

Oznaka poročila: ARRS-RPROJ-ZP-2010-1/46

ZAKLJUČNO POROČILO O REZULTATIH RAZISKOVALNEGA PROJEKTA

A. PODATKI O RAZISKOVALNEM PROJEKTU

1. Osnovni podatki o raziskovalnem projektu

Šifra projekta	L6-9778
Naslov projekta	Slovenski terminološki portal
Vodja projekta	14681 Vojko Gorjanc
Tip projekta	L Aplikativni projekt
Obseg raziskovalnih ur	3.600
Cenovni razred	B
Trajanje projekta	01.2007 - 12.2009
Nosilna raziskovalna organizacija	581 Univerza v Ljubljani, Filozofska fakulteta
Raziskovalne organizacije - soizvajalke	106 Institut "Jožef Stefan" 582 Univerza v Ljubljani, Fakulteta za družbene vede
Družbeno-ekonomski cilj	10. Kultura, rekreacija, religija in sredstva javnega obveščanja

2. Sofinancerji¹

1.	Naziv	Amebis d.o.o. Kamnik
	Naslov	Bakovnik 3, 1241 Kamnik
2.	Naziv	
	Naslov	
3.	Naziv	
	Naslov	

B. REZULTATI IN DOSEŽKI RAZISKOVALNEGA PROJEKTA

3. Poročilo o realizaciji programa raziskovalnega projekta²

Informacijska družba je še posebej v evropskem prostoru postala velik izziv, ki je spodbudil in še spodbuja oblikovanje načel in metod za soočanje z izzivi prenosa raznovrstnih informacij na ravni jezika. Spoznanje o zares svobodni komunikaciji, ki jo pogojuje komunikacija v maternem jeziku, je privedlo do splošno sprejetega načela zagotavljanja možnosti kreativnega uresničevanja vsakega posameznika v svojem jeziku in hkratne izmenjave informacij v tujih jezikih. To velja za vsa področja človekovega delovanja in ustvarjanja, pri čemer je v na znanju temelječi družbi posebej izpostavljena strokovno-znanstvena komunikacija, torej komunikacija, kjer se srečujemo s

specializiranimi znanji. Za tovrstno komunikacijo pa je posebej pomembna terminologija – specializiran slovar določenega strokovnega/predmetnostnega področja. S projektom smo želeli ugotoviti, kakšni terminološki viri za slovenščino so vse na voljo in interes različnih strokovnih javnosti za njihovo povezovanje. Ker je v tujini na voljo več različnih terminoloških portalov, je bil naš namen pregledati rešitve, ki se ponujajo, in jih oceniti z vidika ustreznosti glede na specifikko slovenskega prostora. Projekt je oblikoval slovenski terminološki portal, ki obsega osnovne informacije o terminoloških načelih in terminološkem delu in vzorčne primere terminoloških podatkov v terminološki podatkovni zbirki <http://lojze.lugos.si/stp/index.html>. Ob tem se je pri projektu preverilo možnost povezave besedilnih virov in terminološke podatkovne zbirke.

Projekt je bil glede na zastavljene cilje razdeljen na tri vsebinske sklope. Prvi se je ukvarjal s pregledom področja terminoloških virov po posameznih področjih v slovenskem prostoru in ugotavljal, kakšne so tuje rešitve pri oblikovanju terminoloških portalov. Drugi in tudi osrednji sklop je bil povezan z oblikovanjem slovenskega terminološkega portala, pri tretjem slopu pa smo ugotavljali postopke za povezovanje besedilnih virov in pridobivanja terminološko relevantnih podatkov iz njih za slovenščino.

Prvi sklop: Pregled področja terminoloških virov

V prvem sklopu smo ugotovili, kakšni terminološki viri za slovenščino so vse na voljo in interes različnih strokovnih javnosti za njihovo povezovanje ter pregledali rešitve pri oblikovanju terminoloških portalov v tujini in jih ocenili z vidika ustreznosti glede na specifikko slovenskega prostora.

Terminološki viri za slovenščino

Projekti za zagotavljanje jezikovnih virov za slovenščino so bili do sedaj izrazito usmerjeni v gradnjo besedilnih korpusov, kar je tudi razumljivo, saj ti pomenijo osnovo jezikovne infrastrukture. Korpusno jezikoslovje je v slovenskem prostoru z zaključenimi projekti oblikovanja korpusov uspešno končalo prvo in seveda nujno potrebno fazo za nadaljnji razvoj. Ob tem je zaradi nujnega medstrokovnega sodelovanja pri gradnji korpusov oblikovalo tudi solidno izhodiščno platformo za širok razvoj področja jezikovnih virov za slovenščino. Oblikovani korpusi slovenskega jezika pa so bili pobudni tudi za vrsto celovitih korpusnih študij, tako enojezičnih kot tudi kontrastivnih, prav tako pa postajajo korpusi, še posebej referenčna korpusa FIDA in FIDA PLUS, vse bolj nepogrešljiv del jezikoslovnega raziskovalnega dela sploh, predvsem ko gre za leksikalne oz. leksikalnopomenske študije, manj, ko gre za terminološke, čeprav je tudi tu opazen napredek v smeri korpusno podprtega terminološkega dela.

Vsekakor je za celovitost jezikovnih virov treba obstoječim dodajati nove, med njimi vse bolj tudi terminološke. Smiselno bi bilo, da bi bili oblikovani tako, da omogočajo čim širšo dostopnost, hkrati pa bi bila velika prednost, če bi bili zasnovani na način, ki bi omogočal vključevanje čim širše zainteresirane strokovne javnosti. Terminološko delo namreč zahteva izjemno veliko časa in sredstev, zato je pomembno, da se pri tem uporablja učinkovite metode dela, postopke in orodja. Eden večjih problemov slovenskega prostora, ko gre za terminološko delo, je njegova nepovezanost. Poteka vrsta terminoloških aktivnosti in na voljo je v resnici tudi veliko terminoloških slovarjev in glosarjev, a so nepovezani in neenovito grajeni, velikokrat javnosti tudi nedostopni. Terminološki slovarji in drugi terminološki viri v veliki meri nastajajo v okviru Terminološke sekcije Inštituta za slovenski jezik Frana Ramovša ZRC SAZU, a tudi na vrsti drugih institucij, spodbudno pa je tudi, da so v Sloveniji vse bolj pobudna pri tem tudi neakademska okolja, na primer bančne institucije, mobilni operaterji itd., ne smemo pa zanemariti tudi velikega števila terminoloških virov, ki nastajajo v okviru univerzitetnega pedagoškega procesa kot del seminarских, diplomskih, magistrskih in doktorskih del.

Tuji terminološki portali

Kot izhodišče za oblikovanje slovenskega terminološkega portala je služila analiza stanja s tega področja v tujini. Področje javno dostopnih je terminoloških virov je zelo različno urejeno. Kot najboljši zgled dobre prakse pa lahko izpostavimo projekt *The EuroTermBank* z dostopom na <http://www.eurotermbank.com>, še posebej pri vprašanjih reševanja vprašanj avtorskih pravic in uporabe terminoloških standardov za predstavitev terminologije v podatkovni zbirki je projekt v veliki meri lahko spodbuden tudi za slovensko okolje. Gre za projekt velikega obsega, ki je tudi izjemno finančno podprt, tako da so glede obsega projekti, s katerimi bi se slovenski terminološki portal bistveno lažje primerjal, kanadski *Terminium* <http://www.terminiumplus.gc.ca>, *Nemški terminološki portal* (Deutsches Terminologie-Portal) <http://www.iim.fh-koeln.de/dtp> in *Slovaška*

terminološka podatkovna zbirka (Slovenská terminologická databáza)

<https://data.juls.savba.sk/std/>.

2 Drugi sklop: Oblikovanje slovenskega terminološkega portala

Namen drugega sklopa raziskave je bil oblikovanje slovenskega terminološkega portala.

Osnovna ideja je bila oblikovati izhodiščno platformo za združevanje nepovezanih raznorodnih terminoloških virov v enovito standardizirano podatkovno zbirko.

Glede na analizo trenutnega stanja v zvezi z razpoložljivimi terminološkimi viri lahko zaključimo, da so v slovenskem prostoru največje težave naslednje:

- razpršenost terminoloških virov,
- različnost metod in načel pri njihovem oblikovanju,
- pri elektronskih virih raznorodnost formatov zapisa in
- javna nedostopnost oz. pri elektronsko dostopnih vrsta različnih dostopnih točk.

Med potekom projekta se je za namene združevanja raznorodnih terminoloških virov v enovito podatkovno zbirko zbiralo vire v različnem formatu, pri katerih ni težav z njihovo dostopnostjo in kjer ni tudi nobenih težav glede avtorskih pravic.

Standardizacija

Da bi bila terminološka podatkovna zbirka primerljiva s sorodnimi v tujini, je pomembno, da sledi mednarodnim standardom s področja računalniškega zapisa terminoloških podatkov, kot so ISO 12620:1999 (Computer applications in terminology – Data categories), ISO 12200:1999 (Computer applications in terminology – Machine-readable terminology interchange), ISO 16642:2003 (Computer applications in terminology – Terminological markup framework) in ISO/DIS 30042.2 (Computer applications in terminology – Term-based exchange format specification). Slovenska terminološka podatkovna zbirka s svojo pojmovno orientiranostjo je za zapis uporablja format TBX kot standardizirani format za zapis jezikovnih podatkov s terminološko vrednostjo ([http://www.lisa.org/Term-Base-eXchange.32.0.html?&no_cache=1&sword_list\[\]=tbx](http://www.lisa.org/Term-Base-eXchange.32.0.html?&no_cache=1&sword_list[]=tbx)). Na podlagi standarda TBX je oblikovan zožen nabor podatkovnih kategorij, ki deloma sledijo tudi naboru kategorij TBX-Basic; slednji je bil oblikovan prav za namene manjših in srednje velikih terminoloških projektov (<http://www.lisa.org/Terminology-Special.102.0.html>). Tako TBX kot TBX-basic sta usklajena s standardi ISO.

Terminološka podatkovna zbirka integrira različne terminološke vire s heterogenimi vsebinskimi informacijami in zapisane v zelo različnih formatih, zato je v trenutni fazi razvoja predlagani prvi korak pretvorba v strukturo podatkovne zbirke XML z robustnim DTD-jem, ki omogoča integracijo različnih formalnih struktur obstoječih virov. Uporabljeni DTD temelji na spoznanjih projektov, ki so se ukvarjali prav z integracijo slovarskih podatkov v standardizirano strukturirano podatkovno zbirko, projektu CONCEDE, <http://www.itri.brighton.ac.uk/projects/concede> in TMF, <http://www.loria.fr/projets/TMF>. Osnovna ideja je, da pri avtomatski pretvorbi obstoječih virov v enovito označeno zbirko XML ohranimo vitalni del strukture prvotnega vira in enovito označimo čim več vsebinskih informacij, ki jih potrebujemo v naslednjem koraku, tj. pri pretvorbi podatkovne zbirke v formatu XML v standardizirano terminološko podatkovno zbirko v formatu TBX.

3 Tretji sklop: Povezovanje z besedilnimi viri

Za sodobne podatke o kateremkoli segmentu jezika so ključnega pomena podatki o jezikovni realnosti. Najboljši vir za to so besedilni korpusi, iz katerih s (pol)avtomatskimi tehnikami lahko izluščimo termine. Tehnologijo luščenja terminologije bi lahko opisali kot (pol)avtomatsko strokovno slovaropisje, vendar namen avtomatskega luščenja terminologije iz besedil ni samodejna izdelava slovarjev, temveč podpora različnim dejavnostim in tehnologijam, kjer je pisanje strokovnih slovarjev le ena izmed njih. Metoda omogoča pridobivanje terminološko relevantnih informacij iz besedilnih virov, ki so lahko v veliko pomoč tudi pri pripravi terminoloških slovarjev. Na spletni strani projekta <http://lojze.lugos.si/stp/index.html> je na voljo v okviru projekta inovativno razviti luščilnik terminov, ki temelji na kombinaciji statističnega jezikovno neodvisnega modela in jezikovnoodvisnih jezikovnih vzorcev, za katere se smatra, da so terminološko relevantni.

4. Ocena stopnje realizacije zastavljenih raziskovalnih ciljev³

Zastavljeni raziskovalni cilji so bili v celoti realizirani:

- 1) Pregled področja terminoloških virov: z raziskavo smo ugotovili, kakšni terminološki viri za slovenščino so vse na voljo in interes različnih strokovnih javnosti za njihovo povezovanje. Ker je v tujini na voljo več različnih terminoloških portalov, smo pregledali rešitve, ki se ponujajo, in jih ocenili z vidika ustreznosti glede na specifično slovenskega prostora. Na podlagi zgledov dobrih praks iz tujine je bila oblikovana idejna rešitev slovenskega terminološkega portala.
- 2) Osrednji cilj projekta je bil oblikovati slovenski terminološki portal, ki obsega osnovne informacije o terminoloških načelih in terminološkem delu, predvsem pa ponuja predstavitev terminologije v enovitem formatu, po enotnih načelih in na enovit način.
- 3) Povezovanje besedilnih in terminoloških virov: preizkusili smo, v kolikšni miri iz terminološko bogatih slovenskih besedilnih virov lahko (pol)avtomatsko pridobimo terminološko relevantne podatke, ki so potem osnova za gradnjo terminološke podatkovne zbirke. Na spletni strani projekta je predstavljena inovativna rešitev za luščilnik terminologije.

5. Utemeljitev morebitnih sprememb programa raziskovalnega projekta⁴

Sprememb programa raziskovalnega projekta ni bilo.

6. Najpomembnejši znanstveni rezultati projektne skupine⁵

Znanstveni rezultat		
1.	Naslov	<i>SLO</i> Terminološka veda in računalniško podprta terminografija
		<i>ANG</i> Terminology Management and Computer-aided Terminography
	Opis	<i>SLO</i> Poudarek monografije je na računalniško podprtem terminološkem delu, ki pomeni inovativen prispevek avtorice ne le k razvoju področja terminološke vede, ampak tudi slovenskega jezikoslovja sploh; področje terminološkega dela se tako z monografijo Špela Vintar umešča v sodobne svetovne tokove upravljanja in načrtovanja terminologije.
		<i>ANG</i> The book focuses on computer-based terminology management. Since modern terminology is increasingly concerned with eContent and knowledge management, this book promptly reflects current global trends in the field of terminology management and planning.
	Objavljeno v	VINTAR, Špela. Terminologija : terminološka veda in računalniško podprta terminografija, (Zbirka Prevodoslovje in uporabno jezikoslovje). 1. natis. Ljubljana: Znanstvena založba Filozofske fakultete, Oddelek za prevajalstvo, 2008. 132 str., ilustr. ISBN 978-961-237-245-3.
	Tipologija	2.01 Znanstvena monografija
	COBISS.SI-ID	239732736
2.	Naslov	<i>SLO</i> Slovenski terminološki portal
		<i>ANG</i> Slovene terminology web portal
	Opis	<i>SLO</i> V prispevku je predstavljen projekt "Slovenski terminološki portal", pri katerem je glavni cilj razvoj terminološkega portala s terminološko podatkovno zbirko v standardiziranem formatu TBX, zgrajeno na podlagi obstoječih terminoloških virov, informacijami o načelih terminološkega dela in orodji za luščenej terminologije.
		<i>ANG</i> The paper presents the Slovene Terminology Web Portal project. The main objective of the project is to develop a terminology portal which collects terminological data from various sources, provides basic information on the principles of terminological work, and offers various tools for termbase standardization and conversion into an open standard format TBX.
	Objavljeno v	GORJANC, Vojko, KREK, Simon, VINTAR, Špela. Slovene terminology web portal. V: Proceedings of the XIII EURALEX international congress (Barcelona, 15-19 July 2008). [Barcelona: Documenta universitaria, 2008], str. 971-978.
	Tipologija	1.08 Objavljeni znanstveni prispevek na konferenci
		37643874

	COBISS.SI-ID		
3.	Naslov	SLO	Načrtovanje in upravljanje terminologije
		ANG	Terminology planning and terminology management
	Opis	SLO	V članku predstavimo izhodišča terminološkega načrtovanja, ki ga razumemo kot upravljanje s pojmovnimi inovacijami tako na nadnacionalnem kot nacionalnem nivoju. Okvir terminološkega načrtovanja vključuje niz aktivnosti, ki so del v širšega sklop aktivnosti jezikovnega načrtovanja, prav tako pa tudi izobraževanja, prevajana in seveda razvoja znanosti in tehnologije. V članku ocenimo aktivnosti, povezane s terminološkim načrtovanjem v slovenskem prostoru, in predstavimo predlog, kako priti do koherentnejšega terminološkega načrtovanja pri nas.
		ANG	The article presents the basic principles of terminology planning, which is understood as management of conceptual innovation both on supranational and national levels. The framework of terminology planning includes a series of actions that are part of a broader set of language-planning activities as well as education, translation, and development of science and technology. The author assesses the activities related to terminology planning in Slovenia and puts forward a proposal on how to achieve a more coherent terminology planning.
	Objavljeno v	GORJANC, Vojko. Terminološko načrtovanje in upravljanje terminologije. Slavistična revija. [Tiskana izd.], 2010, letn. 58, št. 1, str. [95]-104.	
	Tipologija	1.01 Izvirni znanstveni članek	
	COBISS.SI-ID	41699682	
4.	Naslov	SLO	Idejna rešitev slovenskega terminološkega portala
		ANG	Conceptual solution for Slovene terminology web portal
	Opis	SLO	V prispevku predstavljamo projekt Slovenskega terminološkega portala. Izhodiščni cilj raziskovalnega projekta je ugotoviti, kakšni terminološki viri za slovenščino so vse na voljo in interes različnih strokovnih javnosti za njihovo povezovanje. Ker je v tujini na voljo več različnih terminoloških portalov, je naš namen tudi pregledati rešitve, ki se ponujajo, in jih oceniti z vidika ustreznosti glede na specifično slovenskega prostora.
		ANG	The paper presents the Slovene Terminology Web Portal project. The goals of the research project are the following: The overview of terminology resources: the study aims to establish which resources are available for the Slovenian language and to determine the interest of different professional fields for their cooperation. As there are various terminology portals available for other languages, our aim is to examine the existing solutions and to evaluate them from the point of view of the specific Slovenian situation.
	Objavljeno v	GORJANC, Vojko. Slovenski terminološki portal. V: LEDINEK, Nina (ur.), ŽAGAR, Mojca (ur.), HUMAR, Marjeta (ur.). Terminologija in sodobna terminografija. Ljubljana: Založba ZRC, ZRC SAZU, 2009, str. 303-310.	
	Tipologija	1.16 Samostojni znanstveni sestavek ali poglavje v monografski publikaciji	
	COBISS.SI-ID	40751970	
5.	Naslov	SLO	Korpusni pristop k izdelavi terminoloških slovarjev
		ANG	Corpus-based compilation of terminological dictionaries
	Opis	SLO	Prispevek opisuje postopke izbora in obdelave gesel za namen gradnje terminološkega slovarja odnosov z javnostmi. Osnova za izbor gesel so samodejno izdelani sezname terminoloških kandidatov, se pravi besed in besednih zvez, ki na podlagi statističnih obdelav in oblikoskladenjskih značilnosti izstopajo kot za stroko pomembne. Ročni pregled tako pridobljenih seznamov pokaže njihovo uporabnost, pa tudi nekatere probleme, povezane z razlikovanjem terminov od neterminov in terminov od terminoloških kolokacij.
		ANG	The article describes the procedures and treatment of entries for the purposes of a PR terminological dictionary. The selection of entries is based on automatically generated lists of field term candidates, i.e., words and collocations that stand out as relevant to a particular field based on statistical analyses and their morphosyntactic features. Manual examination of lists of terms acquired in this way indicates their usefulness but at the

		same time reveals some problems related to distinguishing between proper terms and non-terms, or between terms and terminological collocations.
Objavljeno v		LOGAR BERGINC, Nataša, VINTAR, Špela. Korpusni pristop k izdelavi terminoloških slovarjev: od besednih seznamov in konkordanc do samodejnega luščenja izrazja. Jez. slovst. (Tisk. izd.). [Tiskana izd.], sep.-okt. 2008, letn. 53, št. 5, str. [3]-17.
Tipologija	1.01	Izvirni znanstveni članek
COBISS.SI-ID	28004445	

7. Najpomembnejši družbeno-ekonomsko relevantni rezultati projektne skupine⁶

Družbeno-ekonomsko relevantni rezultat			
1.	Naslov	SLO	Slovenski terminološki portal in s TBX združljiv poenostavljeni DTD
		ANG	Slovene Terminology Web Portal and the TBX-Compatible Simplified DTD/schema
Opis		SLO	Predstavljen je poenostavljeni DTD, ki je eden od modulov terminološkega portala. S svojo enostavno strukturo označevanja obstoječih virov omogoča pretvorbo heterogenih zapisov terminoloških podatkov v novo standardno označeno podatkovno zbirko formata TBX, dostopno na internetu.
		ANG	The article primarily focuses on the DTD/schema which lies at the core of the web portal. It enables the unifying of heterogeneous terminological data and provides a very basic initial markup for new terminology resources: TBX database available on the web.
Šifra		B.03	Referat na mednarodni znanstveni konferenci
Objavljeno v		KREK, Simon, GORJANC, Vojko, ARHAR, Špela. Slovene Terminology Web Portal and the TBX-Compatible Simplified DTD/schema. V: 6th International Conference on Language Resources and Evaluation, Marrakech, Morocco, May 26 - June 1, 2008. LREC 2008 : proceedings. Marrakech: ELRA, 2008, 5 str.	
Tipologija		1.08	Objavljeni znanstveni prispevek na konferenci
COBISS.SI-ID		37644898	
2.	Naslov	SLO	Slovenski jezikovni viri: od korpusov do terminološke podatkovne zbirke
		ANG	Slovene language resources: from corpora to terminology data-base
Opis		SLO	Obstoječi jezikovni viri za slovenščino bi morali biti dopolnjeni z novimi, med drugim tudi terminološkimi. Obstaja vrsta terminoloških virov, a so metodološko heterogeni in velikokrat javnosti nedostopni. Glavna ideja Slovenskega terminološkega portala je združiti obstoječe vire v enovito terminološko podatkovno zbirko.
		ANG	The existing Slovene language resources should be supplemented with new ones, among them also terminological resources. A significant number of terminological resources exist, but they are methodologically heterogeneous and often unavailable to the general public. The main goal of the Slovene Terminology Web Portal project is to unify terminology resources in a single termbase.
Šifra		B.05	Gostujoči profesor na inštitutu/univerzi
Objavljeno v		GORJANC, Vojko. Slovene language resources : from corpora to terminology data-base : [predavanje na oddelku za slavistiko Univerze v Göteborgu 29. 4. 2008]. 29. 4. 2008.	
Tipologija		3.14	Predavanje na tuji univerzi
COBISS.SI-ID		36800354	
3.	Naslov	SLO	Samodejno luščenje terminologije
		ANG	Automatic term extraction
Opis		SLO	Prispevek predstavlja samodejno luščenja terminologije iz eno- in večjezičnih korpusov. V uvodnem delu predavanja naštejemo glavne metode pri luščenju terminologije, ki za ugotavljanje terminološkosti uporabljajo tako statistične kot jezikoslovne podatke iz korpusa. Odkar imamo tudi za slovenščino na razpolago oblikoskladenjsko označene korpusne, temeljijo metode za samodejno luščenje predvsem na t. i. oblikoskladenjskih vzorcih, s katerimi

		iz besedil luščimo potencialno terminološke besedne zveze, te pa se nato s pomočjo statističnih meril razvrščajo glede na svojo terminološko relevantnost.	
	ANG	The paper presents automatic term extraction from mono- and multilingual corpora. The introductory part of the paper outlines the main methods employed in term extraction utilizing both statistical and linguistic corpus data. Since the introduction of the first freely available morphosyntactic annotation tool for Slovene, methods in term extraction have been primarily based on morphosyntactic patterns which help us extract potentially relevant terms from the corpus. These candidate phrases are subsequently ranked according to a statistical measure of termhood.	
	Šifra	B.03 Referat na mednarodni znanstveni konferenci	
	Objavljeno v	VINTAR, Špela. Samodejno luščenje terminologije: izkušnje in perspektive = Automatic term extraction: experiences and perspectives. V: LEDINEK, Nina (ur.), ŽAGAR, Mojca (ur.). Zbornik povzetkov. Ljubljana: Inštitut za slovenski jezik Frana Ramovša, ZRC SAZU, [2008], str. 40.	
	Tipologija	1.12 Objavljeni povzetek znanstvenega prispevka na konferenci	
	COBISS.SI-ID	28535597	
4.	Naslov	SLO	Na besedilnih korpusih temelječa leksikalnopodatkovna zbirka za slovenščino
		ANG	Corpus based lexical database for Slovenian language
	Opis	SLO	Mentorstvo pri doktorski disertaciji mladi raziskovalki
		ANG	Young researcher PhD supervisor
	Šifra	D.09	Mentorstvo doktorandom
	Objavljeno v		ARHAR, Špela. Na besedilnih korpusih temelječa leksikalnopodatkovna zbirka za slovenščino : [doktorska naloga]. Ljubljana: [Š. Arhar], 2009. VIII, 230 f., ilustr.
	Tipologija	2.08	Doktorska disertacija
	COBISS.SI-ID		40243810
5.	Naslov	SLO	Izdelava slovenskega semantičnega leksikona z uporabo eno- in večjezičnih jezikovnih virov
		ANG	Creating Slovene semantic lexicon based on mono- and multilingual resources
	Opis	SLO	Mentorstvo pri doktorski disertaciji mladi raziskovalki
		ANG	Young researcher PhD supervisor
	Šifra	D.09	Mentorstvo doktorandom
	Objavljeno v		FIŠER, Darja. Izdelava slovenskega semantičnega leksikona z uporabo eno- in večjezičnih jezikovnih virov : doktorska disertacija. Ljubljana: [D. Fišer], 2009. X, 155 f., ilustr., graf. prikazi, tabele.
	Tipologija	2.08	Doktorska disertacija
	COBISS.SI-ID		40174690

8. Drugi pomembni rezultati projektne skupine⁷

Ključni rezultati so navedeni v točkah 6 in 7.

9. Pomen raziskovalnih rezultatov projektne skupine⁸

9.1. Pomen za razvoj znanosti⁹

SLO

Razviti jezikoslovni postopki in jezikovne tehnologije za slovenski jezik krepijo jezikovne vire za slovenščino. Tako je za slovenščino oblikovan pomemben segment jezikovne infrastrukture. Projektno delo sodi na področje uporabnega jezikoslovja, v okviru katerega doprinaša k

- (a) razvoju načel za spletno predstavitev terminoloških slovarjev,
- (b) razvoju metodologije korpusnega jezikoslovja v segmentu oblikovanja specializiranih korpusov strokovnih jezikov,
- (c) razvoju sodobnih metod za analizo strukturiranih podatkov v besedilnih korpusih,

(č) razvoju sodobnih metod povezovanja besedilnih virov, njihove analize in predstavitve rezultatov analize v slovarju in
(d) razvoju sodobnih metod dostopa do slovenskih terminoloških virov.
V segmentu uporabnega jezikoslovja projekt krepi pozicijo korpusnega jezikoslovja v slovenskem prostoru. Korpusno jezikoslovje se je v Sloveniji dokončno oblikovalo kot samostojno raziskovalno izhodišče z gradnjo splošnih korpusov. Pri terminološkem delu pa korpusni pristop v slovenskem prostoru do sedaj ni bil prisoten prav v veliki meri, tako da projekt pomeni pomemben zasuk terminološkega dela v smer korpusnega pristopa, ki temelji na analizi velike količine jezikovnih podatkov, zbranih v besedilnih korpusih. Tako projekt terminološko delo v Sloveniji umešča v sodobne smeri razvoja terminološkega dela v Evropi in svetu.

ANG

Linguistic procedures and language technology developed within the project have improved language resources for Slovene and have made an important new resource available. The project belongs to the field of applied linguistics and it generates:

- the development of the principles of web presentation of specialized dictionaries
- the development of corpus linguistics methodology concerning compilation of specialized corpora of professional language
- the development of modern methods for structured data analysis in text corpora
- the development of modern methods of linking language resources, their analysis and the presentation of the results in a dictionary
- the development of modern methods for accessing Slovene terminological resources

In applied linguistics domain it strengthened the role of corpus linguistics in Slovene context. With the compilation of general corpora, corpus linguistics has finally established itself as an independent research field in Slovenia. Until now the corpus approach has not been present in terminological activities in Slovenia and the project represents an important advance towards this type of approach in terminological work which is based on the analysis of extensive data collected from text corpora. The field of terminology in Slovenia has therefore joined the modern trends of terminological work in Europe and the rest of the world.

9.2. Pomen za razvoj Slovenije¹⁰

SLO

V današnji družbi znanja in visoko razvitih tehnologij je za vsak jezik posebnega pomena razvoj jezikovne infrastrukture, za katero moramo poskrbeti sami. Predlagani projekt jo razvija v segmentu terminologije in s tem nudi podporo za učinkovito strokovno komunikacijo tako v slovenskem jeziku kot pri prenosu informacij v slovenščino in obratno. V globalno povezanem svetu postaja dostopnost do informacij v določenem jeziku in povezljivost z drugimi jeziki ključnega pomena, zato je oblikovanje raznorodnih jezikovnih virov za slovenščino velikega pomena.

V svetu se močno podpira terminološko delo, ker spodbuja tehnološki transfer. To prav gotovo velja tudi za slovensko okolje. Z učinkovitim terminološkim upravljanjem namreč lahko veliko pripomoremo k hitrejšemu prenosu novih tehnoloških rešitev v slovensko okolje, prav tako pa učinkovito predstavimo svoje dosežke v tujem jeziku.

Podatki govorijo o tem, da se preko sodobnih medijev prenaša največ, kar 80 % vseh podatkov, povezanih s specializiranimi znanji in vedenji. Tako je za učinkovito izmenjavo tovrstnih informacij terminološko delo izjemnega pomena, pomeni pa tudi možnost ekonomske učinkovitosti gospodarskih subjektov, ki v tem okolju delujejo, saj zagotavlja hitro in učinkovito izmenjavo novih informacij. Podatki kažejo, da je obvladovanje specializirane komunikacije pri prenosu iz enega v drug jezik pomemben dejavnik gospodarskega razvoja.

Komunikacijske težave velikokrat niso posledica neznanja tujih jezikov, ampak terminoloških zadreg. Uporaba neustrezne terminologije pa lahko privede do negativnih učinkov ne le, ko gre za gospodarsko učinkovitost, ampak lahko privede v primerih varnostnih regulativ in navodil za uporabo zaradi napačnih prevodov tudi do poškodb, na ravni izobraževanja slabše učinkovitosti, pri tehnični dokumentaciji zaradi neustreznega medjezikovnega prenosa visokih stroškov, pri pogodbah in pravnih dokumentih zmanjšanja pravne varnosti itd. Na splošno lahko rečemo, da ima zagotavljanje terminoloških virov tako kvalitativne kot kvantitativne učinke. Rezultati projekta zagotavljajo učinkovitejšo komunikacijo in imajo tako velik pomen pri odpravljanju zgoraj navedenih negativnih učinkov, ki jih lahko povzročita neustrezen medjezikovni prenos.

ANG

In the modern society based on knowledge and advanced technologies, it is crucial that adequate infrastructure be developed for every language – something we are required to do ourselves. The project has developed such infrastructure in the field of terminology and the results offer support for efficient professional communication both in the Slovene language as well as in transferring information into Slovene and vice versa. In the globally interconnected world, accessibility of information in a particular language, as well as its potential to interrelate with other languages, is of vital importance. Therefore, the compilation of various language resources for Slovene is very important. In contemporary society, efficient professional and interprofessional communication is vital. Of course, not only within one language but also between pairs of various languages. Professional communication requires accuracy which can be achieved solely through a systematic, concept-oriented terminological type of work, such as put forward in the proposal of the Slovene terminology web portal project.

Work in the field of terminology is widely supported worldwide, for it enhances the transfer of technology. Undoubtedly, this is also true of Slovenia. With efficient terminology management we can not only encourage a faster transfer of technological solutions to the Slovene environment, but also present Slovene technological achievements in a foreign language.

The figures show that the majority, as much as 80%, of all data related to specialized knowledge and skills is transferred through modern media. Developing terminology is therefore extremely important for a successful exchange of relevant information. By ensuring a fast and efficient exchange of new information, it offers economic subjects operating in this environment an opportunity for better performance. According to the relevant data, skilled transfer of specialized information from one language to another is an important factor in economic development.

Communication problems often arise from insufficient knowledge of terminology rather than a lack of general language skills. The incorrect use of terminology can not only produce negative effects in terms of economic performance - due to inadequate translations - but can even cause harm or injury in situations involving safety regulations or instructions for use. Furthermore, occurring on the level of education inadequate terminology can reduce efficiency, in technical documentation unsuccessful interlingual transfer can increase the costs, in contracts and legal documents it can reduce the level of legal protection etc. It can be claimed that developing terminology resources has effects both in terms of quality and quantity. As a publicly available terminology resource, the Slovene Terminology Web Portal project ensures efficient communication and consequently plays an important role in eliminating the negative effects listed above.

10. Samo za aplikativne projekte!

Označite, katerega od navedenih ciljev ste si zastavili pri aplikativnem projektu, katere konkretne rezultate ste dosegli in v kakšni meri so doseženi rezultati uporabljeni

Cilj	
F.01	Pridobitev novih praktičnih znanj, informacij in veščin
	Zastavljen cilj <input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat <input type="text" value="Dosežen"/>
	Uporaba rezultatov <input type="text" value="V celoti"/>
F.02	Pridobitev novih znanstvenih spoznanj
	Zastavljen cilj <input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat <input type="text" value="Dosežen"/>
	Uporaba rezultatov <input type="text" value="V celoti"/>
F.03	Večja usposobljenost raziskovalno-razvojnega osebja
	Zastavljen cilj <input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat <input type="text" value="Dosežen"/>
	Uporaba rezultatov <input type="text" value="V celoti"/>
F.04	Dvig tehnološke ravni

	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.05	Sposobnost za začetek novega tehnološkega razvoja	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.06	Razvoj novega izdelka	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.07	Izboljšanje obstoječega izdelka	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.08	Razvoj in izdelava prototipa	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.09	Razvoj novega tehnološkega procesa oz. tehnologije	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.10	Izboljšanje obstoječega tehnološkega procesa oz. tehnologije	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.11	Razvoj nove storitve	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.12	Izboljšanje obstoječe storitve	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.13	Razvoj novih proizvodnih metod in instrumentov oz. proizvodnih procesov	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>

	Uporaba rezultatov	<input type="text"/>
F.14	Izboljšanje obstoječih proizvodnih metod in instrumentov oz. proizvodnih procesov	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.15	Razvoj novega informacijskega sistema/podatkovnih baz	
	Zastavljen cilj	<input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat	Dosežen <input type="text"/>
	Uporaba rezultatov	Uporabljen bo v naslednjih 3 letih <input type="text"/>
F.16	Izboljšanje obstoječega informacijskega sistema/podatkovnih baz	
	Zastavljen cilj	<input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat	Dosežen <input type="text"/>
	Uporaba rezultatov	Uporabljen bo v naslednjih 3 letih <input type="text"/>
F.17	Prenos obstoječih tehnologij, znanj, metod in postopkov v prakso	
	Zastavljen cilj	<input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat	Dosežen <input type="text"/>
	Uporaba rezultatov	Uporabljen bo v naslednjih 3 letih <input type="text"/>
F.18	Posredovanje novih znanj neposrednim uporabnikom (seminarji, forumi, konference)	
	Zastavljen cilj	<input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat	Dosežen <input type="text"/>
	Uporaba rezultatov	V celoti <input type="text"/>
F.19	Znanje, ki vodi k ustanovitvi novega podjetja ("spin off")	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.20	Ustanovitev novega podjetja ("spin off")	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.21	Razvoj novih zdravstvenih/diagnostičnih metod/postopkov	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.22	Izboljšanje obstoječih zdravstvenih/diagnostičnih metod/postopkov	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>

F.23	Razvoj novih sistemskih, normativnih, programskih in metodoloških rešitev	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.24	Izboljšanje obstoječih sistemskih, normativnih, programskih in metodoloških rešitev	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.25	Razvoj novih organizacijskih in upravljavskih rešitev	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.26	Izboljšanje obstoječih organizacijskih in upravljavskih rešitev	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.27	Prispevek k ohranjanju/varovanje naravne in kulturne dediščine	
	Zastavljen cilj	<input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat	Dosežen <input type="text"/>
	Uporaba rezultatov	V celoti <input type="text"/>
F.28	Priprava/organizacija razstave	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.29	Prispevek k razvoju nacionalne kulturne identitete	
	Zastavljen cilj	<input checked="" type="radio"/> DA <input type="radio"/> NE
	Rezultat	Dosežen <input type="text"/>
	Uporaba rezultatov	V celoti <input type="text"/>
F.30	Strokovna ocena stanja	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.31	Razvoj standardov	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.32	Mednarodni patent	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE

	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.33	Patent v Sloveniji	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.34	Svetovalna dejavnost	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>
F.35	Drugo	
	Zastavljen cilj	<input type="radio"/> DA <input checked="" type="radio"/> NE
	Rezultat	<input type="text"/>
	Uporaba rezultatov	<input type="text"/>

Komentar

-

11. Samo za aplikativne projekte!

Označite potencialne vplive oziroma učinke vaših rezultatov na navedena področja

	Vpliv	Ni vpliva	Majhen vpliv	Srednji vpliv	Velik vpliv	
G.01	Razvoj visoko-šolskega izobraževanja					
G.01.01.	Razvoj dodiplomskega izobraževanja	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
G.01.02.	Razvoj podiplomskega izobraževanja	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
G.01.03.	Drugo:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02	Gospodarski razvoj					
G.02.01	Razširitev ponudbe novih izdelkov/storitev na trgu	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.02.	Širitev obstoječih trgov	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.03.	Znižanje stroškov proizvodnje	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.04.	Zmanjšanje porabe materialov in energije	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.05.	Razširitev področja dejavnosti	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.06.	Večja konkurenčna sposobnost	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.07.	Večji delež izvoza	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.08.	Povečanje dobička	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.09.	Nova delovna mesta	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.10.	Dvig izobrazbene strukture zaposlenih	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.02.11.	Nov investicijski zagon	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

G.02.12.	Drugo:		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.03	Tehnološki razvoj						
G.03.01.	Tehnološka razširitev/posodobitev dejavnosti		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.03.02.	Tehnološko prestrukturiranje dejavnosti		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.03.03.	Uvajanje novih tehnologij		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.03.04.	Drugo:		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.04	Družbeni razvoj						
G.04.01	Dvig kvalitete življenja		<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.04.02.	Izboljšanje vodenja in upravljanja		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.04.03.	Izboljšanje delovanja administracije in javne uprave		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.04.04.	Razvoj socialnih dejavnosti		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.04.05.	Razvoj civilne družbe		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.04.06.	Drugo:		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.05.	Ohranjanje in razvoj nacionalne naravne in kulturne dediščine in identitete		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
G.06.	Varovanje okolja in trajnostni razvoj		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.07	Razvoj družbene infrastrukture						
G.07.01.	Informacijsko-komunikacijska infrastruktura		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	
G.07.02.	Prometna infrastruktura		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.07.03.	Energetska infrastruktura		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.07.04.	Drugo:		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.08.	Varovanje zdravja in razvoj zdravstvenega varstva		<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	
G.09.	Drugo:		<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	

Komentar

-

12. Pomen raziskovanja za sofinancerje, navedene v 2. točki¹¹

1.	Sofinancer	Amebis d.o.o. Kamnik		
	Vrednost sofinanciranja za celotno obdobje trajanja projekta je znašala:		26.484,00	EUR
	Odstotek od utemeljenih stroškov projekta:		25,00	%
	Najpomembnejši rezultati raziskovanja za sofinancerja			Šifra
		1.	KREK, Simon, GORJANC, Vojko, ARHAR, Špela. Slovene Terminology Web Portal and the TBX-Compatible Simplified DTD/schema. V: 6th International Conference on Language Resources and Evaluation, Marrakech.	B.03
	2.	VINTAR, Špela. Samodejno luščenje terminologije: izkušnje in perspektive = Automatic term extraction: experiences and	B.03	

		perspectives. V: LEDINEK, Nina (ur.), ŽAGAR, Mojca (ur.). Zbornik povzetkov.	
	3.	Mentorstvo mlade raziskovalke. ARHAR, Špela. Na besedilnih korpusih temelječa leksikalnopodatkovna zbirka za slovenščino : [doktorska naloga]. Ljubljana: [Š. Arhar], 2009. VIII, 230 f., ilustr.	D.09
	4.		
	5.		
	Komentar	-	
	Ocena	<p>Rezultati projekta so relevantni za gradnjo leksikalnih podatkovnih zbirk, s katerimi se ukvarja podjetje Amebis. Gre za metodološke postopke, razvite v okviru projektnega dela, ki omogočajo (a) (pol)avtomatsko pretvarjanje raznorodnih leksikalnih podatkov v enovito strukturirane in zapisane podatke in (b) analizo besedil z novimi tehnikami in postopki, ki omogočajo luščenje leksikalnih podatkov iz besedilnih korpusov. Za uporaba je slednje še posebej pomembno, in sicer tako avtomatsko pridobivanje terminoloških podatkov iz terminološko bogatih besedilnih virov, kar omogoča razvoj luščilnika terminologije, kot tudi drugih leksikalnih podatkov iz referenčnih korpusov. Pri slednjih gre za sodelovanje tudi v okviru dela mlade raziskovalke, zaposlene v podjetju Amebis, s katerim se vpeljuje celostno uporabo korpusnih podatkov za namene gradnje leksikalne podatkovne zbirke s (pol)avtomatskimi analitičnimi tehnikami in postopki. Jasna umestitev na korpusno jezikoslovno raziskovalno področje z namenom njegovega razvoja za slovenščino je usmeritev, ki jo lahko samo pozdravimo, še posebej zato, ker je namen njenega dela oblikovati in ovrednotiti obstoječe nekorpusne leksikalne podatkovne zbirke ter jih sopostavi tistim, ki nastajajo z analizo korpusa. Ker gre za uporabno (pol)avtomatskih analitičnih tehnik in postopkov, njeno raziskovalno delo slovenistično raziskovanje povezuje s področjem avtomatske obdelave naravnega jezika ter tako na področje slovenistične računalniške leksikografije vnaša svež pristop, ki ga v podjetju Amebis lahko samo pozdravimo, saj bodo rezultati neposredno uporabni pri nadgradnji obstoječih Amebisovih leksikalnih podatkovnih zbirk.</p>	
2.	Sofinancer		
		Vrednost sofinanciranja za celotno obdobje trajanja projekta je znašala:	EUR
		Odstotek od utemeljenih stroškov projekta:	%
		Najpomembnejši rezultati raziskovanja za sofinancerja	Šifra
	1.		
	2.		
	3.		
	4.		
	5.		
	Komentar		
	Ocena		
3.	Sofinancer		
		Vrednost sofinanciranja za celotno obdobje trajanja projekta je znašala:	EUR

Odstotek od utemeljenih stroškov projekta:		%
Najpomembnejši rezultati raziskovanja za sofinancerja		Šifra
1.		
2.		
3.		
4.		
5.		
Komentar		
Ocena		

C. IZJAVE

Podpisani izjavljam/o, da:

- so vsi podatki, ki jih navajamo v poročilu, resnični in točni
- se strinjamo z obdelavo podatkov v skladu z zakonodajo o varstvu osebnih podatkov za potrebe ocenjevanja, za objavo 6., 7. in 8. točke na spletni strani <http://sicris.izum.si/> ter obdelavo teh podatkov za evidence ARRS
- so vsi podatki v obrazcu v elektronski obliki identični podatkom v obrazcu v pisni obliki
- so z vsebino zaključnega poročila seznanjeni in se strinjajo vsi soizvajalci projekta

Podpisi:

Vojko Gorjanc	in	
podpis vodje raziskovalnega projekta		zastopnik oz. pooblaščen oseba RO

Kraj in datum:

Ljubljana

18.4.2010

Oznaka poročila: ARRS-RPROJ-ZP-2010-1/46

¹ Samo za aplikativne projekte. [Nazaj](#)

² Napišite kratko vsebinsko poročilo, kjer boste predstavili raziskovalno hipotezo in opis raziskovanja. Navedite ključne ugotovitve, znanstvena spoznanja ter rezultate in učinke raziskovalnega projekta. Največ 18.000 znakov vključno s presledki (približno tri strani, velikosti pisave 11). [Nazaj](#)

³ Realizacija raziskovalne hipoteze. Največ 3.000 znakov vključno s presledki (približno pol strani, velikosti pisave 11). [Nazaj](#)

⁴ Samo v primeru bistvenih odstopanj in sprememb od predvidenega programa raziskovalnega projekta, kot je bil zapisan v predlogu raziskovalnega projekta. Največ 3.000 znakov vključno s presledki (približno pol strani, velikosti pisave 11). [Nazaj](#)

⁵ Navedite največ pet najpomembnejših znanstvenih rezultatov projektne skupine, ki so nastali v času trajanja projekta v okviru raziskovalnega projekta, ki je predmet poročanja. Za vsak rezultat navedite naslov v slovenskem in angleškem jeziku (največ 150 znakov vključno s presledki), rezultat opišite (največ 600 znakov vključno s presledki) v slovenskem in angleškem jeziku, navedite, kje je objavljen (največ 500 znakov vključno s presledki), izberite ustrezno šifro tipa objave po Tipologiji dokumentov/del za vodenje bibliografij v sistemu COBISS ter napišite ustrezno COBISS.SI-ID številko bibliografske enote. Navedeni rezultati bodo objavljeni na spletni strani <http://sicris.izum.si/>.

PRIMER (v slovenskem jeziku):

Naslov: Regulacija delovanja beta-2 integrinskih receptorjev s katepsinom X;

Opis: Cisteinske proteaze imajo pomembno vlogo pri nastanku in napredovanju raka. Zadnje študije kažejo njihovo povezanost s procesi celičnega signaliziranja in imunskega odziva. V tem znanstvenem članku smo prvi dokazali... (največ 600 znakov vključno s presledki)

Objavljeno v: OBERMAJER, N., PREMZL, A., ZAVAŠNIK-BERGANT, T., TURK, B., KOS, J.. Carboxypeptidase cathepsin X mediates $\beta 2$ - integrin dependent adhesion of differentiated U-937 cells. Exp. Cell Res., 2006, 312, 2515-2527, JCR IF (2005): 4.148

Tipologija: 1.01 - Izvirni znanstveni članek

COBISS.SI-ID: 1920113 [Nazaj](#)

⁶ Navedite največ pet najpomembnejših družbeno-ekonomsko relevantnih rezultatov projektne skupine, ki so nastali v času trajanja projekta v okviru raziskovalnega projekta, ki je predmet poročanja. Za vsak rezultat navedite naslov (največ 150 znakov vključno s presledki), rezultat opišite (največ 600 znakov vključno s presledki), izberite ustrezen rezultat, ki je v Sifrantu raziskovalnih rezultatov in učinkov (Glej: <http://www.arrs.gov.si/sl/gradivo/sifranti/sif-razisk-rezult.asp>), navedite, kje je rezultat objavljen (največ 500 znakov vključno s presledki), izberite ustrezno šifro tipa objave po Tipologiji dokumentov/del za vodenje bibliografij v sistemu COBISS ter napišite ustrezno COBISS.SI-ID številko bibliografske enote.

Navedeni rezultati bodo objavljeni na spletni strani <http://sicris.izum.si/>. [Nazaj](#)

⁷ Navedite rezultate raziskovalnega projekta v primeru, da katerega od rezultatov ni mogoče navesti v točkah 6 in 7 (npr. ker se ga v sistemu COBISS ne vodi). Največ 2.000 znakov vključno s presledki. [Nazaj](#)

⁸ Pomen raziskovalnih rezultatov za razvoj znanosti in za razvoj Slovenije bo objavljen na spletni strani: <http://sicris.izum.si/> za posamezen projekt, ki je predmet poročanja. [Nazaj](#)

⁹ Največ 4.000 znakov vključno s presledki [Nazaj](#)

¹⁰ Največ 4.000 znakov vključno s presledki [Nazaj](#)

¹¹ Rubrike izpolnite/prepišite skladno z obrazcem "Izjava sofinancerja" (<http://www.arrs.gov.si/sl/progproj/rproj/gradivo/>), ki ga mora izpolniti sofinancer. Podpisan obrazec "Izjava sofinancerja" pridobi in hrani nosilna raziskovalna organizacija – izvajalka projekta. [Nazaj](#)

Obrazec: ARRS-RPROJ-ZP/2010 v1.00a

8D-C9-0C-4F-CA-D8-1D-D6-F2-77-F6-EA-08-11-B7-0E-F8-5D-29-F3