadnog Balkana daje odgovor na jednu od glavnih kritika OER, da su izrazito naklonjeni takozvanim "velikim" jezicima. Kako bi se uspešno implementirala višejezičnost unutar mreže, posebno kada je u pitanju terminologija, razvijen je sistem za lingvističku podršku u okviru portala BAEKTEL metapodataka. Osim podržavanja stručne terminologije u višejezičnom okruženju ovaj sistem takođe unapređuje pretragu i prelistavanje na portalu metapodataka.

Sistem za jezičku podršku na portalu metapodataka, čija je struktura prikazana na slici 1, se zasniva na elektronskim jezičkim resursima, konkretno, gramatikama, leksičkim i tekstualnim resursima. Najjednostavniji višejezični leksički resursi su dvojezični rečnici u elektronskom obliku. Međutim, za njihovu punu funkcionalnost u jezicima sa složenom morfologijom, kao što su srpski, bosanski i crnogorski jezik, oni se moraju kombinovati sa jezički specifičnim morfološkim rečnicima.

Morfološki rečnici u tzv. LADL formatu (Krstev, Stanković, Obradović, Vitas, & Utvić, 2010) su tako neophodan deo leksičkih resursa koje koristi BMP sistem za jezičku podršku. Osim za srpski, takvi resursi postoje i za mnoge druge jezike, uključujući ruski i engleski, koji su predviđeni kao jezici za obrazovne resurse u našoj mreži. Osim morfoloških rečnika, za punu funkcionalnost ovog sistema potrebne su i lokalne gramatike, koje su implementirane takozvanim konačnim automatima, transduktorima, i pravilima za određivanje flektivnih svojstava polileksičkih jedinica (Krstev, Processing of Serbian – Automata, Texts and Electronic dictionaries, 2008).

(Krstev, Processing of Serbian – Automata, Texts and Electronic dictionaries, 2008).

Another important lexical resource offering support for multilingual terminology is the Serbian wordnet. In brief, a wordnet consists of sets of synonymous words representing specific concepts, called synsets. A semantic network is formed on basis of semantic relations between them.

Akin to standard dictionaries, each synset word, or literal, is composed of a literal string and a sense tag, representing the sense of the literal string specific to that particular concept. The multilinguality potential of wordnets stems from the interlingual index (ILI), which establishes relations between synsets representing the same concept in different languages (Krstev, Pavlović-Lažetić, Vitas, & Obradović, 2004).

Drugi važan leksički resurs koji pruža podršku višejezičnoj terminologiji je srpska mreža reči (vordnet). Ukratko, vordnet se sastoji od skupa sinonima koji predstavlja određeni pojam, a koji se nazivaju sinseti. Na osnovu semantičkih relacija između njih formira se semantička mreža.

Kao kod standardnih rečnika, svaka reč u sinsetu, ili literal, sastoji se od niske znakova i oznake koja predstavlja značenje niske znakova za taj konkretni pojam. Višejezičnost vordneta proizlazi iz međujezičkog indeksa (ILI), koji uspostavlja relacije između sinseta koji predstavljaju isti pojam na različitim jezicima (Krstev, Pavlović-Lažetić, Vitas, & Obradović, 2004).

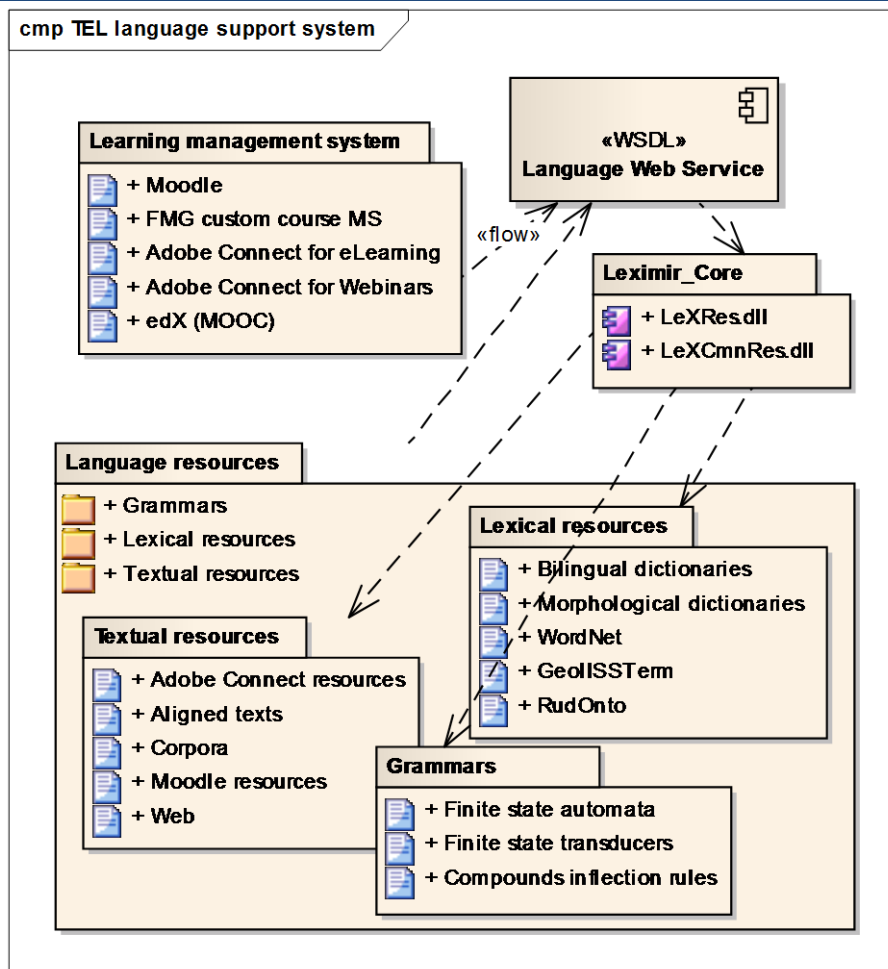


Figure 1: The BMP language support system

Slika 1. Jezička podrška za BMP

Finally, the language support system features domain specific terminological resources such as GeolISSTerm for geology and RudOnto for mining (Stanković, Obradović, Kitanović, & Kolonja, Building Terminological Resources in an e-Learning Environment, 2012). GeolISS is a thesaurus of geological terms with entries in Serbian and English, developed at University of Belgrade, Faculty of Mining and Geology within the GeolISS project (<http://geoliss.mre.gov.rs>).

Konačno, sistem za jezičku podršku obuhvata i terminološke resurse specifične za pojedine domene kao što su GeolISSTerm za geologiju i RudOnto za rudarstvo (Stanković, Obradović, Kitanović, & Kolonja, Building Terminological Resources in an e-Learning Environment, 2012). GeolISS je tezaurus geoloških termina sa odrednicama na srpskom i engleskom, razvijen na Rudarsko-geološkom fakultetu Univerziteta u Beogradu, u okviru projekta GeolISS (<http://geoliss.mre.gov.rs>).

Thesauruses are complex terminological resources, usually related to a specific domain, with a semantic structure formed by semantic relations between terms, and thus in some of their features resembling wordnets.

Tezaurusi su kompleksni terminološki resursi, koji se obično odnose na određenu oblast, sa semantičkom strukturom koju kreiraju semantičke relacije između termina, nalik vordnetu.

RudOnto (<http://rudonto.rgf.bg.ac.rs/>) is another complex terminological resource, also developed at University of Belgrade, Faculty of Mining and Geology. The ultimate goal is to gradually evolve RudOnto into the reference Serbian resource in e-format for mining terminology. Currently RudOnto comprises of concepts in Serbian, some with their English equivalents, and a small number of equivalents in other languages.

The BMP language support system also offers the possibility of storing specific textual resources, such as aligned texts and corpora. Aligned texts are pairs of texts in different languages, mainly an original and its translation, aligned on some structural level, most often the sentence. Aligned texts in BMP are in the standard, Translation Memory eXchange (TMX) format, which is XML-compliant.

Corpora are large and structured sets of texts, both monolingual and multilingual, the latter often composed of aligned texts. Finally, the World Wide Web itself represents a textual resource that BMP language support system makes use of.

The language support system handles various types of requests issued by users, usually in the form of a query. The requests are handled by WSDL (Web Services Description Language) described Language Web Service, basically composed of a web application and a web service. When handling a query, the Language Web Service uses a component of the multipurpose tool for lexical resources management and query expansion LeXimir, developed at University of Belgrade, Faculty of Mining and Geology (Stanković, Obradović, Krstev, & Vitas, 2011). Queries are accepted by the web application, which forwards it to the web service. The web service then invokes LeXimir_Core, a function library of LeXimir,

RudOnto je još jedan kompleksan terminološki resurs, takođe razvijen na Rudarsko-geološkom fakultetu Univerziteta u Beogradu. Krajnji cilj je da se RudOnto postepeno razvije u referentni elektronski terminološki resurs za oblast rudarstva. Trenutno RudOnto sadrži pojmove na srpskom jeziku, od kojih deo ima prevod na engleski, a sasvim mali broj ima prevode i na drugim jezicima.

Sistem za jezičku podršku BMP omogućava skladištenje specifičnih tekstualnih resursa, kao što su paralelni (poravnati) tekstovi i korpusi. Paralelni tekstovi su parovi tekstova na različitim jezicima, uglavnom original i njegov prevod, poravnati na nekom strukturnom nivou, najčešće rečenici. Paralelni tekstovi na BMP su kreirani u XML formatu saglasno standardu za prevodilačke memorije TMX (Translation Memory eXchange).

Korpusi su veliki i strukturirani skupovi tekstova, mogu biti jednojezični i višejezični, pri čemu se višejezični obično sastoje od paralelnih tekstova. Konačno, sam World Wide Web se može posmatrati kao tekstualni resurs koji koristi sistem za jezičku podršku BMP.

Sistem za jezičku podršku obrađuje različite vrste korisničkih zahteva, najčešće u obliku upita. Zahteve obrađuje jezički veb servis opisan pomoću WSDL (Web Services Description Language) jezika, a koji se sastoji od veb aplikacije i veb servisa. Pri rukovanju upitima, jezički veb servis koristi komponentu višenamenskog alata za upravljanje leksičkim resursima i ekspanziju upita LeXimir, koji je razvijen na Rudarsko-geološkom fakultetu Univerziteta u Beogradu, (Stanković, Obradović, Krstev, & Vitas, 2011). Upite prihvata veb aplikacija, koja ih potom prosleđuje veb servisu. Veb servis onda poziva LeXimir_Core, biblioteku funkcija

which processes the query using the available resources.

During this processing the system can perform a morphological expansion of the query to improve recall, which is especially important for morphologically rich languages such as Serbian, Bosnian and Montenegrin. In order to support the multilinguality of the BMP, the language support system can also expand a query in one language to another language, e.g. a query in Serbian to English or Russian, and vice-versa.

With all the aforementioned features the language support system takes a prominent place within the BMP, offering invaluable support for better understanding and handling of the multilingual OER content.

LeXimir-a, koji obrađuje upit koristeći raspoložive resurse.

Tokom ove obrade sistem može da proširi upit morfološki kako bi se poboljšao odziv, što je posebno važno za morfološki kompleksne jezike, kao što su srpski, bosanski i crnogorski. U cilju podrške višejezičnosti BMP, sistem za jezičku podršku može proširiti upit i sa jednog jezika na drugi jezik, npr. da proširi upit zadat na srpskom na engleski ili ruski, i obrnuto.

Uz sve navedene funkcije podrške sistem za jezičku podršku zauzima značajno mesto u BMP, nudeći dragocenu podršku boljem razumevanju i rukovanju višejezičnim OER sadržajem.

3 Multilingual resources

3.1 Computer-lexicographical process for online dictionaries

Three phases of any lexicographical process can be distinguished: planning, writing, producing, each following one another. In practice, working on a dictionary usually (but especially for long-term academic projects) means that while still writing the dictionary, new material may be added to the corpus, corrections for entries already published are gathered, headwords and cross-references will be supplemented, and modifications in the original concept will become necessary. Permanent corrections and actualizations are usually part of a dictionary's writing phase (Klosa, 2013).

The phases of planning and writing, thus, do not really follow each other: for some time, they run parallel. Only the production phase of paper dictionaries or completed electronic dictionaries will not start before the other phases are completed.

Možemo razlikovati tri osnovne faze bilo kog leksikografskog procesa koje se odvijaju jedna za drugom: planiranje, pisanje, proizvodnja. U praksi, rad na rečniku obično (a naročito kod dugoročnih akademskih projekata) znači da se, dok se piše rečnik, može dodavati novi materijal u korpus, mogu prikupljati korekcije za već objavljene odrednice se, dopunjavati rečničke odrednice i unakrsno referenciranje, pri čemu modifikacije u originalnim konceptima postaju neophodne. Stalne korekcije i osavremenjavanje su obično deo faze pisanja rečnika (Klosa, 2013).

Faze planiranja i pisanja, dakle, ne slede u potpunosti jedna za drugom, već se delimično odvijaju paralelno. Jedino faza proizvodnje papirnog rečnika ili upotpunjenog elektronskog rečnika ne otpočinje pre završetka ostalih faza.

Producing an online dictionary may begin before the phase of writing is finished: online dictionaries can be published step-by-step.

Izrada online rečnika može početi pre završetka faze pisanja: online rečnici se mogu objavljivati deo-po-deo.

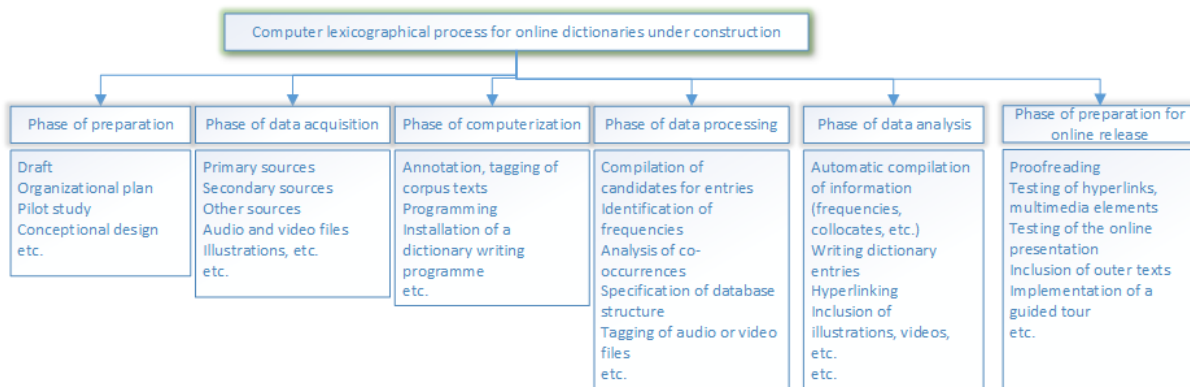
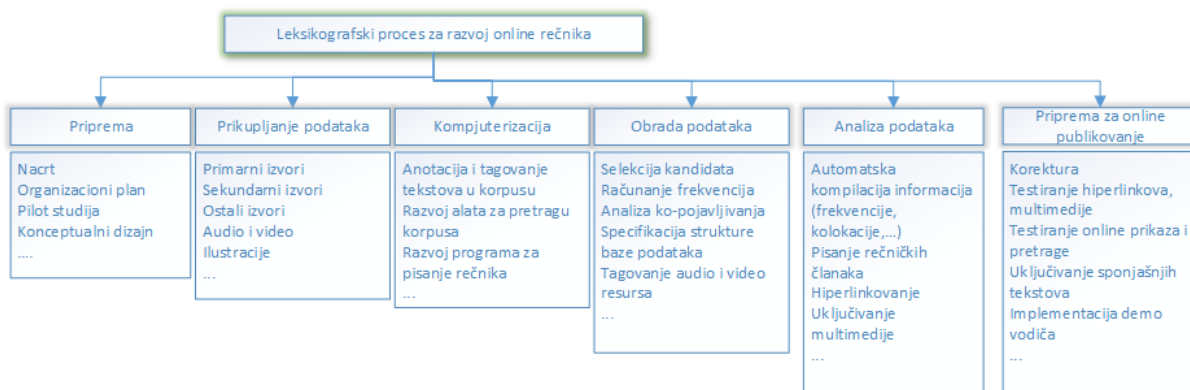


Figure 2 Computer lexicographical process for online dictionaries under construction



Slika 2. Leksikografski proces za razvoj online rečnika

Thus, all phases of the computer-lexicographical process (planning & writing & producing) merge, giving yet unknown flexibility to the lexicographer.

Dakle, sve faze računarski podržanog leksikografskog procesa (planiranje, pisanje i proizvodnja) se spajaju, omogućavajući leksikografu do sada nepoznatu fleksibilnost u radu.

3.2 Terminological standards and guidelines

The standards related to terminology management and translational technologies are commonly used by the localization and translation industry as well as by public translation and terminology units and organizations. This is the reason why terminological standards are probably the most widespread and consolidated comparing to the other language technology standards.

Standardi u oblasti upravljanja terminologijom i prevodilačkih tehnologija se najčešće koriste za lokalizaciju i u prevodilačkoj industriji, a koriste ih kako samostalni prevodioci tako i oni u prevodilačkim društvima i preduzećima. Ovo je razlog što je primena terminoloških standarda verovatno najrasprostranjenija i najjednačenija u poređenju sa standardima

Standards are used to enable portability, reusability and consistency, which is especially important given the variety of tools used for Computer Aided Translation (CAT).

Organizations responsible for main language resources standards are International Organization for Standardization (ISO) and European Telecommunications Standards Institute (ETSI). ISO/TC 37 – Technical Committee on Terminology and Other Language and Content Resources is responsible for developing and maintaining standards. One of those standards is widespread used ISO 30042:2008 Systems to manage terminology, knowledge and content – TermBase eXchange (TBX). It was prepared by Localization Industry Standards Association (LISA), special interest group, OSCAR (Open Standards for Container/Content Allowing Re-use) and adopted by ISO/TC 37 Subcommittee.

LISA OSCAR was also responsible for Translation Memory eXchange (TMX) standard. In 2011 LISA was taken over by European Telecommunications Standards Institute (ETSI).

koji su namenjeni drugim oblastima jezičkih tehnologija.

Standardi se koriste kako bi se postigli prenosivost, ponovno korišćenje i konzistentnost podataka, što je naročito važno zbog raznovrsnosti alata za računarski potpomognuto prevođenje (CAT, Computer Aided Translation).

Organizacije odgovorne za glavne standarde u oblasti jezičkih resursa su Međunarodna organizacija za standardizaciju (ISO) i Evropski institut za telekomunikacione standarde (ETSI). ISO/TC 37 Tehnički komitet za terminologiju i druge jezike i sadržajne resurse je odgovoran za razvoj i održavanje standarda. Jedan od tih standarda je široko korišćeni ISO 30042:2008 Sistemi za upravljanje terminologijom, znanjem i sadržajem – TermBase eXchange (TBX). Ovaj standard je pripremio udruženje Localization Industry Standards Association (LISA) i specijalna interesna grupa OSCAR (Open Standards for Container/ Content Allowing Re-use), a usvojio je podkomitet ISO/TC 37.

Udruženje LISA OSCAR je takođe bilo odgovorno za Translation Memory eXchange (TMX) standard. Od 2011. godine LISA-om upravlja European Telecommunications Standards Institute (ETSI).

3.2.1 TBX as standard for data exchange

Term Base eXchange (TBX) defines an XML-based framework for representing terminological data in structured form, used for terminological data interchange, dissemination, analysis and representation. TBX is designed to support the flow of terminological data between technologies and systems, their integration from multiple sources and data conversion in case of change in applications or technologies. It is

Term Base eXchange (TBX) definiše okvir za predstavljanje terminoloških podataka u strukturiranoj formi zasnovan na XML-u. Koristi se za razmenu, diseminaciju, analizu i predstavljanje terminoloških podataka. TBX je osmišljen sa idejom da podrži tokove terminoloških podataka između tehnologija i sistema, njihovo uklapanje iz različitih izvora i konverziju podataka u slučaju promena u aplikacijama ili tehnologijama. U okviru njega

possible to define a variety of Terminological Markup Languages (TMLs) within this framework by selecting data-categories.

TBX TML includes two modules, a core structure and a set of data-categories and their constraints represented as an eXtensible Constraint Specification (XCS). Both modules are defined with the appropriate DTDs (Document Type Definition).

There are four general types of data-category:

- a core structure module data-category – any data-category defined in structure module DTD as an XML element.
- a meta data-category – core structure data-category that takes a type attribute. It is used for grouping similar data categories. The type attribute values are defined according to ISO 12620 (ISOcat - a Data Category Registry).
- a data-category implemented using an attribute – it is a terminological data-category that is defined according to ISO 12620 and one that is specified as value of the name attribute in the default XCS file.
- a data-category implemented as the content of an element – it is simple data category, represents one value of a closed set of values defined in an XCS file.

Hierarchy of TMX document consists of a <martiHeader> element and a <text> element, according to MARTIF (Machine-readable terminology interchange format).

The <martiHeader> element consists of a description of data collection in <fileDesc> element, info about used XCS file in <encodingDesc> element and the history of

je moguće definisati mnoštvo različitih terminoloških jezika za obeležavanje (TML) odabirom kategorija podataka.

TBX TML uključuje dva modula, jezgro i skup kategorija podataka i njihova ograničenja predstavljena kao eXtensible Constraint Specification (XCS). Oba modula su definisana pomoću odgovarajućeg DTD-a (Document Type Definition).

Postoje četiri glavna tipa kategorije podataka:

- kategorija podataka jezgra strukture modula – bilo koja kategorija podataka deklarirana u DTD-u kao XML element
- kategorija metapodataka – osnovna struktura koja uključuje atribut type. Koristi se za grupisanje sličnih kategorija podataka. Vrednosti atributa type su definisane u skladu sa standardom ISO 12620 (ISOcat - a Data Category Registry).
- kategorija podataka implementirana korišćenjem atributa – to je terminološka kategorija podataka koja je definisana u skladu sa ISO 12620 i koja je specifikovana kao vrednost imena atributa u podrazumevanom XCS dokumentu.
- kategorija podataka implementirana kao sadržaj elementa – to je prosta kategorija podataka, koja uzima vrednost iz zatvorenog skupa vrednosti definisanih u XCS datoteci.

Hijerarhija TMX dokumenta sastoji se od <martiHeader> i <text> elemenata, u skladu sa MARTIF (Machine-readable terminology interchange format).

Element <martiHeader> sastoji se od kolekcije podataka u okviru elementa <fileDesc>, informacije o korišćenju XCS datoteke u <encodingDesc> elementu i

revisions to the collection in the <revisionDesc> element.

The <text> element consists of terminological entries surrounded by <body> element and complementary information in <back> element.

Each terminological entry in the <body> is a <termEntry> and it follows the structure of TMF metamodel (ISO 16642). It has three levels, entry-level, language-level and term-level.

TBX is widely supported and recognized by tool providers, companies and public organizations as one of concrete results of Terminological Markup Framework (TMF) or ISO 16642:2003 which exists in XML, UML and in RDF.

There is also TBX-Basic, lighter version of TBX. It is adjusted to language applications that require soft approach to terminology management. Its purpose is formalizing the translation in an XML in order to increase the ability to exchange language resources between users and to use it in different electronic environments.

pregleda promena u kolekciji u okviru elementa <revisionDesc>.

Element <text> se sastoji od terminološkog unosa okruženog etiketom <body> i dopunskih informacija u okviru elementa <back>.

Svaki pojedinačni unos termina se unosi u element <termEntry> i prati strukturu TMF metamodela (ISO 16642). Svaki unos ima tri nivoa, nivo unosa, nivo jezika i nivo termina.

TBX je široko podržan i prepoznat od strane proizvođača alata, firmi i javnih organizacija kao jedan od konkretnih rezultata Terminological Markup Framework (TMF) ili ISO 16642:2003 standarda koji postoji u XML, UL i RDF formatima.

Postoji i TBX-Basic, jednostavnija verzija TBX-a koja je prilagođena jezičkim aplikacijama koje zahtevaju blaži pristup upravljanju terminologijom. Njegova svrha je da formalizuje prevod u XML-u kako bi se povećala sposobnost razmene jezičkih resursa među korisnicima, kao i njihovo korišćenje u različitim elektronskim okruženjima.

```
<?xml version="1.0" encoding="utf-8" standalone="no" ?>
<!DOCTYPE martif SYSTEM "TBXcoreStructV02.dtd">
<martif type="TBX" xml:lang="en">
  <martifHeader>
  </martifHeader>
  <text>
    <body>
      <termEntry id="c110691">
        <descrip type="relatedConceptNarrower">aritmetika</descrip>
        <descrip type="relatedConceptNarrower">algebra</descrip>
        <descrip type="relatedConceptNarrower">analiza</descrip>
        <descrip type="relatedConceptNarrower">geometrija</descrip>
        <langSet xml:lang="sr">
          <ntig>
            <termGrp>
              <term>matematika</term>
              <termNote type="termType">entryTerm</termNote>
            </termGrp>
            <descripGrp>
              <descrip type="definition">Matematika je egzaktna nauka koja se na osnovu deduktivnog zaključivanja bavi izučavanjem apstraktnih pojmova (brojeva, geometrijskih figura, funkcija, prostora..) i odnosa među njima. Matematika se izučava kroz naučne discipline kao što su aritmetika, algebra, geometrija, analiza...</descrip>
            </descripGrp>
          </ntig>
        </langSet>
        <langSet xml:lang="en">
          <descrip type="relatedConceptNarrower">arithmetic</descrip>
          <descrip type="relatedConceptNarrower">algebra</descrip>
          <descrip type="relatedConceptNarrower">mathematical analysis</descrip>
          <descrip type="relatedConceptNarrower">geometry</descrip>
          <ntig>
            <termGrp>
              <term>mathematics</term>
              <termNote type="termType">entryTerm</termNote>
            </termGrp>
          </ntig>
        </langSet>
      </termEntry>
    </body>
  </text>
</martif>
```

Figure 3. TBX example

Slika 3. TBX primer

3.2.2 TMX as standard for data exchange

Translation memory eXchange is vendor-neutral XML-based standard for the exchange of Translatory Memory (TM) data created by CAT. It provides easier exchange of TM data between tools and/or translation vendors. During that process information loss is little or none. This allows tools users to not be cornered in using a specific tool.

TM is a language technology that enables the translation of segments (sentences, paragraphs, or phrases) of documents. It is based on search for similar segments in a database and suggesting matches that are found in the database.

TMX file should be validated against the appropriate DTD. It is enclosed in <tmx>

Translation memory eXchange je otvoreni XML standard za razmenu prevodilačkih memorija (TM) kreiranih CAT tehnologijom. On omogućava lakšu razmenu TM među alatima i/ili prevodiocima. Tokom tog procesa ne dolazi ili skoro da ne dolazi do gubljenja informacija. Ovo omogućava korisnicima alata da ne budu ograničeni na jedan alat.

TM je jezička tehnologija koja omogućava prevođenje segmenata (rečenica, pasusa ili fraza) dokumenta. Zasnovana je na pretrazi sličnih segmenata u bazi podataka i predlaganju prevoda koji su pronađeni u bazi.

Validnost TMX datoteka treba da se proveri u odnosu na predefinisani DTD. On se nalazi

root element that contains two elements, <header> and <body>.

The <header> element contains meta-data about the document. In addition to its attributes (required creationtool, creationtoolversion, segtype, o-tmf, adminlang, srclang and datatype) it can also store information about document in <note> and <prop> elements.

The <body> element in TMX document consists of <tu> (Translation Unit) and <tuv> (Translation Unit Variant) elements, where each <tuv> is a segment in one of the translation languages. Figure 4 illustrates a single aligned segment (TU) of a document.

u <tmx> korenom elementu koji sadrži elemente <header> i <body>.

Element <header> sadrži metapodatke o dokumentu. Uz attribute (obavezni creationtool, creationtoolversion, segtype, o-tmf, adminlang, srclang and datatype) može da skladišti i informacije o dokumentu u elementima <note> i <prop>.

Element <body> u TMX dokumentu počiva na <tu> (prevodilačkoj jedinici) i <tuv> (varijanta prevodilačke jedinice) elementima, gde je svaki <tuv> zapravo segment na jednom od jezika koji se koriste za prevođenje. Slika 4 prikazuje jedan poravnati segment <tu> dokumenta.

```
<tu>
  <prop type="Domain">BAEKTEL 1.2, 2014, ID: 5.2014.1.2</prop>
  <tuv xml:lang="en" creationid="n3 " creationdate="20150327T152832Z">
    <seg>Open educational resources policies are principles that support the use of
    open content and practices in educational institutions, especially open
    educational resources (OER), and are adopted mostly by governing bodies. </seg>
  </tuv>
  <tuv xml:lang="sr" creationid="n3 " creationdate="20150327T152832Z">
    <seg>Politiku otvorenih obrazovnih resursa čine principi i politike koji
    podržavaju korišćenje otvorenog sadržaja i prakse u obrazovnim institucijama,
    posebno otvorenih obrazovnih resursa (OER), i uglavnom su usvojeni od strane
    upravljačkih tela. </seg>
  </tuv>
</tu>
```

Figure 4 Aligned segment (TU)

Slika 4. Poravnati segment <tu>

3.3 Translation tools

3.3.1 Tools overview

The automatic machine translation systems available today are not able to produce high-quality translations unaided: their output must be edited by a human to correct errors and improve the quality of translation. Computer-assisted translation (CAT) incorporates a manual editing stage, making translation an interactive process between human and computer (Barrachina & al., 2009).

Sistemi za automatsko mašinsko prevođenje koji su danas dostupni još uvek nisu u stanju da proizvedu kvalitetne prevode bez korekcije prevodioca (čoveka) koji treba da ispravi greške i poboljša kvalitet prevoda. Računarski podržano prevođenje (CAT) uključuje fazu ručnog editovanja, tako da prevođenje postaje interaktivni proces između čoveka i računara (Barrachina & al., 2009).

Some computer-assisted translation solutions include controlled machine translation, with complex set of tools which may include terminology management features and various other linguistic tools. Carefully customized user dictionaries based on correct terminology significantly improve the accuracy of MT, and thus, aim at increasing the efficiency of the entire translation process.

Terminology managers allow translators to manage their own terminology bank in an electronic form. This can range from a simple table created in the translator's word processing software or spreadsheet, a database, and even more robust solutions, specialized software packages.

Full-text search tools (or indexers) allow the translator to query already translated texts of various kinds that might be useful.

Concordancers are programs that retrieve instances of a word or an expression and their respective context in a monolingual, bilingual or multilingual corpus, such as a biterm or a translation memory. Biterm aligners are tools that align a source text and its translation which can then be analyzed using a full-text search tool or a concordance.

Project management software allows linguists to structure complex translation projects, assign the various tasks to different people, and track the progress of each of these tasks.

Translation memory tools (TM tools), consists of a database of text segments in a source language and their translations in one or more target languages. In this document we will present opensource tool OmegaT and mention two commercial software solutions: MemoQ and Trados, most frequently used software of this type.

Neki softveri za računarski pomognuto prevođenje uključuju kontrolisano mašinsko prevođenje, sa složenim skupom alata koji može uključiti upravljanje terminologijom i razne druge jezičke alate. Pažljivo prilagođeni korisnički (lični) rečnici zasnovani na ispravnoj terminologiji značajno poboljšavaju preciznost mašinskog prevođenja, čime se postiže povećanje efikasnosti celokupnog procesa prevođenja.

Alati za upravljanje terminologijom omogućavaju prevodiocu da kreira sopstvenu terminološku banku u elektronskom obliku. Formati mogu da budu različite kompleksnosti: od jednostavne tabele kreirane u softveru za uređivanje teksta ili tabela, preko baze podataka, do još robusnijih rešenja, specijalizovanih programskih paketa.

Alati za pretraživanje punog teksta (ili indeksi) omogućavaju prevodiocu da pretražuje prethodno prevedene tekstove koji mu mogu biti od koristi.

Konkordansi su programi koji pronalaze primere upotrebe reči ili izraza i prikazuju njihov kontekst u jednojezičnim, dvojezičnim ili višejezičnim korpusima, kao što su biterm ili prevodilačke memorije. Biterm paralelizatori su alati koji poravnavaju izvorni tekst i prevod, koji se potom mogu analizirati pomoću alata za pretraživanje punog teksta ili konkordansi.

Softver za upravljanje projektima omogućuje strukturiranje složenih prevodilačkih projekata, dodeljivanje zadataka različitim ljudima, i praćenje napretka zadataka.

Alati sa prevodilačkim memorijama (TM alati), se sastoje od baze podataka segmenata teksta u izvornom jeziku i njihovih prevoda u jednom ili više ciljnih jezika. U ovom dokumentu ćemo predstaviti alat otvorenog koda OmegaT i spomenuti dva komercijalna softvera: MemoQ i Trados, koji su najčešće korišćeni alati ovog tipa.

3.3.2 OmegaT CAT tool

OmegaT is a computer-assisted translation tool written in the Java programming language

OmegaT is intended for professional translators. Its features include customisable segmentation using regular expressions, translation memory with fuzzy matching and match propagation, glossary matching, dictionary matching, translation memory and reference material searching, and inline spell-checking using Hunspell spelling dictionaries.

OmegaT handles a translation job as a project, a hierarchy of folders with specific names. The user copies non-translated documents into one named `/source/` (or subfolders thereof). The Editor pane displays the source documents as individual “segments” for translation, one segment at a time. OmegaT, when directed, generates the (partially) translated versions in the `/target/` subfolder.

Other named folders include ones for automatic consultation within the program: `/tm/` for existing translation pairs in `.tmx` format, `/tm/auto/` for automatic translation of 100% matches, `/glossary/` for glossaries, `/dictionary/` for StarDict (and `.tbx`) dictionaries.

When the user goes to translate a segment in the Editor pane, OmegaT automatically searches the `.tmx` files in the `/tm/` hierarchy for previous translation pairs with similar source sentences and displays them in the Fuzzy Matches pane for insertion into the Editor pane with a keyboard shortcut. The Glossary and Dictionary panes provide similar automatic look-up functions for any glossaries and dictionaries in the corresponding named folders in the project. The optional Machine Translation pane shows machine translations from Google Translate and similar services.

Omega T je alat za računarski podržano prevođenje, napisan programskim jezikom Java.

Omega T je namenjen profesionalnim prevodiocima. Njegove karakteristike uključuju prilagodljivu segmentaciju pomoću regularnih izraza, prevodilačku memoriju sa fazi uparivanjem i propagacijom uparivanja, uparivanje sa glosarijem i rečnikom, prevodilačku memoriju i pretragu referentnog materijala uz neposrednu proveru pravopisa korišćenjem pravopisnih rečnika Hunspell.

Omega T upravlja poslom prevođenja kroz projekat, koji čini hijerarhija foldera sa specifičnim nazivima. Korisnik kopira neprevedene dokumente u jedan sa imenom `/source/` (ili njegove podfoldere). Urednički panel prikazuje izvorne dokumente kao individualne “segmente” koji se prevode jedan po jedan. Omega T, na zahtev, generiše (delimično) prevedene verzije u `/target/` podfolderu.

Ostali imenovani folderi uključuju one za automatsku konsultaciju unutar programa: `/tm/` za postojeće parove prevoda u `.tmx` formatu, `/tm/auto/` za 100% uparen prevod, `/glossary/` za glosare, `/dictionary/` za StarDict (i `.tbx`) rečnike.

Kada korisnik prevodi segment u Editor panelu, Omega T automatski pretražuje `.tmx` fajlove u `/tm/` hijerarhiji za prethodne parove prevoda sa sličnom izvornom rečenicom i prikazuje ih u Fuzzy Matches panelu za unos u Editor panelu prečicom na tastaturi. Paneli za glosar i rečnik pružaju slične automatske pretrage za bilo koji glosar i rečnik u odgovarajućem imenovanom folderu u projektu. Opcioni panel za automatsko prevođenje (Machine Translation) pokazuje mašinske prevode sa Google prevodioca i sličnih servisa.

When the user leaves a segment, OmegaT normally first adds the source-target pair to its database in memory. It subsequently saves that database to disk in Translation Memory eXchange (.tmx) format for use another day, in other projects, by other translators, and even with other CAT tools. No change, naturally enough, means no such update.

At any point, the user can create partially translated versions of the source files. Note that OmegaT copies source segments verbatim if they have yet to be translated. Before doing so, however, the user is advised to use the Validate menu command to check for tag and other errors.

OmegaT shares many features with mainstream CAT tools. These include creating, importing and exporting translation memories, fuzzy matching from translation memories, glossary look-up, and reference and concordance searching (keywords with context).

To create a project in OmegaT, select Project, then click New. The “Create a New Project” dialog appears. Select the desired source and target languages by clicking on the arrows to the right of these fields (Figure 5).

Kada korisnik napusti segment, Omega T obično prvo dodaje izvorno-ciljni par svojoj bazi podataka u memoriji. Nakon toga čuva tu bazu podataka na disku u Translation Memory eXchange (.tmx) formatu za upotrebu drugog dana, u drugim projektima, od strane drugih prevodilaca, čak i sa drugim CAT alatima. Ukoliko nema promene, nema ni ažuriranja.

U svakom trenutku korisnik može da kreira delimično prevedene verzije izvornih datoteka. Treba imati na umu da Omega T kopira izvore segmenta doslovno ako oni tek treba da budu prevedeni. Međutim, korisniku se savetuje da pre toga koristi komandu Validate menija da proveri etikete i druge greške.

Omega T ima mnogo sličnosti sa najpoznatijim CAT alatima. One obuhvataju kreiranje, uvoz i izvoz prevodilačkih memorija, fazi uparivanje iz prevodilačkih memorija, pretragu glosara, referenci i konkordansi (ključnih reči sa prikazanim kontekstom).

Da biste kreirali projekat u Omega T, selektujte Project pa kliknite New. Pojavljuje se dijalog “Create a New Project”. Izaberite željeni izvorni i ciljni jezik klikom na strelice sa desne strane odgovarajućih polja (slika5).

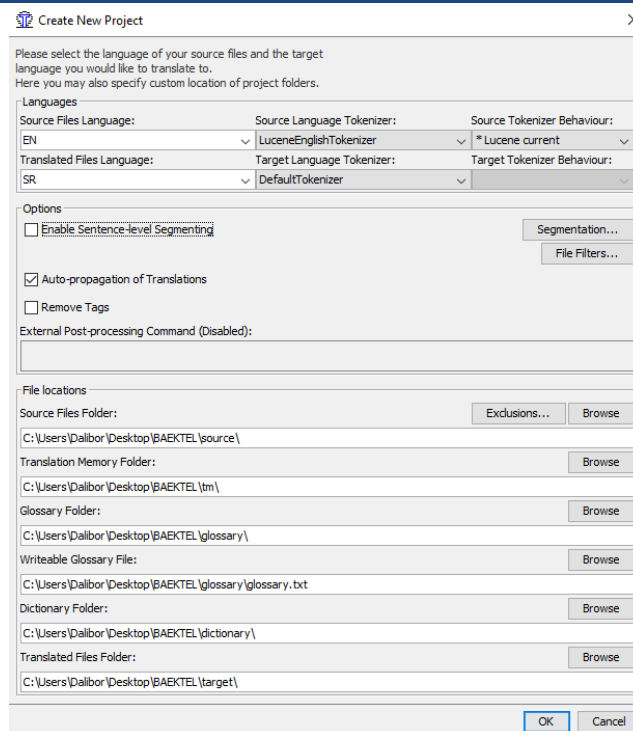


Figure 5. Create New Project

Slika 5. Kreiranje novog projekta

The project can be found in file manager: it is a folder with the name of the project, containing the sub-folders /dictionary, /glossary, /omegat, /source, /target, and /tm. They are all empty. You will put your document to be translated into the /source folder; the /target folder will be filled with your translated text; the /omegat folder need not concern you now. The /tm and /glossary files will remain empty until you create a .tmx file and a glossary and put them into those folders. In the /dictionary folder you can put bilingual dictionaries, in the appropriate format

After creating a project, a document for translation is imported. When it is loaded into the editor, duplicated text display appears: the original and the other is for translation. If there are certain segments of the translational memory, they will be automatically translated. All other segments are translated manually and added to the memory. (Figure 6)

Projekat se može naći i u fajl menadžeru: to je folder sa nazivom projekta, sadrži podfoldere /dictionary, /glossary, /omegat, /source, /target, i /tm. Na početku su svi prazni. Dokument koji se prevodi se smešta u /source folder, dok će /target folder biti punjen prevedenim tekstom. Folder /omegat nije bitan za početak. Folderima /tm i /glossary će biti prazni dok se ne kreiraju .tmx fajl i glosar i stave u odgovarajuće foldere. U /dictionary folder se opciono mogu staviti dvojezični rečnici u odgovarajućem formatu.

Nakon kreiranja projekta potrebno je importovati dokument za prevođenje. Kada se dokument učitava, u editoru se pojavi dupliran prikaz teksta: jedan je originalni, a drugi se prevodi. Ako određeni segmenti postoje u prevodilačkoj memoriji biće automatski prevedeni. Svi ostali segmenti se prevode ručno i dodaju se u memoriju. (Slika 6)

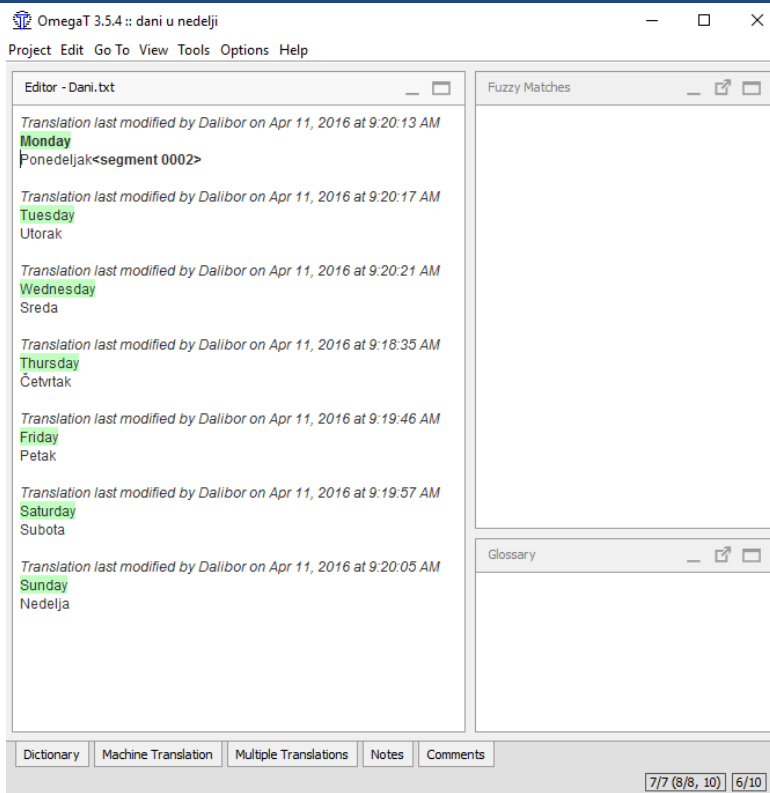


Figure 6. Translation in OmegaT

Slika 6. Prevođenje u OmegaT

After the completion of the translation it is necessary to create a translated document Project/ Create Translated Documents. The translated document is available in the /target directory.

Nakon završetka prevođenja potrebno je kreirati prevedeni dokument Project/ Create Translated Documents. Prevedeni dokument je dostupan u direktorijumu /target.

3.3.3 MemoQ

MemoQ is a computer-assisted translation software suite which runs on Microsoft Windows operating systems. It is developed by the Hungarian software company Kilgray Fordítástechnológiai Kft. (Kilgray Translation Technologies), a provider of translation management software established in 2004 (<https://www.memoq.com/about-us/background>) and cited as one of the fastest growing companies in the translation technology sector. MemoQ provides translation memory, terminology, machine translation integration and reference information management in desktop, client/

MemoQ je softverski paket za prevođenje uz pomoć računara prevođenje koji radi na Microsoft Windows operativnim sistemima. Razvijen je od strane mađarske softverske kompanije Kilgray Fordítástechnológiai Kft. (Kilgray Translation Technologies), koja proizvodi softver za upravljanje prevođenjem a koja je osnovana 2004. godine (<https://www.memoq.com/about-us/background>) i navodi se kao jedna od najbrže rastućih kompanija u sektoru prevodilačkih tehnologija. MemoQ obezbeđuje rad sa prevodilačkim memorijama, terminologijom, mašinsko

server and web application environments. (Figure 7).

prevođenje, integraciju i mogućnost upravljanja informacijama u desktop, klijent/server i veb aplikativnim okruženjima. (slika 7).

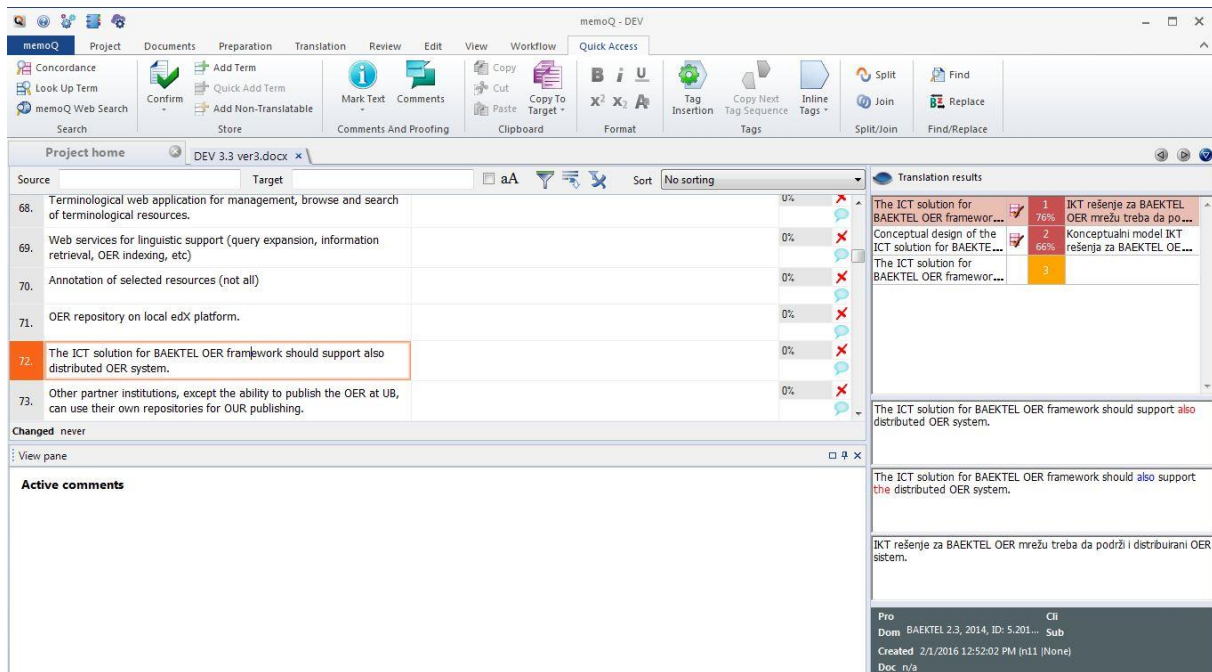


Figure 7. MemoQ interface

Slika 7. Interfejs alata MemoQ

3.3.4 SDL Trados

SDL Trados is a computer-assisted translation software suite, originally developed by the German company Trados GmbH and currently available from SDL plc, a provider of customer experience cloud solutions.

It is providing translation software solutions across the entire translation supply chain, including freelance translators, language service providers, corporate language departments and academic institutions.

Alongside SDL Trados, SDL also develops productivity tools SDL MultiTerm (for terminology management), SDL Passolo (for software localization) and SDL Language

SDL Trados je alat za računarski potpomognuto prevođenje koji je inicijalno razvila nemačke kompanija Trados GmbH, a aktuelnu verziju razvija firma SDL plc, provajder korisničkih aplikacija u oblaku.

Ona nudi prevodilačke alate različitih namena i izdanja za: samostalne prevodioce, prevodilačka društva, korporacije i akademske institucije.

Pored SDL Trados, SDL razvija i alate SDL MultiTerm (upravljanje terminologijom), SDL Passolo (lokalizacija softvera) i SDL jezik Cloud (platforma u oblaku za klasično i mašinsko prevođenje).

Cloud (cloud translation platform for human and machine translation).

More details about SDL Trados available at <http://www.translationzone.com/>

Više detalja o programu SDL Trados je dostupno na <http://www.translationzone.com>

4 Services for multilingual support

4.1 Termi

Application Termi for terminology management is available at <http://termi.rgf.bg.ac.rs/>. The verified terms are free for browsing to all users, without login, but inserts and updates are only available to authorized users. The person willing to join the community can request access to Termi and enter his/her terminology as desired. The reviewer or supervisor checks and verifies the entered terminology, whereupon it becomes publicly available. The application was developed by the team of Mining and Geology, University of Belgrade, in support of the BAEKTEL project.

The application consists of three basic web pages to manage terminology: browse, search, update, as well as pages for managing profiles and bibliography, and a page to log on to the system (login). The application has a bilingual interface (Serbian and English). The application home page has a search box for query terms in Serbian and in English (Figure 8).

Each term is described with the name, definition, synonyms, abbreviations and bibliographic source. Each term, except the top term in dictionary tree, has a hypernym term, while each term can have an arbitrary number of hyponym terms.

Aplikacija Termi za upravljanje terminologijom je dostupna na <http://termi.rgf.bg.ac.rs/>. Prelistavanje overenog sadržaja je otvoreno za sve korisnike, bez prethodnog logovanja, dok unos u bazu mogu da rade ovlašćeni korisnici. Osoba koja je voljna da se pridruži zajednici, može zatražiti pristup Termi aplikaciji i potom unositi terminologiju po želji. Redaktor, odnosno supervizor pregleda i overava unetu terminologiju, nakon čega ona postaje javno dostupna. Aplikaciju je razvio tim Rudarsko-geološkog fakulteta Univerziteta u Beogradu, za podršku BAEKTEL projektu.

Aplikacija se sastoji od tri osnovne veb stranice za rukovanje terminologijom: prelistavanje, pretraga, ažuriranje, potom stranice za upravljanje profilima i bibliografijom, kao i stranicu za prijavljivanje na sistem (logovanje). Aplikacija sadrži dvojezični interfejs (srpski i engleski). Početna strana aplikacije prikazuje polje za pretragu termina na srpskom i engleskom jeziku (Slika 8).

Svaki termin se opisuje nazivom, definicijom, sinonimima, skraćenicom i bibliografskim izvorom. Izuzev čvora (termina u stablu) koji predstavlja polazni čvor rečnika, svaki termin ima jedan nadređen termin, dok svaki termin može imati proizvoljan broj podređenih termina.

for selected term that is to be displayed (in our case this term is terminological dictionary).

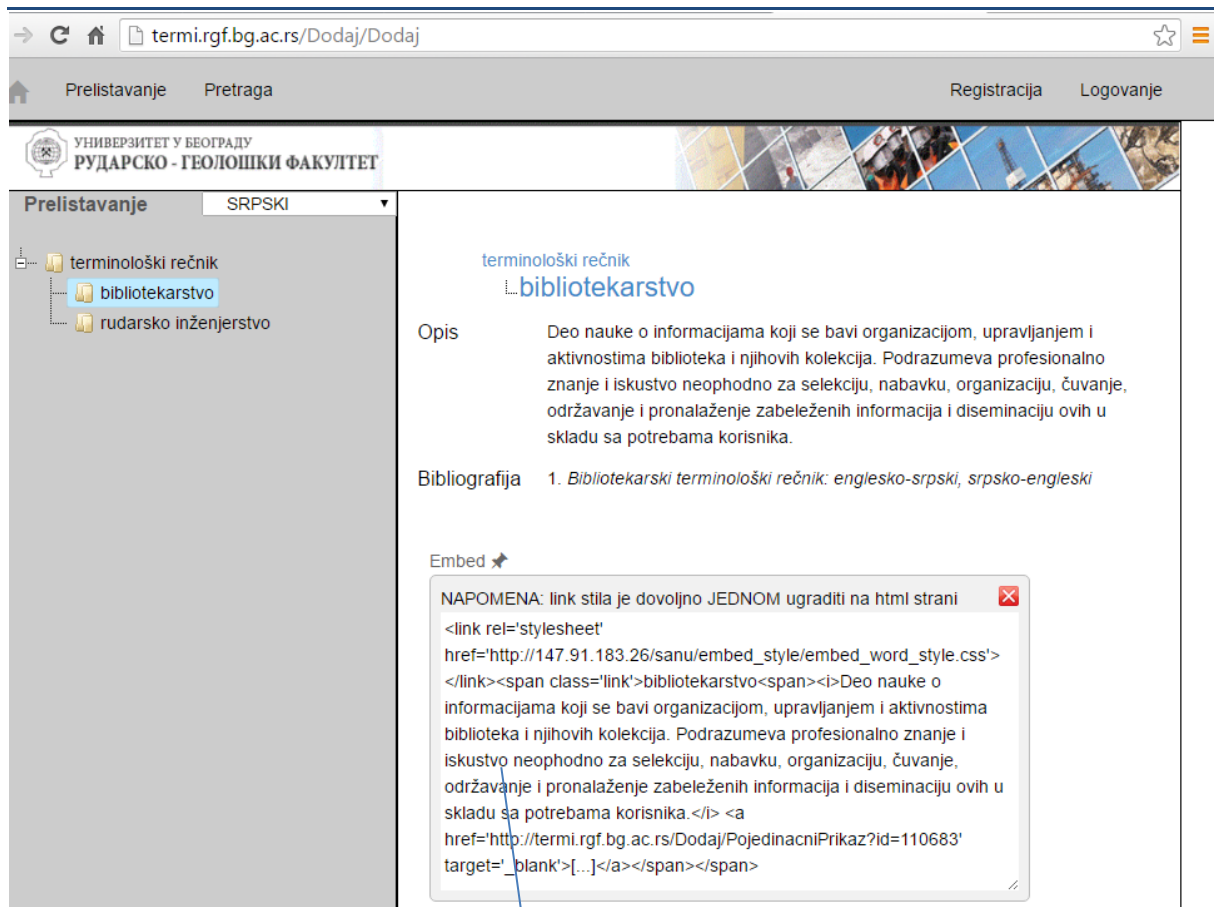
As the user does not know the specific ID code for the term, an option for generating HTML code is created. Generated code can be embedded in other web pages in order to display a tooltip with a definition and translation of term with the link to the page with full term preview.

Note: edX users who create course materials should be reminded that settings for HTML editor should be “Row”, not “Visual”.

konkretnog termina čiji prikaz tražimo (u našem slučaju taj termin je terminološki rečnik).

Kako korisnik ne zna konkretnu šifru, kreirana je opcija za generisanje HTML koda. Takav generisani kod može da se ugradi u druge veb strane i da se onda prikaže kao “oblačić” sa definicijom i prevodom, uz link ka strani sa punim prikazom svih detalja za odabrani termin.

Napomena: za korisnike edX-a koji kreiraju kurseve, napomena da treba da podešavanja za HTML editor treba da budu “Row” a ne “Visual”.



The screenshot shows a web browser window with the URL `termi.rgf.bg.ac.rs/Dodaj/Dodaj`. The page is for the University of Belgrade, Faculty of Mining and Geology. On the left, there is a navigation menu under 'Prelistavanje' with options for 'terminološki rečnik', 'bibliotekarstvo', and 'rudarsko inženjerstvo'. The main content area displays the 'terminološki rečnik' for 'bibliotekarstvo'. It includes an 'Opis' (Description) and a 'Bibliografija' (Bibliography) section. A 'Bibliografija' entry reads: '1. Bibliotekarski terminološki rečnik: englesko-srpski, srpsko-engleski'. Below the main content, there is an 'Embed' section with a code block containing HTML and CSS snippets for embedding the term. A blue arrow points from the 'Embed' section to the table below.

<i>bibliotekarstvo</i>	
Naziv:	bibliotekarstvo
Opis:	Deo nauke o informacijama koji se bavi organizacijom, upravljanjem i aktivnostima biblioteka i njihovih kolekcija. Podrazumeva profesionalno znanje i iskustvo neophodno za selekciju, nabavku, organizaciju, čuvanje, održavanje i pronalaženje zabeleženih informacija i diseminaciju ovih u skladu sa potrebama korisnika.
Naziv(Eng):	librarianship
Opis(Eng):	The profession devoted to applying theory and technology to the creation, selection, organization, management, preservation, dissemination, and utilization of collections of information in all formats. In the United States, often used synonymously with library science.
Nadređeni koncepti:	terminološki rečnik
Bibliografija:	Bibliotekarski terminološki rečnik: englesko-srpski, srpsko-engleski, Ljiljana Kovačević, Vesna Injac, Dobriša Begenišić, 2004

Figure 9 Terms browse and preview form

Slika 9. Panel za prelistavanje i prikaz termina

4.1.2 Search/ Pretraga

Figure 10 presents the web page for terms search, available either with or without login. In the search box the desired term or part of

Na panelu sa slike 10 je prikazana mogućnost pretraživanja termina dostupna sa i bez logovanja. U polju za pretragu se unosi željeni

the term are entered, but after only 3 typed characters result in the boxes below appear. The left box shows the list of terms in Serbian, which corresponds to the given key string, while the right side displays a list of terms that correspond to the given key string in the English part of the dictionary.

Further input i.e. change of specified search string is automatically reflected in the result. Comparison of specified key string and terms in the database can be done in three ways, so that the term: includes, begins with, or is identical to the entered string. One of three options is selected in the drop-down list on the right before searching.

termin ili deo termina, a već posle 3 uneta slova prikazuje se rezultat u poljima ispod. U levom polju se prikazuje spisak termina u srpskom koji odgovaraju zadatoj niski karaktera dok se desno pojavljuju termine u engleskom delu rečnika koji odgovaraju zadatoj niski.

Dalji unos tj. promena zadate niske za pretragu automatski se reflektuje na rezultat. Poređenje zadate niske karaktera i termina u bazi se može vršiti na tri načina: tako da termin sadrži, počinje ili je identičan unetom tekstu. Jedna od tri opcije poređenja se bira u padajućoj listi sa desne strane, pre startovanja pretrage.

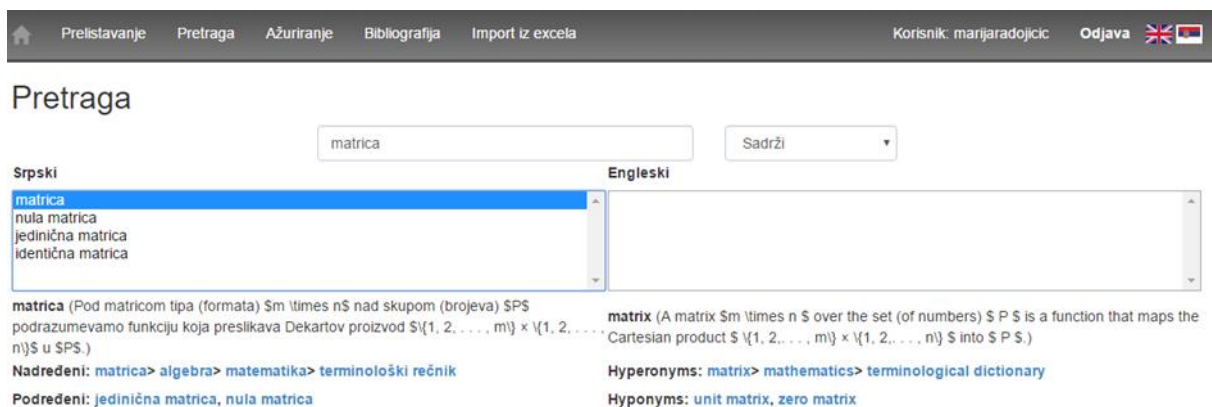


Figure 10 Search panel

Slika 10. Panel za pretragu

Click on one of the terms displayed retrieves the detailed view of the term.

Kada se klikne neki od ponuđenih termina dobije se detaljan prikaz tog termina.

4.1.3 Get account for Termi/ Dobijanje naloga za Termi

To join the dictionary writing team one should send a request for registration. Registration itself can be done in two ways: via an existing Gmail account on the page <http://termi.rgf.bg.ac.rs/Account/Login> or via the panel for registration, which is located in the main application menu.

Da bi se neko priključio timu koji se bavi pisanjem rečnika potrebno je da pošalje zahtev za registraciju. Sama registracija se može izvršiti na dva načina: preko postojećeg Gmail naloga na stranici <http://termi.rgf.bg.ac.rs/Account/Login> ili preko panela za registraciju koji se nalazi u glavnom meniju aplikacije.

If the user wants to join the team through their Gmail account, they should click on the Google button on the aforementioned page, which leads to a procedure with simple steps to enter basic info. If the user does not have a Gmail account, the panel for registration by entering basic information is available. After the completion of this process, the administrator will assign to user appropriate role and privileges in the system.

Ako korisnik želi da se timu pridruži preko svog Gmail naloga to može učiniti klikom na dugme na kome piše Google na gore pomenutoj stranici, nakon čega će biti potrebno da isprati jednostavne korake, da osnovne podatke. Ukoliko korisnik ipak nema Gmail nalog može posetiti panel za registraciju i na njemu uneti osnovne podatke o sebi. Posle ovoga nalog obrađuje administrator koji korisniku dodeljuje odgovarajuću ulogu u sistemu.

4.1.4 Writing/ Ažuriranje

Page for entering terms (Figure 11) looks like the page for browse, but as for its functionality, it is used for term record update (add, modify, and delete). In the right part of the panel there is a toolbar with buttons at the top, and below them there are text fields. This page is available only for logged in users.

Stranica za dodavanje i ažuriranje termina (slika 11) izgledom podseća na stranicu *Prelistavanje*, a što se tiče funkcionalnosti, ona služi za ažuriranje zapisa termina (dodavanje, izmena i brisanje). U desnom radnom delu panela se nalaze dugmići menija na vrhu, a ispod njih se nalaze tekstualna polja. Ova stranica je namenjena samo ulogovanim korisnicima.

Vocabulary articles about terms are followed by a status, which encodes the degree of completion and verification of term defining.

Rečničke članke o terminima prati status koji govori o stepenu završenosti i verifikovanosti definisanja termina.

There are four statuses:

Moguća su četiri statusa:

B - deleted (not visible via application, but still remains in database),

B – obrisan (članak nije vidljiv putem aplikacije ali i dalje postoji u bazi),

U - updatable (editor began writing vocabulary article but it is not completed),

U – u radu (editor je počeo pisanje rečničkog članka ali nije završio),

Z - locked (editor has finished writing and handed over to reviewer for review),

Z – zaključan (editor je završio pisanje i predao redaktoru na proveru),

V - verified (reviewer finished review, verified the vocabulary article and it becomes publicly available and visible).

V – verifikovan (redaktor završio proveru, overio rečnički članak i on postaje javno dostupan i vidljiv).

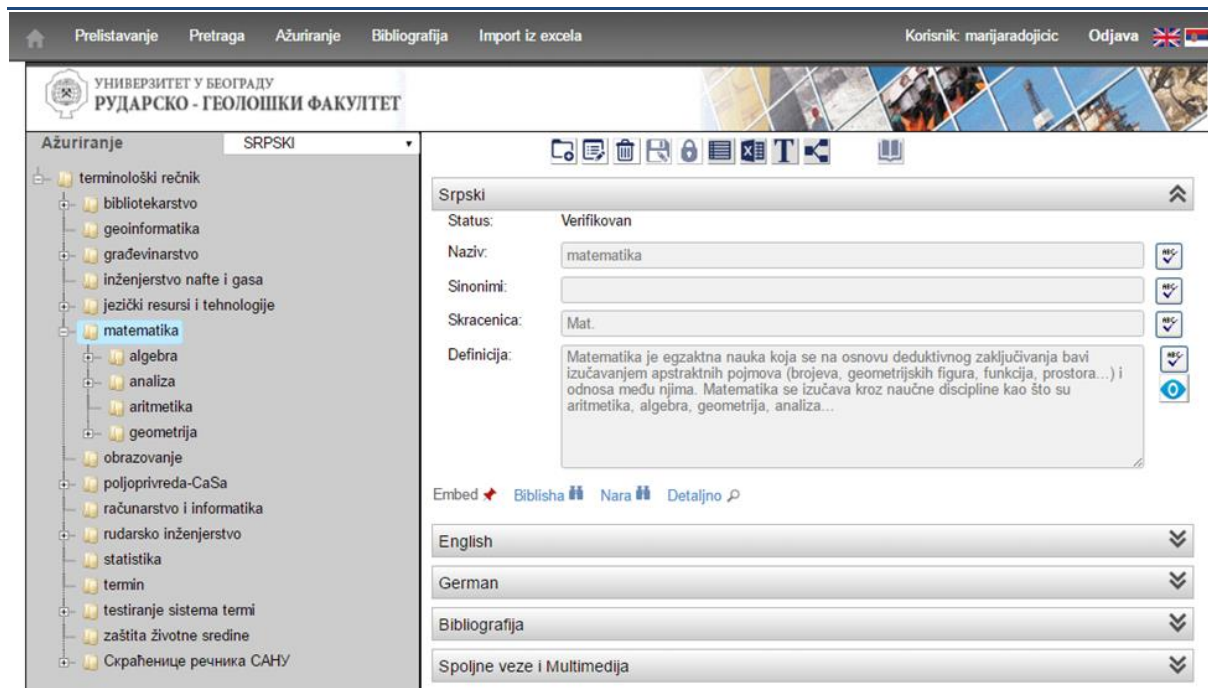








Figure 11 Page for terms writing

Slika 11 Ažuriranje termina

Modify: By clicking on the term in a tree the text fields are filled in with details of the term (definition, synonyms, abbreviation bibliographic sources, multimedia content (Figure 12)), which are by default read only. If the user has the right to change the chosen record, it is possible to click on the button  and enter into mode for editing. After completion of editing, a user can save the changes in two ways:

- The first method is to click the button  record to execute save to database (recommended in case the user wants to continue editing before sending to reviewer)
- The other is to click on the end-changes button  to save the final version of the term.

Izmena: Klikom na neki termin u stablu popunjavaju se tekst polja podacima o tom terminu (definicija, sinonimi, skraćenica, bibliografski izvori, multimedijalni sadržaji (Slika 12)), koja su podrazumevano needitabilni. Ukoliko korisnik ima pravo da menja izabrani termin, omogućeno je da klikom na taster **Izmena-termina**  pređe u mod za izmenu. Po završetku izmene korisnik može da sačuva unete podatke na dva načina:

- prvi način je da klikom na dugme  snimi izvrši čuvanje termina u bazi (preporučuje se u slučaju da korisnik želi još nešto da promeni na terminu pre nego što da pošalje redaktoru),
- druga je da klikom na dugme  kraj-izmena izvrši čuvanje svoje finalne verzije termina.

Spoljne veze i Multimedia

Tip: Spoljne veze (URL, odrednice u nekom ć ▾)

Naziv: Wikipedia

Autor(i): Marija Radojčić



Godina: 2016

Opis: Link ka stranici wikipedije |

URL: https://sh.wikipedia.org/wiki/Glavna_stranica

Sačuvaj

Fotografije

Linkovi

1. <http://baektel.eu> (BAEKTEL: Blending academic and entrepreneurial knowledge in technology enhanced learning)

PDF dokumenta


1. Uputstvo za upotrebu (Uputstvo za upotrebu Termini aplikacije.) [pdf]


Zvuk


▶ 0:01 / 0:01


Figure 12 Adding multimedia content


Slika 12. Dodavanje multimedijalnog sadržaja


Adding new: Each term (except root) has hypernym term, and consequently adding a new terms begins with selection of the hypernym by clicking in the tree. When a hypernym is chosen next step is to click on the button add-term , followed by filling text boxes with definition and other information related to term, as previously explained. The term is added first in Serbian language, saved and then its translation into other languages is added.

Dodavanje novog: Svaki termin (osim korena) ima svoj nadređeni termin, pa se shodno tome dodavanje novog termina započinje biranjem nadređenog, klikom u stablu. Sledeći korak je klik na dugme dodavanje-termina  nakon čega se tekstualna polja popunjavaju podacima o terminu i izvrši snimanje kako je na prethodno objašnjeno. Termin se unosi na prvo na srpskom jeziku, sačuva, a potom se prilikom ažuriranja dodaju njegovi prevodi na engleski i druge jezike.

Delete: Akin to modify, in order to delete a record, the user must have the right to do so. If user has the right for delete, he must first turn into the mod for modify (see modify), and then to click the delete term button . The process of term deleting can be labeled as logical deletion, which means that the term remains in the database, but since then has status B (being deleted).

Bibliography: When we are in the mode for modifying a term, that term can have a bibliography attached (Figure 13). Clicking the button add-bibliography  opens a pop-up with the form for adding new bibliography, and while typing in the title the application will offer an autocomplete to choose from existing bibliography. Bibliography is deleted (removed from a term), by removing it from the bibliography table, which is located below the forms for data entry and modification of terms. If a bibliography does not exist in the database, it must first be added by means of the Bibliography panel, which shall be discussed in the next section.

Brisanje: Analogno, kao kod izmena da bi se obrisao pojam korisnik mora da ima pravo na to. Ako korisnik ima pravo da briše termin, prvo treba da selektuje taj termin, a zatim da klikne na dugme brisanje termina . Sam proces brisanja termina može se nazvati logičkim brisanjem, to znači da termin ipak ostaje u bazi, ali od tada ima status B (u brisanju).

Bibliografija: Kada se nađemo u modu za izmenu termina, tom terminu možemo pridružiti bibliografiju (Slika 13). Klikom na dugme dodaj-bibliografiju  otvara se prozor sa formom za dodavanje nove bibliografije, i dok budemo kucali naslov aplikacija će nam iz baze nuditi postojeće bibliografske reference od kojih treba da izaberemo neku postojeću bibliografiju. Bibliografije brišemo (uklanjamo sa termina) tako što ih uklonimo iz tabele bibliografija koja se nalazi ispod forme za unos i izmenu termina. Ako neka bibliografija ne postoji u bazi, potrebno je prvo dodati bibliografiju na panelu Bibliografija, o tome će biti reči u narednom odeljku.

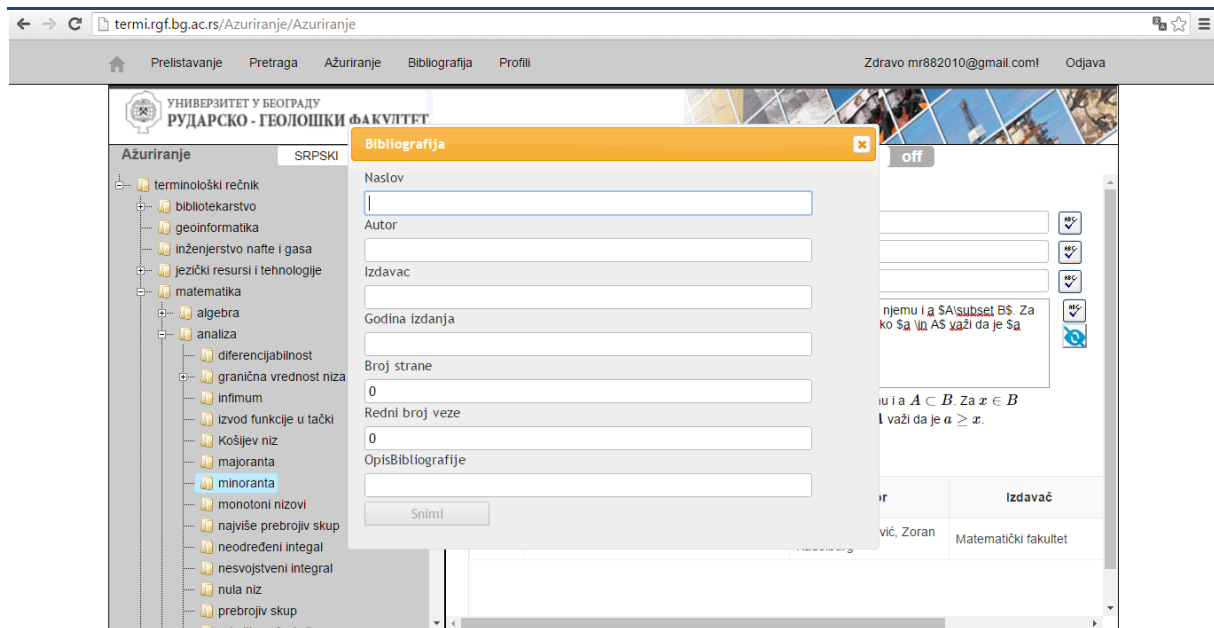






Figure 13 Assignment of bibliography to a term Slika 13 Pridruživanje bibliografije terminu



Help for update: In order to facilitate the entry and update of the terms two types of assistance are provided:


- Spell check is added in order to facilitate text entry. Next to each text field there is a button  for checking whether the entered words are properly spelled. Clicking on this button opens a new window to perform the aforementioned check. Of course, it is important to note that spelling is controlled only on the basis of the words that exist in the dictionary (spell check dictionary). If a word is underlined with a red line that means that it is not well spelled.
- When giving the term definition, especially mathematical and technical, some parts of the description cannot be presented only as simple text. This is why latex fragments are introduced and they are infiltrated in the description of the term itself. Each latex fragment must be entered between

Pomoć pri ažuriranju: Kako bi se olakšao unos i ažuriranje termina predviđene su 2 vrste pomoći:


- Spell check je dodat da bi se korisnicima olakšao unos teksta. Pored svakog tekstualnog polja nalazi se dugme  za proveru da li su unete reči dobro spelovane. Klikom na ovo dugme otvara se novi prozor u kome se vrši pomenuta provera. Naravno bitno je napomenuti se samo spelovanje kontroliše na osnovu reči koje postoje u rečniku (spell check rečniku). Ako je neka reč podvučena crvenom linijom to znači da nije dobro spelovana.
- Pri unosu definicije termina, naročito matematičkih i tehničkih, pojedine delove opisa ne možemo predstaviti samo uz pomoć teksta. Zbog ovoga pribeglo se tome da se delovi latex fragmenata infiltriraju u sam opis termina. Svaki latex fragment mora biti upisan između 2 znaka



two \$ characters. While entering, the user needs to have an insight into how the description looks like if fragments of between two \$ characters are interpreted as latex, which can be done by clicking the button . To hide of preview, the user should click the button .



Export: On the page for terminology update there is a possibility for exporting a selected part of the dictionary to Excel or tbx file. Exporting is done in the following way: the user first selects the term and the tree for export, and then by clicking the Excel button  the selected term and its hyponyms will be exported to an Excel file. Export to tbx file is the same, only the button TBX export  needs to be clicked on.


In addition to these two types of export, there is also a form for tabular data display. When the user selects a term in tree and clicks the button for Tabular data , data in tabular form appears on the screen instead of the hierarchical display.

In addition to everything mentioned so far, the page for update provides yet another option. If the user wants to focus only on a certain part of the vocabulary, it can be done in the following way: it is necessary that the user chooses the term they want to focus on.


For example, if the user is interested only in mathematics, he/she will select the term: mathematics, then click on the Subtree button . Thereafter, a new page will open with the hierarchical view with selected term and terms that are below it in the hierarchy (for example, mathematics and everything under that term).

\$. Pri unosu potrebno je da korisnik ima uvid u to kako bi njegov opis izgledao kada bi se fragmenti između 2 znaka \$ interpretirali kao latex, a to se može učiniti klikom na dugme . Kada želimo da ugasimo ovakav prikaz potrebno da kliknuti na dugme .

Izvoz: Na stranici ažuriranja postoji mogućnost da se ceo rečnik ili neki njegov deo eksportuje u Excel ili tbx file. Eksportovanje se izvodi na sledeći način: prvo se selektuje termin u stablo koji želimo da eksportujemo, klikom na dugme Excel izvoz  izvršiće se izvoz selektovanog termina i termina koji su mu hijerarhijski podređeni u jednu Excel datoteku. Što se tiče tbx datoteka sve je isto, jedino potrebno kliknuti na dugme TBX izvoz .

Pored ova dva vida izvoza postoji i još jedan način za prikazivanje podataka, a to je tabelarni prikaz. Naime korisnik izabere termin u stablu i klikne na dugme Tabelarni prikaz  na ekranu se umesto hijerarhijskog prikaza pojavljuje tabelarni prikaz podataka.





Pored svega do sada pomenutog na stranici ažuriranja korisniku se pruža još jedna mogućnost. Ako korisnik želi da se fokusira samo na određeni deo rečnika to može učiniti na sledeći način: potrebno je da korisnik izabere termin na koji želi da se fokusira.

Na primer ako korisnika isključivo zanima matematika, on će selektovati termin matematika, zatim je potrebno da se klikne na dugme Podstablo . Kao proizvod ovog postupka dobiće se nova stranica koja će sa leve strane u hijerarhijskom prikazu imati samo selektovani termin i one termine koji se





nalaze u hijerarhiji ispod njega (na primer matematika i sve što se nalazi ispod nje).

4.1.5 Bibliographic sources/ Bibliografski izvori

Page for bibliographic sources management (Figure 14) is designed for full management of bibliography data. For this panel use a table of existing bibliographical sources will be displayed. On this page we can add a new bibliography, modify existing, delete bibliography (in case that it is not connected to a term), and finally see which terms are "connected" to the selected bibliography.

Adding a new bibliography is done by clicking the Add bibliography button . After this it is necessary to fill in the form and save the data related to the newly added bibliography. Bibliography editing is started by clicking on the button  located in the table. Deleting of a bibliography record is performed by clicking on the button  next to the desired bibliography, while the overview of the terms that are related to a bibliography can be obtained by clicking on the button .

Panel Bibliografija (Slika 14) predviđen je za upravljanje bibliografijama. Pri samoj poseti ovog panela biće prikazana tabela postojećih bibliografskih izvora. Na ovoj stranici možemo da dodamo novu bibliografiju, izmenimo postojeću, izbrišemo bibliografiju (pod uslovom da nije vezana ni za jedan termin) i na kraju da pogledamo koji se to termini "pozivaju" na neku bibliografiju.

Dodavanje nove bibliografije se vrši klikom na dugme Dodaj bibliografiju . Posle ovoga potrebno je popuniti formu i sačuvati podatke vezane za novododatu bibliografiju. Izmena bibliografije vrši se klikom na dugme  koje se nalazi u samoj tabeli. Brisanje bibliografije se vrši klikom na dugme  pored željene bibliografije, dok se pregled termina koji su vezani za neku bibliografiju može dobiti klikom na dugme .

l.ac.rs/Bibliografija/Bibliografija?grid-page=2

[Prestavanje](#)
[Pretraga](#)
[Ažuriranje](#)
[Bibliografija](#)
[Profil](#)
Zdravo mr882010@gmail.com! [Odjava](#)

Bibliografija

Naslov	Skraćenica	Autor	Godina izdanja	Izdavac	Apstrakt	Br strana
Online Dictionary for Library and Information Science	ODLIS	Joan M. Reitz	2004-2014	ABC-CLIO		
Visokoškolske biblioteke		Aleksandra Vraneš	2004	Univerzitetna biblioteka "Svetozar Marković"		
Matematička analiza	analiza	Dušan Adnađević, Zoran Kadelburg	2004	Matematički fakultet		
Matematika 1	Mat 1	Ivan Obradović		Rudarsko-geološki fakultet		
Euklidska i hiperbolička geometrija	Geometrija	Zoran Lučić	1997	Total design i Matematički fakultet		
Elementi više matematike II		Nada Miličić, Miloš Miličić	2005	Akademski misao		

< 1 2

Figure 14. Bibliography data management

Slika 14 Panel Bibliografija

4.1.6 User profiles/ Korisnički profili

Page for user profiles management is shown in Figure 15. Access to this web page is available only for administrators.

This page shows the list of users with basic information about users and their privileges.

Stranica namenjena rukovanju korisničkim profilima je prikazana na Slici 15. Pristup ovoj stranici imaju jedino administratori.

Stranica prikazuje listu korisnika sa osnovnim podacima o korisnicima i njihovim ovlašćenjima.



ac.rs/ProfilPanel/ProfilPanel


Korisničko ime	Ime	Prezime	Pozicija	Institucija	Grad	Država	Uloga u sistemu	Status naloga	Domen
aleksandra.trtovac@gmail.com	Aleksandra			Univerzitetska biblioteka "Svetozar Marković"	Beograd	Srbija	Editor		110683
aleksandra@gmail.com	aleksandra			Univerzitetska biblioteka	Beograd	Srbija	Editor	Aktivan	110683
bojan.zlatic@rgf.bg.ac.rs									
bojan.zlatic1@tt.com	Bojan	Zlatic	Levo smetalo	RGF	Cacak	Beograd	Admin	Aktivan	1
bojan.zlatic1@tt.com	Bojan	Zlatic	Levo smetalo	RGF	Cacak	Beograd	Redaktor	Aktivan	1
bojan.zlatic1@tt.com	Bojan	Zlatic	Levo smetalo	RGF	Cacak	Beograd	Redaktor	Aktivan	101
CvetanaJK@gmail.com	Cvetana	Krstev		Filološki fakultet	Beograd	Srbija	Redaktor	Aktivan	1
dalibor.vorkapic@rgf.bg.ac.rs	Dalibor	Vorkapic		Rgf	Beograd	Srbija	Editor	Aktivan	1
daliborvorkapic@gmail.com	Dalibor			RGF	BG	SR	Editor	Aktivan	1
IvanObradovic@gmail.com	Ivan	Obradovic		Univerzitet u Beogradu, Rudarsko-geološki fakultet	Beograd	Srbija	Redaktor	Aktivan	1



© 2016 - RGF


Figure 15. List of users

Slika 15 Lista korisnika


Administrator can change data about users related to privileges (Figure 16) by clicking the edit button . Clicking on this button enables the change of settings (specific roles in the system are: Admin-101, 102-Reader, Editor-103, 104-Reviewer). The admin can add a new role to the existing user, by click on the image  and filling in the role in the system and the ID of term from witch (in hierarchy) the privilege (permission) is assigned.

Admin also has the option to add a new user on main page by clicking on the image , which opens the form for adding a new user (required fields are user name (e-mail) and password of minimum 6 characters)!

Administrator ima mogućnost da izmeni podatke o korisnicima vezanim za privilegije (Slika 16) klikom na dugme za edit . Klikom na ovo dugme se omogućava podešavanje profila (karakteristične uloge u sistemu su: 101-Admin, 102-Čitač, 103-Editor, 104-Redaktor). Ako admin želi da korisniku doda novu ulogu pored već postojeće potrebno je da klikne na sliku  i da popuni polja Uloga u sistemu i šifru termina od kog kreće (u hijerarhiji) ovlašćenje.

Admin takođe ima mogućnost da doda novog korisnika tako što će na osnovnoj strani da klikne na sliku  time se otvara forma za dodavanje novog korisnika (obavezno dodati

korisničko ime (e-mail) i lozinku od minimum 6 karaktera)!

Podešavanje naloga 

Korisničko ime (e-mail)	<input type="text" value="bojan.zlatic@rgf.rs"/>
Ime	<input type="text"/>
Prezime	<input type="text"/>
Pozicija	<input type="text"/>
Institucija	<input type="text"/>
Grad	<input type="text"/>
Država	<input type="text"/>
Jezik	<input type="text"/>
Datum otvaranja naloga	<input type="text" value="6/19/2015 12:00:00 AM"/>
Datum isticanja naloga	<input type="text" value="8/31/2020 12:00:00 AM"/>
Status naloga	<input type="text" value="Aktivan"/>

Figure 16. User account page

Slika 16 Forma korisničkog naloga

4.1.7 Tips and tricks

Instructions for writing definitions

- Definition: Definition must contain only important features of the concept being defined, i.e. only those characteristics that separate the concept from related concepts (etymology, origin, inventor are not listed, or they are listed in parentheses after the definition).
- The definition should not be negative. (E.g. Equilateral triangle is not isosceles nor angles asymmetrical triangle, but rather: an equilateral triangle is a triangle with all sides of equal length.)
- There may be exceptions when the profession requires (E.g. Mathematical definitions)

Uputstava za pisanje definicija

- Definicija: Definicija mora sadržati samo bitna obeležja pojma koji se definiše, tj. samo ona obeležja koja dati pojam odvajaju od srodnih pojmova (etimologija, poreklo, pronalazač i sl. se ne navode ili se navode u zagradi nakon definicije).
- Definicija ne bi trebala da bude negativna. (Npr. Jednakostranični trougao nije ni jednakokraki ni raznostranični već: Jednakostranični trougao je trougao kome su sve stranice jednake dužine.)
- Mogu postojati izuzeci kada to struka zahteva. (Npr. matematičke definicije)

- Sometimes the term does not have to be described in too many details (e.g. instead of a derivative of the function f at point M it should be derivative of the function in point, because the derivative in the point M and S is being calculated in the same way, as well as in any other point).
- All names which are used in a definition must be those that the authors considered the best for a term.
- In the same definition or in related terms, different names must not appear for the same concept.
- In the definition, the same name should be used as for the term itself (if the term is e.g. notebook, laptop should not be used in the definition, but rather recorded as a synonym).
- If one term has several definitions, one is selected. In general lexica dictionaries, when several meanings are recorded, the definitions are numbered with ordinal numbers or they are separated with a comma. This is an open question in terminology, while terminologists propose that there is one preferred term and one preferred definition.
- Errors: You should not start a definition with: The term denotes ..., definitions should never start with etymology and acronyms.
- Definitions from the same domain must be aligned with each other (in a similar way we define related concepts).
- It is not good to use demonstrative pronouns.
- Nekada termin ne treba suviše detaljno opisivati (npr: Umesto izvod funkcije f u tački M treba pisati Izvod funkcije u tački, jer se na isti način računa izvod f u M i u S tački ili bilo kojoj drugoj).
- Svi nazivi koji se koriste u nekoj definiciji moraju biti oni koje autori smatraju najboljim nazivom za određeni pojam.
- U istoj definiciji ili povezanim pojmovima ne smeju se pojaviti različiti nazivi za isti pojam (računar, kompjuter-odabрати računар; prelistač, brauzer odabрати jedan termin)
- U definiciji treba koristiti naziv koji se koristi i u samom terminu (ako je termin recimo beli luk u definiciji ne treba koristiti termin češnjak, već termin beli luk a češnjak navesti kao sinonim).
- Ako jedan termin ima više definicija, bira se jedna. U rečnicima opšteg jezika kada se navodi više značenja tada se definicije navode rednim brojevima ili se odvajaju zarezom. Ovo je predmet rasprava, terminolozi generalno predlažu da treba da postoje jedan preferirani termin za pojam i jedna opšteprihvaćena definicija.
- Greške: Ne treba početi definiciju sa “:” “Termin označava...”, definicije se ne počinju etimologijom i akronimima.
- Definicije iz istog područja moraju biti međusobno usklađene (na sličan način definišemo srodne pojmove).
- Nije dobro koristiti pokazne zamenice. (Čačanska rodna je *ona* šljiva...već: Čačanska rodna je šljiva...)

- Provide a definition, not an explanation (Do not write explanations, e.g. The most common variety of plums grown in Dragačevo region is...)
- Initial letters of the terms are lowercase letters, while the definitions start with a capital letter.
- Do not use circular definitions: The mobile phone is a phone that is mobile.
- Keep in mind that this is a dictionary not a lexicon and that the definitions in a dictionary are shorter than descriptions in a lexicon.
- Analyze related online terminological dictionaries in other languages
- Keywords from ResourceSpace metaportal should always be described in Termi database.
- Keywords are generally listed from specific to general.
- Dati definiciju, a ne objašnjenje (Ne pisati objašnjenja npr. Najčešća sorta šljiva koja se uzgaja u dragačevskom kraju je ...)
- Početna slova naziva termina su mala slova dok se nazivi definicija pišu velikim početnim slovom.
- Ne navoditi kružne definicije: Mobilni telefon je telefon koji je mobilan.
- Imati na umu da je ovo rečnik a ne leksikon i da su definicije u rečniku kraće od opisa u leksikonu.
- Analizirati srodne on-line terminološke rečnike na drugim jezicima.
- Ključne reči navedene u Resource Space metaportalu treba obavezno da budu opisana u Termi bazi.
- Ključne reči se generalno navede od specifičnih ka opštim.

Spelling specifies that for the use of foreign words the phonetic transliteration rule is applied (for example, instead of writing "web" in Serbian one should write "veb").

Pravopis određuje da se za tuđice koristi pravilo fonetske transliteracije (na primer pisati "veb" umesto "web").

5 OER tagging

5.1 Elements for BMP: Description and cataloguing guidelines

On the web page of Termi application with tutorials for the basic use of the application: <http://termi.rgf.bg.ac.rs/Dodaj/Help>, there is a video in Serbian entitled "TERMI - Ugradnja termina na edX portal", about integration of term definitions into OER resources

Figure 17 presents a code prepared for insertion into a HTML page or an OER page segment.

Na adresi sa video materijalima za osnovno korišćenje aplikacije Termi (<http://termi.rgf.bg.ac.rs/Dodaj/Help>) je i video sa uputstvom za integrisanje zapisa iz terminološke baze u lekcije. Video se zove: "TERMI - Ugradnja termina na edX portal".

Slika 17 prikazuje kod spreman za ubacivanje u HTML stranicu ili deo stranice obrazovnog resursa.

Embed ↗

```
NAPOMENA: link stila je dovoljno JEDNOM ugraditi na html strani
<link rel='stylesheet'
href='http://147.91.183.26/sanu/embed_style/embed_word_style.css'>
</link><span class='link'>bibliotekarstvo<span><i>Deo nauke o
informacijama koji se bavi organizacijom, upravljanjem i aktivnostima
biblioteka i njihovih kolekcija. Podrazumeva profesionalno znanje i
iskustvo neophodno za selekciju, nabavku, organizaciju, čuvanje,
održavanje i pronalaženje zabeleženih informacija i diseminaciju ovih u
skladu sa potrebama korisnika.</i> <a
href='http://termi.rgf.bg.ac.rs/Dodaj/PojedinacniPrikaz?id=110683'
target='_blank'>[...]</a></span></span>
```

<i>bibliotekarstvo</i>	
Naziv:	bibliotekarstvo
Opis:	Deo nauke o informacijama koji se bavi organizacijom, upravljanjem i aktivnostima biblioteka i njihovih kolekcija. Podrazumeva profesionalno znanje i iskustvo neophodno za selekciju, nabavku, organizaciju, čuvanje, održavanje i pronalaženje zabeleženih informacija i diseminaciju ovih u skladu sa potrebama korisnika.
Naziv(Eng):	librarianship
Opis(Eng):	The profession devoted to applying theory and technology to the creation, selection, organization, management, preservation, dissemination, and utilization of collections of information in all formats. In the United States, often used synonymously with library science.
Nadređeni koncepti:	terminološki rečnik
Bibliografija:	Bibliotekarski terminološki rečnik: englesko-srpski, srpsko-engleski, Ljiljana Kovačević, Vesna Injac, Dobrila Begenišić, 2004

Figure 17. HTML code for embedding in OER

Slika 17. HTML kod za ugradnju u OER

Figure 18 presents OER with embedded tooltip with term definition segment and link to the term web page that appears during mouse over event.

Slika 18 prikazuje HTML stranicu obrazovnog resursa sa ugrađenim oblakom sa definicijom koja se pojavljuje prilikom prevlačenja miša.

Course Updates & News

📅 **DECEMBER 29, 2015**

Dobrodošli na kurs Leksicko prepoznavanje o obradi prirodnih jezika. Kursom su obuhvaćene morfološka, leksička i sintaksička analiza u obradi prirodnih jezika. Obrada prirodnog jezika razvija metode, alate i tehnike dobijene na osnovu nivoa plitkog parsiranja. Neke od pokrivenih tema su upotreba automata u tekstu, konačni automati, transdudtori, elektronski rečnici i potrebnosti veštine i znanja za korišćenje predstavljenih jezika. Uči potrebne informacije od polaznika će možda poželeti da učestvuju u razvoju postojećih metoda i alata.

Obrada prirodnog jezika razvija metode, alate i tehnike dobijene na osnovu teorijskih rezultata računarske lingvistike, informatike, matematike i drugih kako bi pomogla u obradi, razumevanju i proizvodnji izraza ljudskog jezika, kao i u omogućavanju brojnih formi interakcije čovek-mašina. [...]

Figure 18. Embedded term definition in OER

Slika 18. Ugrađena definicija termina u OER

Automatic procedure is used to speed up manual annotation. System for automatic annotation is based on Unitex system for text processing (Krstev, Processing of Serbian – Automata, Texts and Electronic dictionaries, 2008)

Automatski postupak se koristi za ubrzanje ručne anotacije. Sistem za automatsko anotiranje se zasniva na Unitex sistemu za obradu teksta (Krstev, Processing of Serbian – Automata, Texts and Electronic dictionaries, 2008).

The first task was to prepare a specific electronic dictionary for annotation that covers terms from Termini database. Part of its structure is presented in Fig.19. First, all inflectional forms (for example, all grammatical cases) were produced for each term in the Termini database, and then their definitions from the Termini database were associated.

Prvi zadatak je bio da se pripremi specifičan elektronski rečnik za anotacije koji pokriva termine iz Termini baze. Deo strukture rečnika je predstavljen na slici 19. Za svaki termin iz baze Termini su prvo proizvedeni svi flektivni oblici (na primer svi padeži), a potom su im pridružene definicije iz baze Termini.

Procedure of production of this kind of electronic dictionaries is also automated and it is embedded in LeXimir tool used for lexical resources management.

Postupak proizvodnje ove vrste elektronskih rečnika je takođe automatizovan i ugrađen u LeXimir alat za upravljanje jezičkim resursima.

matematici,matematika.N+ *Matematika je egzaktna nauka koja se na osnovu deduktivnog zaključivanja bavi izučavanjem apstraktnih pojmova (brojeva, geometrijskih figura, funkcija, prostora...) i odnosa među njima. Matematika se izučava kroz naučne discipline kao što su aritmetika, algebra, geometrija, analiza...*

matematika,matematika.N+ *Matematika je egzaktna nauka koja se na osnovu deduktivnog zaključivanja bavi izučavanjem apstraktnih pojmova (brojeva, geometrijskih figura, funkcija, prostora...) i odnosa među njima. Matematika se izučava kroz naučne discipline kao što su aritmetika, algebra, geometrija, analiza.....*

matematikama,matematika.N+ *Matematika je egzaktna nauka koja se na osnovu deduktivnog zaključivanja bavi izučavanjem apstraktnih pojmova (brojeva, geometrijskih figura, funkcija, prostora...) i odnosa među njima. Matematika se izučava kroz naučne discipline kao što su aritmetika, algebra, geometrija, analiza.....*

matematike,matematika.N+ *Matematika je egzaktna nauka koja se na osnovu deduktivnog zaključivanja bavi izučavanjem apstraktnih pojmova (brojeva, geometrijskih figura, funkcija, prostora...) i odnosa među njima. Matematika se izučava kroz naučne discipline kao što su aritmetika, algebra, geometrija, analiza....*

Figure 19. Examples of dictionary entries for automatic annotation

Slika 19. Primeri zapisa iz rečnika za automatsku anotaciju

5.2 Annotation graph

To apply dictionary and perform annotation, a special graph was produced, as shown in figure 20. When the system finds the word form (e.g. matematike), it associates it with the term lemma, that is the form stored in dictionary (e.g. matematika), usually nominative case singular. Term appearance is amended with SPAN snippet, which includes term definition and URL, similar as in manual annotation.

Da biste se primenio rečnik i izvršile anotacije, odnosno obeležavanje teksta, konstruisan je poseban graf, prikazan na slici 20. Kada sistem nađe oblik određene reči (npr. matematike), pridruži se lema termina, odnosno oblik u kom se termin nalazi u rečniku (npr. matematika) obično nominativ jednine. Pojavljivanje termina u tekstu se upotpunjuje SPAN etiketom koja sadrži definiciju i URL termina, slično kao u ručnom režimu anotiranja.

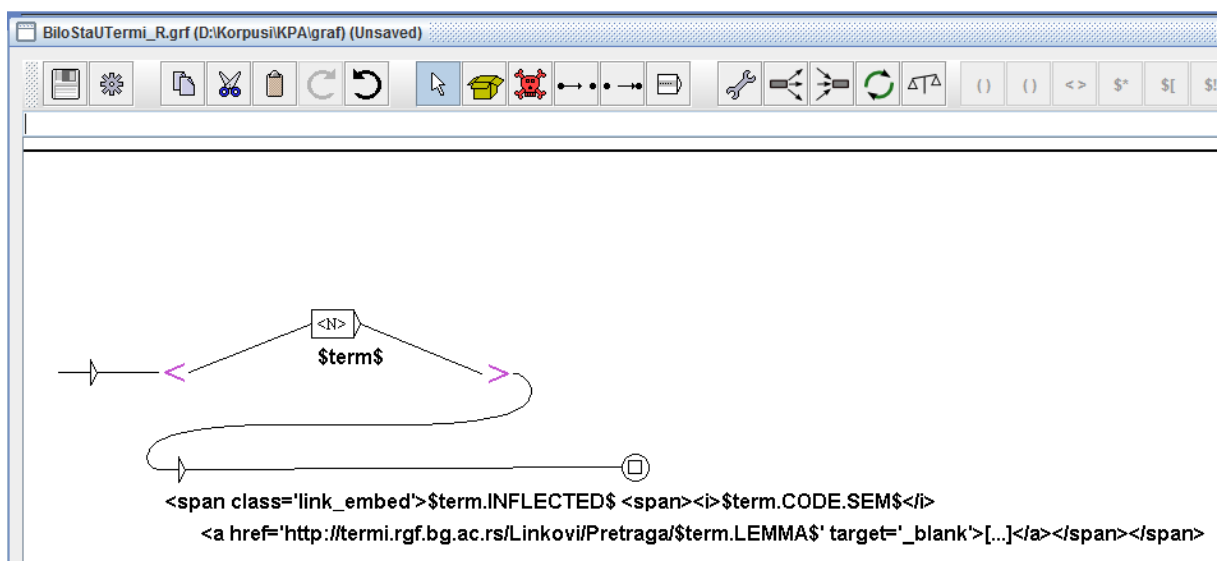


Figure 20. Annotation graph

Slika 20. Graf za anotaciju

5.3 Annotated text

Figure 21 presents an excerpt of text example with xml annotations.

Slika 21 prikazuje isečak teksta sa ubačenim xml anotacijama.

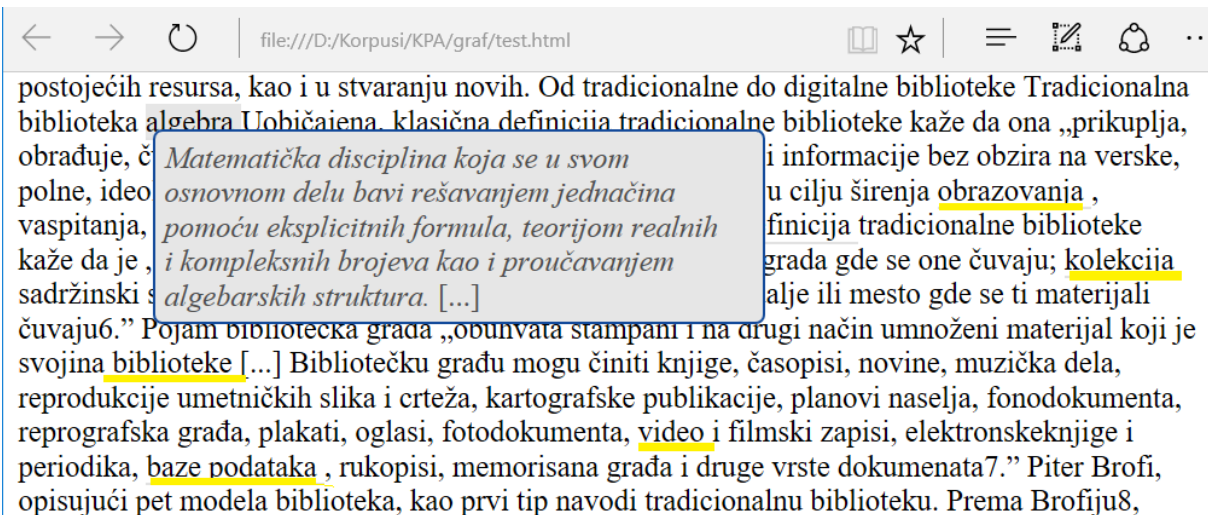
```
<link rel='stylesheet' href='http://termi.rgf.bg.ac.rs/embed_style/embed_word_style.css'></link>
Dobrodošli na kurs <span class='link_embed'>leksičko prepoznavanje </span><i>U prirodnim jezicima, značenje
kompeksne izgovorene rečenice može biti shvaćeno kao raščlanjavanje u manje leksičke segmente (reči jezika),
udruženo sa značenjem svakog segmenta, kombinujući značenja sa gramatičkim pravilima jezika.</i> <a href=
'http://termi.rgf.bg.ac.rs/Linkovi/Pretraga/leksičko_prepoznavanje' target='_blank'>[...]</a></span></span> o
obradi prirodnih jezika. {S} Kursom su obuhvaćene morfološka, leksička <span class='link_embed'>i <span><i>Broj
započetih epizoda obolevanja ili novoobolelih jedinki tokom datog perioda u određenoj populaciji</i> <a href=
'http://termi.rgf.bg.ac.rs/Linkovi/Pretraga/I' target='_blank'>[...]</a></span></span> sintaksička <span class=
'link_embed'>analiza <span><i>Proces identifikacije pitanja ili teme koje treba obraditi, modeliranje
problema, ispitivanje rezultata modela, interpretiranje rezultata i moguće davanje preporuka.</i> <a href=
'http://termi.rgf.bg.ac.rs/Linkovi/Pretraga/analiza' target='_blank'>[...]</a></span></span> u obradi
prirodnih jezika. {S} Sintaksička <span class='link_embed'>analiza <span><i>Proces identifikacije pitanja ili
teme koje treba obraditi, modeliranje problema, ispitivanje rezultata modela, interpretiranje rezultata i
moguće davanje preporuka.</i> <a href='http://termi.rgf.bg.ac.rs/Linkovi/Pretraga/analiza' target='_blank'>
[...]</a></span></span> je obuhvaćena do nivoa plitkog parsiranja. {S} Neke od pokrivenih tema su upotreba
regularnih izraza za prepoznavanje obrazaca u tekstu, konačni automati, transduktori, elektronski rečnici,
kaskade <span class='link_embed'>i <span><i>Broj započetih epizoda obolevanja ili novoobolelih jedinki tokom
datog perioda u određenoj populaciji</i> <a href='http://termi.rgf.bg.ac.rs/Linkovi/Pretraga/I' target='_blank'
>[...]</a></span></span> višesložne reči. {S} Polaznici će steći potrebne veštine <span class='link_embed'>i
```

Figure 21. Annotated text

Slika 21. Anotiran tekst

The same text, but previewed in a web browser is presented in figure 22, showing a cloud with definition and examples of annotated terms highlighted in yellow.

Isti tekst, ali prikazan u veb čitaču je predstavljen na slici 22 gde se može videti oblak sa definicijom, kao i podvučeni primeri anotiranih termina.



postojećih resursa, kao i u stvaranju novih. Od tradicionalne do digitalne biblioteke Tradicionalna biblioteka algebra Uobičajena, klasična definicija tradicionalne biblioteke kaže da ona „prikuplja, obrađuje, čuva i distribuira informacije bez obzira na verske, političke, rasne, etničke, spolne, ideološke i druge karakteristike, u cilju širenja obrazovanja, kulture i nauke”. Definicija tradicionalne biblioteke kaže da je „sistematska i kompleksnih brojeva kao i proučavanjem sadržinskih i strukturalnih algebarskih struktura. [...] čuvaju6.” Pojam bibliotečka građa „obuhvata stampani i na drugi način umnoženi materijal koji je svojina biblioteke [...] Bibliotečku građu mogu činiti knjige, časopisi, novine, muzička dela, reprodukcije umetničkih slika i crteža, kartografske publikacije, planovi naselja, fonodokumenta, reprografska građa, plakati, oglasi, fotodokumenta, video i filmski zapisi, elektronskeknjige i periodika, baze podataka, rukopisi, memorisana građa i druge vrste dokumenata7.” Piter Brofi, opisujući pet modela biblioteka, kao prvi tip navodi tradicionalnu biblioteku. Prema Brofiju8,

Figure 22: Learner as the user of BAEKTEL

Slika 22. Učenik kao korisnik BAEKTEL-a

6 Conclusions

The described terminological support for BAEKTEL network represents a step forward in OER by blending two major sources of knowledge: academic and entrepreneurial. Its aim is to provide a learning continuum, from enrichment of university students' knowledge with entrepreneurial knowledge, which might prove invaluable in their future jobs, to keeping their academic knowledge up-to-date once they have graduated, by offering them a live-long learning opportunity as employees.

In this text elements of the language support system to open educational resources on BAEKTEL portal are given, through a review of lexical resources that are in the function of providing multilingualism of the BAEKTEL project. First, the process of computer-lexicographical production of online dictionaries is outlined, and then special attention is paid to terminology standards and translation tools. The text further describes in detail the Termi application developed within the BAEKTEL project for management of multilingual terminology. At its end, the document describes how to tag the open educational resources in BAEKTEL project, namely, how to connect them to the terminology application.

Multilinguality of open educational resources within the BAEKTEL project is strongly supported by the system for language support developed within the project and described in this document.

Predstavljena terminološka podrška BAEKTEL mreži je korak napred u OER sprezanju dva glavna izvora znanja: akademskog i preduzetničkog. Njen cilj je da obezbedi kontinualno obrazovanje, od obogaćivanja znanja studenata sa preduzetničkim, praktičnim znanjima, što bi im moglo biti od neprocenjive važnosti u njihovim budućim poslovima, kao mogućnost da osvežavaju svoje akademsko znanje i nakon diplomiranja nudeći im celoživotno učenje.

U ovom tekstu dati su elementi sistema za jezičku podršku otvorenim obrazovnim resursima na BAEKTEL portalu, kroz pregled leksičkih resursa koji su funkciji obezbeđivanja višejezičnosti BAEKTEL projekta. Najpre je izložen proces računarsko-leksikografske produkcije online rečnika, a zatim je posebna pažnja posvećena terminološkim standardima i prevodilačkim alatima. U tekstu je zatim detaljno opisana aplikacija Termi razvijena u okviru BAEKTEL projekta za upravljanje višejezičnom terminologijom. Na kraju dokumenta opisan je način obeležavanja otvorenih obrazovnih resursa u BAEKTEL projektu, odnosno njihovo povezivanje sa terminološkom aplikacijom.

Višejezičnost otvorenih obrazovnih resursa BAEKTEL projekta podržana je snažno sistemom za jezičku podršku razvijenom u okviru projekta i opisanom u ovom dokumentu.

7 References

- Barrachina, S., & al., e. (2009). Statistical Approaches to Computer-Assisted Translation. *Computational Linguistics*, 35(1), 3–28. doi:10.1162/coli.2008.07-055-r2-06-29
- Cenoz, J., & Gorter, D. (2011). A holistic approach to multilingual education: Introduction. *The Modern Language Journal*, 95(3), 339-343.
- Klosa, A. (2013). The lexicographical process (with special focus on online dictionaries). U R. H. Gouws, *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin/Boston: Walter de Gruyter.
- Krstev, C. (2008). *Processing of Serbian – Automata, Texts and Electronic dictionaries*. Belgrade: Faculty of Philology, University of Belgrade.
- Krstev, C., Pavlović-Lažetić, G., Vitas, D., & Obradović, I. (2004). Using Textual and Lexical Resources in Developing Serbian Wordnet. *Romanian Journal of Information Science and Technology*, 7(1-2), 147-161.
- Krstev, C., Stanković, R., Obradović, I., Vitas, D., & Utvić, M. (2010). Automatic Construction of a Morphological Dictionary of Multi-Word Units. U E. R. Hrafn Loftsson (Ur.). (str. 226-237). Reykjavik, Iceland: Lecture Notes in Computer Science 6233 Springer 2010.
- Stanković, R., Obradović, I., Kitanović, O., & Kolonja, L. (2012). Building Terminological Resources in an e-Learning Environment. *Proceedings of The Fourth International Conference on e-Learning (eLearning-2012), Belgrade, Serbia, 2012*, (str. 114-119).
- Stanković, R., Obradović, I., Krstev, C., & Vitas, D. (2011). Production of morphological dictionaries of multi-word units using a multipurpose tool. *Proceedings of the Computational Linguistics-Applications Conference, CLA '11*, 77-84.

This project has been funded with support from the European Commission. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Ovaj projekat se finansira uz podršku Evropske komisije. Publikacija odražava stavove autora, a Evropska komisija ne snosi odgovornost za bilo kakvu upotrebu informacija iz ove publikacije.