


**ZAKLJUČNO POROČILO**  
**O REZULTATIH OPRAVLJENEGA RAZISKOVALNEGA DELA**  
**NA PROJEKTU V OKVIRU CILJNEGA RAZISKOVALNEGA**  
**PROGRAMA (CRP) »KONKURENČNOST SLOVENIJE 2006 – 2013«**


**REPUBLIKA SLOVENIJA**  
**NOSILEC JAVNEGA POOBLASTILA**  
**AGENCIJA ZA RAZISKOVALNO DEJAVNOST**  
**REPUBLIKE SLOVENIJE, LJUBLJANA**

**I. Predstavitev osnovnih podatkov raziskovalnega projekta**

1. Naziv težišča v okviru CRP:

Težišče 1: Konkurenčnejše gospodarstvo in hitrejša rast  
 Tematski sklop: Informacijska družba

Prejeto: 20-05-2009	Sig. z.: OKO
Šifra zadeve: 63113-4/2007	Pril.:
	Vrednost: 13

2. Šifra projekta:

V2-0380

3. Naslov projekta:

Elektronsko besedilno središče z multimedijško komunikacijo

3. Naslov projekta

3.1. Naslov projekta v slovenskem jeziku:

Elektronsko besedilno središče z multimedijško komunikacijo

3.2. Naslov projekta v angleškem jeziku:

Electronic multimedial text center

4. Ključne besede projekta

4.1. Ključne besede projekta v slovenskem jeziku:

Slovensko slovstvo, elektronske izdaje, znanstvene izdaje, XML, TEI, Unikod

4.2. Ključne besede projekta v angleškem jeziku:

Slovenian literature, digital editions, scholarly editing, XML, TEI, Unicode

5. Naziv nosilne raziskovalne organizacije:

Znanstvenoraziskovalni center Slovenske akademije znanosti in umetnosti (ZRC SAZU)

5.1. Seznam sodelujočih raziskovalnih organizacij (RO):

Inštitut Jožef Stefan

6. Sofinancer/sofinancerji:

Javna agencija za raziskovalno dejavnost Republike Slovenije (ARRS)  
Ministrstvo za visoko šolstvo, znanost in tehnologijo

7. Šifra ter ime in priimek vodje projekta:

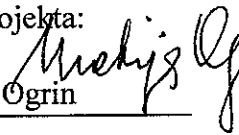
16207

Matija Ogrin

Datum: 19. maj 2009

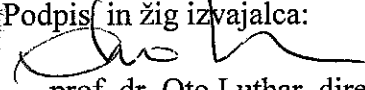
Podpis vodje projekta:

doc. dr. Matija Ogrin



Podpis in žig izvajalca:

prof. dr. Oto Luthar, direktor



## II. Vsebinska struktura zaključnega poročila o rezultatih raziskovalnega projekta v okviru CRP

### 1. Cilji projekta:

1.1. Ali so bili cilji projekta doseženi?

- a) v celoti  
 b) delno  
 c) ne

Če b) in c), je potrebna utemeljitev.

1.2. Ali so se cilji projekta med raziskavo spremenili?

- a) da  
 b) ne

Če so se, je potrebna utemeljitev:

## 2. Vsebinsko poročilo o realizaciji predloženega programa dela<sup>1</sup>:

Projekt Elektronsko besedilno središče z multimedijško komunikacijo je zasnovan kot izrazito večdisciplinaren projekt. Ima močno tehnološko podlago, njegova vsebina pa povezuje v sodelovanje več humanističnih disciplin: literarno zgodovino, splošno zgodovino, narodopisje in jezikoslovje.

Temeljna ambicija projekta je bila izdelati zameetek slovenske elektronske besedilne zbirke, to je zbirke elektronskih izdaj slovenskih besedil, ki bi bila zasnovana po premišljenih edicijskih načelih in podprta z enotno označevalno tehnologijo. Temeljila naj bi na odprtih mednarodnih standardih računalniškega označevanja (XML, TEI, Unikon), kar omogoča:

- da so notranje strukture besedil maksimalno operabilne – t.j. pripravljene za raznovrstne analize in prezentacije, vključno z večmedijskimi povezavami (zlasti povezava besedil in faksimilov ali besedil in njihovih avdioposnetkov);
- da so e-izdaje ne le javno dostopne, ampak tudi interoperabilne, t.j. združljive z drugimi podobnimi besedilnimi zbirkami;
- celotna zbirka izdaj naj bo odporna na tehnološke spremembe in neodvisna od strojne in programske opreme (platform-independent).

Ta temeljni cilj naj bi projekt uresničil z naslednjimi konkretnimi nalogami:

1. Izdelati vsaj dva primera zgledne, metodološko že izpopolnjene elektronske monografske izdaje, kjer se besedilo bralcu odpira v raznih zahtevnostnih prikazih in v vsaj dveh medijih (slika, besedilo ali zvok, besedilo). Takšna izdaja naj služi kot zameetek besedilne zbirke, uporabne tako v pedagoške kakor raziskovalne namene.

2. Ob tem izgotoviti tudi poglobljena orodja in mehanizme, ki bodo potrebni za bodočo hitro in cenejšo širitev vsebine besedilnega središča v večjih razsežnostih, kot so pretvorni vmesnik, iskalni mehanizmi.

Ambicija projekta je bila tudi, da javna dostopnost in uporabnost tu izdelanih orodij omogoči tudi drugim sorodnim institucijam, da začnejo izdelovati podobne izdaje.

Načrt dela NI predvideval posebnega repozitorija (strežniškega sistema) elektronskih besedil; dejstvo, da smo tak repozitorij kljub temu izdelali, kakor bo razloženo v nadaljevanju, je pomemben presežek projektnega načrta.

ad 1

Humanistični del projekta je potekal kot temeljna raziskava več vrst slovstvenega gradiva, iz katerega naj bi po programu nastali dve znanstveni elektronski izdaji. Z racionalizacijo in sodelovanjem je bil ta cilj presežen: namesto dveh izdaj je bila vsebinsko docela pripravljena še tretja, ki je trenutno v sklepnih fazah in bo objavljena jeseni 2009. Pri tem je vsaka od sodelujočih ved analizirala svoje besedilno gradivo; nato je ugotovila, katere temeljne strukture besedila je potrebno elektronsko označiti in jih s tem računalniku eksplicirati; na tej osnovi je vsaka sodelujoča disciplina napravila za svoj besedilni objekt ustrezen edicijski koncept. Na tej podlagi je potekala podrobnejša raziskava in obdelava

<sup>1</sup> Potrebno je napisati vsebinsko raziskovalno poročilo, kjer mora biti na kratko predstavljen program dela z raziskovalno hipotezo in metodološko-teoretičen opis raziskovanja pri njenem preverjanju ali zavračanju vključno s pridobljenimi rezultati projekta.

besedil: metodološko različne transkripcije, komentiranje besedil, analitične študije. Raziskana, edicijsko obdelana in analizirana so bila besedila, ki so kot načrtovani neposredni rezultat projekta javno dostopna v naslednjih elektronskih izdajah.

ad 1 a)

Mestna prisežna besedila v slovenskem jeziku do začetka 19. stoletja  
Prepise, opombe in spremna besedila napisal Boris Golec. Uredil M. Ogrin  
Mesto objave: <http://nl.ijs.si/e-zrc/prisege/>

Objava se kot prva loteva celotnega nabora slovenskih priseg enega tipa ustvarjalcev. Zajema vsa znana – že objavljena, še neobjavljena in novo odkrita prisežna besedila –, ki so nastala pri poslovanju mestnih organov sedmih slovenskih mest do vključno prvega desetletja 19. stoletja. Časovni razpon obravnavanih besedil obsega vsaj dve stoletji in pol, skrajni možni razpon pa razmejujeta letnici 1493 in 1808. Besedila izvirajo iz šestih kranjskih mest (Ljubljana, Novo mesto, Kranj, Metlika, Višnja Gora in domnevno Kamnik) ter iz enega štajerskega (Ormož).

Po mestih nastanka so razvrščena v osem poglavij, ki vključujejo različno število besedil, od enega do sedemnajst, skupaj 36. Vsako poglavje se deli na naslednje vsebinske sklope: Zvrsti besedil, Diplomatični in kritični prepis, Sedanje in prejšnja hranišča, Dosedanje objave in obravnave, Oblikovne značilnosti, Okoliščine in čas nastanka, Vsebina, Pravopisne in jezikovne značilnosti, Pomen in vloga v družbenem okolju, Literatura.

Razen tistih sedmih besedil, katerih izvorniki veljajo za pogrešane oziroma izgubljene, so vsa besedila objavljena v treh različicah: kot faksimile ter v obliki diplomatičnega in kritičnega prepisa.

Glede na zvrst je kar 29 priseg ali 80 % službenih oziroma vdanostnih, pri čemer gre praviloma za prisežne obrazce za večkratno uporabo in le izjemoma za enkratni zapis prisežnega besedila. Preostalih sedem priseg ali manj kot 20 % vseh je sodnih, od tega trije prisežni obrazci (vključno z dvema poukoma prisežencem) in štirje zapisi iz sodne prakse.

Primarni namen prisežnih besedil je bila njihova razumljivost v ciljnem, lokalnem okolju. V strukturi, skladnji in leksiki se jim pozna, da so dobesedni prevodi nemških predlog oziroma da so sestavljeni po nemških zgledih. Močnih vplivov uradovalne nemščine prav tako ne morejo skriti prisege iz sodne prakse, med obravnavanimi sicer najbolj izvorna besedila. Posebno dragoceno je politično, sodno in upravno besedišče, ki ne glede na mesto nastanka kaže opazne skupne poteze in priča o splošni uveljavljenosti izrazja v osrednjeslovenskem prostoru. Močno jezikovno odstopanje je zaznati pri besedilih iz vzhodnoštajerskega mesteca Ormož.

Pomen raziskave in e-izdaje kot končnega rezultata je mdr. v tem, da so s to izdajo slovenska uradovalna besedila iz dobe med 16. in 19. stoletjem sploh prvič v celoti zbrana, hkrati pa posredovana javnosti s faksimili, natančnimi prepisi ter študijami v prostem dostopu na spletu.

ad 1 b)

Pripovedi iz Bavšice

Posnela in komentirala Barbara Ivančič Kutin. Fonetični prepis in študije napisali Karmen Kenda Jež in Barbara Ivančič Kutin. Uredil M. Ogrin

Mesto objave: <http://nl.ijs.si/e-zrc/pripovedi>

V elektronski objavi Pripovedi iz Bavšice je predstavljen del repertoarja folklornih

pripovedi pripovedovalke Tine Kravanja (1923–2005) iz Bavšice pri Bovcu. Monografija obsega dvaindvajset folklornih in spominskih pripovedi v narečnem bovškem govoru, ki so bile na terenu zvočno dokumentirane leta 2001 in bodo uporabnikom na voljo v treh različnih oblikah:

- v digitaliziranem zvočnem posnetku
- v dosledni fonetični transkripciji posnetega narečnega govora
- v prepisu, približanem knjižnemu jeziku, vendar z ohranjenimi lokalizmi.

Monografijo začena teoretični del, ki ob primeru tu objavljenih folklornih pripovedi obravnava metodološko problematiko prenosa živega govora v zapisano besedilo. Pomemben element teoretičnega dela predstavljata poglavji z jezikoslovnim (dialektološkim) orisom bovškega govora in s tem povezano poglavje o redakciji zvočnih postnekov, fonetičnih prepisov in knjižne ustreznice besedil. Drugi del monografije obsega same pripovedi in komentar.

Vsaka faza obravnave je zahtevala svojstven pristop: pri prekodiranju živega pripovedovanja v vse ostale naknadne, tehnično posredovane oblike je bila pozornost usmerjena najprej na to, kako ohraniti kar se da avtentično podobo pripovedi v smislu procesa pripovedovanja (še le sekundarno tudi pripovedi kot vsebinsko zaokroženo enote). Ker je tonski posnetek prikrajšan za vizualne elemente, grafični zapis pa tudi za akustične elemente, ki v živem govoru predstavljajo eno od bistvenih sredstev za estetsko in dramatično oblikovanje, je bilo treba poiskati rešitev, ki bo bralcu omogočala kolikor moč sintetičen pogled na ravnino izvirne besedilne teksture (t.j. načina, kako je bila pripoved v živo izvedena). Poleg opomb te vrste, ki so vpletene v transkripcijo, spremljajo prepise pripovedi še druge opombe (o samem tekstu, njegovem kontekstu in redakciji) in sezname narečnega besedja. Poleg opisa vseh okoliščin pripovedovanja in dokumentiranja ter značilnosti in posebnosti bovškega narečnega jezika so v spremnih tekstih izpostavljeni konkretni problemi in rešitve pri pretvorbi (prekodiranju) na vseh stopnjah dela.

V tehničnem pogledu izhaja besedilo pripovedi iz zvočnih posnetkov slabše kvalitete na avdiokasetah, posnetih na navadnem prenosnem kasetniku ali diktafonu z vgrajenim mikrofonom. Pri nadaljnji digitalni obdelavi je bilo pomembno predvsem to, da se pridobi čim večja koristna glasnost (govora pripovedovalcev). V ta namen smo najprej ročno odstranili nekaj šumov (kašljanje, ropot), da je bilo moč povečati glasnost celotnega posnetka (normalizacija). Nekaj koristne glasnosti smo pridobili še s kompresijo, k razumljivosti govora pa je precej pripomoglo rezanje basov (pod 120 Hz), ki niso nosili govorne informacije (bobnenje strojev). Odstranili smo tudi nekaj previsokih frekvenc izven govornega območja in uporabili določeno mero redukcije šuma.

Tako je bila za vsako pripoved izdelana po ena večja zvočna datoteka. Te so v izdaji integrirane s prepisi, tako da je mogoče bodisi ob fonetičnem prepisu bodisi ob knjižni redakciji s klikom poklicati zvočno datoteko in potem brati pripoved ob hkratnem poslušanju njene prvotne zvočne podobe. Možnost takšne primerjave obogati tako fonetični prepis kakor knjižno ustreznico in sproži v bralcu predstavo o razponu razlik med živim govorom in njegovim zapisom.

Kombinacija besedila in zvočnega posnetka (kasete, zgoščenka) v narodopisnih študijah ni novost, ampak uveljavljena metoda predstavljanja besedil. Novost elektronske izdaje Pripovedi iz Bavšice je v tem, da kot ena prvih prinaša pravi dialektološki fonetični prepis v sopostavitvi s knjižno izpeljanko pripovedi, oboje pa integrira z zvočnim posnetkom. To omogoča – glede na doslej znano prakso na tem področju – najpopolnejšo predstavitev celotnega upovedovalnega loka od izvirnega govora na terenu prek njegove jezikoslovne fonetične določitve vse do njegove knjižne rekonstruirane, pravzaprav interpretirane možnosti.

ad 1 c)

Najstarejši pregledi slovstva na Slovenskem. Elektronska znanstvenokritična izdaja  
Prepisi, prevodi, spremne študije, uredništvo: Luka Vidmar s sodelavci

Leta 2007 so se začele priprave na znanstvenokritično izdajo prepisov in prevodov štirih najstarejših pregledov slovstva na Slovenskem iz dobe reformacije, protireformacije in baroka, in sicer spisov Register vnd summarischer Inhalt (1560) Primoža Trubarja, Memoriale scribendorum (1605) Tomaža Hrena, poglavje o kranjskih pisateljih iz Slave vojvodine Kranjske Janeza Vajkarda Valvasorja (1689) in Bibliotheca Labacensis publica (1715–1719) Janeza Gregorja Dolničarja. Že v istem letu (2007) je bilo pripravljenih okoli 200 skenogramov Dolničarjevega rokopisa iz Semeniške knjižnice v Ljubljani in okoli 50 skenogramov originalnega Valvasorjevega tiska. Izvirna teksta Trubarja in Hrena, natisnjena v več desetletjih starih izdajah, sta bila pretvorjena v elektronsko besedilno obliko. Istega leta sta bila dokončana prevod Valvasorjevega nemškega (J. Moder) in prepis Dolničarjevega latinskega spisa (V. Pavlin).

Leta 2008 je bil pripravljen kritični prepis Valvasorjevega tiska (M. Deželak Trojar), dokončani pa so bili tudi vsi prevodi iz latinščine v slovenščino: prevod Hrenove spomenice (M. Deželak Trojar), prevod latinskih odlomkov v Valvasorjevem delu (P. Simoniti) in prevod Dolničarjevega dela (N. Vilhelm). Prav tako je bila sestavljena bibliografija, ki je zajela izdaje štirih pregledov in študije o njih (M. Grum). Vsi prepisi in prevodi so bili pregledani, urejeni in usklajeni ter opremljeni z ustreznim aparatom, in sicer z opombami, kazali in študijo (L. Vidmar). Do aprila leta 2009 je bila celota pripravljena za elektronsko izdajo. Trenutno je monografija v uredniški fazi (strokovni pregled, korekture, elektronska obdelava besedil) in bo izšla jeseni 2009.

Elektronska monografija o najstarejših pregledih slovenskega slovstva je tretji vsebinski rezultat projekta (namesto načrtovanih dveh!). Z njo je delovni načrt projekta tudi po vsebinski strani presežen. Zaradi kompleksne in zelo zahtevne vsebine se je proces uredniške priprave podaljšal in bo monografija objavljena jeseni. Z njo se potrjuje, da je slovenska literatura že v najzgodnejših obdobjih postala objekt misli, predmet zgodovinopisja in duhovne refleksije.

ad 2

Tehnološki del projekta je za vse zgoraj navedene znanstvene izdaje zagotovil ustrezno podlago; ta je izvedena tako, da se bo s projektom začeto delo lahko v naslednjih letih smotno nadaljevalo: pretvorna in prezentacijska orodja bodo omogočala (točka ad 2 a), da se v delo vključi več sodelavcev, tudi z drugih inštitucij. Skupaj z repozitorijem besedil (točka ad 2 b) je s tem ustvarjena kompaktna, urejena sistemska platforma za nadaljnje vsebinsko delo.

ad 2 a)

Spletni pretvornik RTF→XML v prostem dostopu

Implementirali smo pretvornik iz formata RTF v format XML-TEI, prosto dostopen na spletnem naslovu <http://nl.ijs.si/e-zrc/rtf2tei/>. Namen pretvornika je omogočiti raziskovalcem, sami pretvorijo svoja znanstvena besedila v bogato strukturiran, semantično obogaten zapis XML, primeren za elektronsko objavo in trajnejšo hrambo. To lahko storijo tako, da besedilo in strukturo svoje znanstvene izdaje zapišejo v urejevalniku Word, ki omogoča hitro in enostavno delo in na katerega so raziskovalci iz humanistike

navajeni, nato pa preko našega spletnega servisa to besedilo pretvorijo v osnovni format XML-TEI in nato še v prikaz tega formata v zapisu HTML. S tem se raziskovalci izognejo zamudnemu in bolj okornemu pisanju besedil neposredno v formatu XML z rabo oznak konzorcija TEI.

Word je urejevalnik besedil, ki je predvsem usmerjen k videzu, k prikazu besedil, medtem ko oznake XML po priporočilih TEI opisujejo pomen delov besedila, nad gnezdenjem oz. strukturiranjem teh oznak pa je pri tem postavljenih bistveno več omejitev. Zato se lahko zgodi, da se besedilo v zapisu Word ne pretvori v pravilno zapisan TEI. Da se takim situacijam izognemo, so na spletni strani pretvornika podana podrobnejša navodila, poleg tega pa omogoča pretvornik tudi validacijo zapisa TEI in ogled besedila v formatu HTML, ki je izveden iz formata XML-TEI.

ad 2 b)

Repozitorij in poizvedovalni sistem

Poleg objave teh besedil v obliki statičnih datotek na spletnem strežniku (znanstvenokritične izdaje eZISS – <http://nl.ijs.si/e-zrc/>) smo pripravili tudi njihovo prezentacijo in shranjevanje v obliki repozitorija na dokumentacijskem strežniku. S tem smo hkrati ustvarili robusten repozitorij za raznolike besedilne zbirke kakor tudi poizvedovalni sistem za iskanje po shranjenih besedilih. Uporabnikova vstopna točka v repozitorij je portal eZISS.

Sodobne možnosti digitalizacije dokumentov so spodbudile razvoj standardov in implementacij dokumentacijskih strežnikov, namenjenih za delo z arhivskimi, rokopisnimi, jezikoslovnimi, slikovnimi in drugimi viri. Uporabili smo sistem Fedora Commons, ki je nastal na osnovi dela na področju digitalizacije arhivskih, dokumentalističnih, jezikoslovnih in drugih humanističnih virov na Univerzi v Virginiji in na Institute for Advanced Technology in the Humanities v Virginiji, ZDA, v sodelovanju z raziskovalno skupino za digitalne knjižnice na Univerzi Cornell (Flexible Extensible Digital Object Repository Architecture). Fedora Commons je danes ena od vodilnih rešitev za digitalne knjižnice dokumentov v humanistiki.

Fedora Commons je fleksibilno ogrodje za gradnjo dokumentnih strežnikov. Podpira shranjevanje dokumentov, upravljanje z dokumenti, diseminacijo kompleksnih dokumentnih objektov ter upravljanje z odnosi med objekti. Strežnik omogoča združevanje lokalnih virov z oddaljenimi viri v združene objekte s skupnimi storitvami za prikaz in obdelavo. Tako ima lahko posamezen objekt več načinov prikaza, ki so lahko tudi dinamični. Arhitektura strežnika omogoča tudi dinamičen sistem odnosov med objekti in komponentami (uporabljena je tehnologija podatkovne zbirke trojčkov RDF za zapis relacij), ki ima podporo za iskalni jezik. Strežnik je razvit kot prenosna spletna aplikacija v jeziku Java in podpira standardne spletne aplikativne vmesnike REST (preprostejši vmesnik, ki ga podpirajo vsi brskalniki) in SOAP (aplikacijski vmesnik s podporo za XML) ter vmesnik Open Archive Interchange (OAI) za zajemanje metapodatkov iz digitalnih zbirk in knjižnic.

Poleg osnovnega sistema za iskanje po metapodatkovni bazi objektov ima strežnik Fedora tudi odprt vmesnik za kompleksnejše poizvedovalne rešitve – trenutno so dostopni gonilniki za iskalno-poizvedovalna sistema Lucene in Solr, ki podpirata vrsto naprednih tipov poizvedb po celotnem besedilu izdaj, ter iskalni sistem Zebra s podporo za vrsto iskalnih in poizvedovalnih jezikov in standardov, med drugim standard Z39.50 ANSI/NISO za poizvedovanje po elektronskih knjižnicah.

Ogrodje Fedora Commons z vsemi dodatnimi komponentami ter strežniškim spletnim okoljem Tomcat je dostopno z izvorno kodo in pod prostoprogramsko licenco Apache



(izjema je Zebra z licenco GPL). Tudi naša implementacija in naše specifične programske rešitve so prosto dostopne zainteresirani javnosti pod licenco Apache.

Naša implementacija ima obliko večnamenskega dokumentnega strežnika. Dokumenti so ločeni na zbirke glede na kontekst objave, tako da bo strežnik uporaben tudi za tekoče in prihodnje projekte (del implementacije je že nastal v okviru projekta Neznani rokopisi slovenskega slovstva 17. in 18. stoletja – NRSS in je bila racionalno uporabljena tudi za pričujoči projekt). Posamezna zbirka ima svoj iskalni in poizvedovalni vmesnik, prilagojen specifičnim potrebam projekta in lastnostim objavljenih dokumentov. Za razliko od dosedanjih uporab sistema (zlasti NRSS), kjer so bili dokumenti zaradi narave podatkov predstavljeni kot zbirka elementov z medsebojnimi odnosi, so zaradi kompleksnosti kritične in drugih znanstvenih izdaj posamezne publikacije v strežniku shranjene kot en objekt, ki vsebuje izvorni zapis v obliki datoteke XML v zapisu TEI, ter vrsto zunanjih podatkovnih zapisov, ki predstavljajo posamezne segmente publikacije v obliki HTML, primerni za prikaz na spletu.

Prednosti takšne objave poleg statične spletne objave so možnost sledenja posodobitvam publikacije, sistem za poizvedovanje, programski vmesniki REST in SOAP ter vmesnik za zajem metapodatkov OAI.

## Sklep

Namen večdisciplinarnega projekta Elektronsko besedilno središče je bil, po eni strani, izpopolniti metodologijo priprave elektronskih znanstvenih edicij na humanističnem področju in to metodo tudi praktično uresničiti z dvema izgotavljenima, kompleksnima znanstvenima izdajama. S tehnološke strani je bil cilj projekta izdelati elektronska orodja, ki bodo razmeroma težavno tehnološko pot do bogato strukturiranega zapisa XML–TEI kot temelja modernih elektronskih znanstvenih izdaj poenostavila in izdelavo takšnih izdaj približala širšemu krogu humanističnih znanstvenikov.

Oba cilja sta bila kljub zelo nizkemu proračunu projekta dosežena; delovni program projekta je bil ne le uresničen, ampak z racionalnim usklajevanjem človeških ter materialnih virov tudi presežen.

Literarnozgodovinski, historični in narodopisni cilji so bili doseženi z raziskavami, iz katerih sta nastali v projektu dve znanstveni monografiji: Mestna prisežna besedila v slovenskem jeziku do začetka 19. stoletja (avtor Boris Golec, ur. M. Ogrin) in Folklorne pripovedi iz Bavšice (avtorica Barbara Ivančič Kutin, ur. M. Ogrin). V prvi elektronski izdaji je avtor zbral vse razpoložljive oz. ohranjene uradovalne dokumente v slovenskem jeziku mestnega izvora od 16. do začetka 19. stoletja, kar je prvi tovrsten poskus doslej. Monografija uporablja celotno historiografsko metodologijo, dopolnjeno tudi s filološko metodo: od iskanja ter analize virov (rokopisov), prek dveh metodološko različnih prepisov (diplomatični, kritični prepis) do ocene historičnega pomena teh listin v originalnem socialnem okolju. Metodološko sega torej Golčeva monografija o mestnih prisegah od diplomatike in paleografije do intrepeticije; s sodelovanjem jezikoslovcev je pridobila tudi pomembno historično-jezikoslovno vrednost (kritični prepis, analiza pravopisnih posebnosti, slovar historičnega besedja). Vsem tem metodičnim postopkom je koncept elektronske izdaje v zapisu TEI odgovoril z ustreznimi strukturami in jih praktično uresničil. Z drugo elektronsko monografijo – Folklorne pripovedi iz Bavšice – je realizirana kompleksna metodologija narodopisnega raziskovanja, tesno povezanega s slovensko dialektologijo: sega od teoretične razčlenbe odnosov med živim pripovedovanjem, posnetkom in zapisom govora prek jezikoslovnega (dialektološkega) opisa govora do samega terenskega snemanja folklornih pripovedi (Bavšica nad Bovcem)

in njihove večmedialne predstavitve (digitalizirani zvočni posnetki, fonetični prepis, knjižno-jezikovni zapis). Tako elektronska izdaja predstavi folklorne pripovedi v treh metodološko različnih modusih (posnetek in dva prepisa), hkrati pa v največji možni meri posreduje njihovo izvorno fakturo. Tudi tem metodološkim potrebam je elektronska izdaja lahko odgovorila z vsemi potrebnimi možnostmi. Tretja izdaja, ki predstavlja presežek programa dela – najstarejši pregledi slovenskega slovstva (Trubar, Hren, Valvasor, Dolničar) – je v uredniški proceduri in bo izšla jeseni 2009.

Tehnološka komponenta projekta Elektronsko besedilno središče nudi tem izdajam več kot le trdno funkcionalno podlago. Deluje marveč tudi kot servis, ki z javnim pretvornikom iz zapisa RTF v XML-TEI ( <http://nl.ijs.si/e-zrc/rtf2tei/>) olajša pripravo kompleksnih elektronskih znanstvenih edicij širšemu krogu humanističnih in družboslovnih raziskovalcev, s tem pa nudi kvalitetno javno (in brezplačno) storitev. Enako pomembna tehnološka sestavina je repozitorij besedil in poizvedovalni sistem Fedora Commons (nameščen za portalom eZISS – <http://nl.ijs.si/e-zrc/> ), ki omogoča hitre in kvalitetne poizvedbe po obsežnih podatkovnih ali besedilnih zbirkah. (Pripominjamo, da je bilo obsežno programje Fedora Commons mogoče implementirati v okviru našega projekta le zato, ker je bil deloma že implementiran v drugačni obliki za projekt NRSS – Neznani rokopisi slovenskega slovstva 17. in 18. stoletja; s takšno racionalno izrabo sicer pičlih sredstev predstavlja ta rezultat pomemben presežek prvotnega projektnega načrta.) Z obema tehnološkima sistemoma se portal eZISS iz manjše zbirke znanstvenih izdaj spreminja v obsežnejši, kompakten sistem, namenjen shranjevanju raznolikih elektronskih besedil ter inteligentnemu iskanju po tekstih kot tudi njihovi uspešni javni prezentaciji ter diseminaciji. Namen projekta Elektronsko besedilno središče z multimedijško komunikacijo je s tem v polni meri dosežen.

### 3. Izkoriščanje dobljenih rezultatov:

3.1. Kakšen je potencialni pomen<sup>2</sup> rezultatov vašega raziskovalnega projekta za:

- a) odkritje novih znanstvenih spoznanj;
- b) izpopolnitev oziroma razširitev metodološkega instrumentarija;
- c) razvoj svojega temeljnega raziskovanja;
- d) razvoj drugih temeljnih znanosti;
- e) razvoj novih tehnologij in drugih razvojnih raziskav.

3.2. Označite s katerimi družbeno-ekonomskimi cilji (po metodologiji OECD-ja) sovpadajo rezultati vašega raziskovalnega projekta:

- a) razvoj kmetijstva, gozdarstva in ribolova - Vključuje RR, ki je v osnovi namenjen razvoju in podpori teh dejavnosti;
- b) pospeševanje industrijskega razvoja - vključuje RR, ki v osnovi podpira razvoj industrije, vključno s proizvodnjo, gradbeništvom, prodajo na debelo in drobno, restavracijami in hoteli, bančništvom, zavarovalnicami in drugimi gospodarskimi dejavnostmi;
- c) proizvodnja in racionalna izraba energije - vključuje RR-dejavnosti, ki so v funkciji dobave, proizvodnje, hranjenja in distribucije vseh oblik energije. V to skupino je treba vključiti tudi RR vodnih virov in nuklearne energije;
- d) razvoj infrastrukture - Ta skupina vključuje dve podskupini:
  - transport in telekomunikacije - Vključen je RR, ki je usmerjen v izboljšavo in povečanje varnosti prometnih sistemov, vključno z varnostjo v prometu;
  - prostorsko planiranje mest in podeželja - Vključen je RR, ki se nanaša na skupno načrtovanje mest in podeželja, boljše pogoje bivanja in izboljšave v okolju;
- e) nadzor in skrb za okolje - Vključuje RR, ki je usmerjen v ohranjanje fizičnega okolja. Zajema onesnaževanje zraka, voda, zemlje in spodnjih slojev, onesnaženje zaradi hrupa, odlaganja trdnih odpadkov in sevanja. Razdeljen je v dve skupini:
- f) zdravstveno varstvo (z izjemo onesnaževanja) - Vključuje RR - programe, ki so usmerjeni v varstvo in izboljšanje človekovega zdravja;
- g) družbeni razvoj in storitve - Vključuje RR, ki se nanaša na družbene in kulturne probleme;
- h) splošni napredek znanja - Ta skupina zajema RR, ki prispeva k splošnemu napredku znanja in ga ne moremo pripisati določenim ciljem;
- i) obramba - Vključuje RR, ki se v osnovi izvaja v vojaške namene, ne glede na njegovo vsebino, ali na možnost posredne civilne uporabe. Vključuje tudi varstvo (obrambo) pred naravnimi nesrečami.

<sup>2</sup> Označite lahko več odgovorov.

3.3. Kateri so **neposredni rezultati** vašega raziskovalnega projekta glede na zgoraj označen potencialni pomen in razvojne cilje?

Takšen neposredno koristen rezultat je zgoraj opisani pretvornik iz formata RTF v obogaten zapis XML, ki je prosto dostopen na naslovu <http://nl.ijs.si/e-zrc/rtf2tei/>. S pomočjo tega spletnega servisa lahko zainteresirani humanisti in družboslovci sami pretvorijo svoja znanstvena besedila v trajno uporaben in elektronski objavi primeren format XML - s tem pa si olajšajo pot do znanstvene objave. Enako pomembna je razvita metodologija, praktično udejanjena z dvema zgoraj navedenima znanstvenima monografijama v elektronski obliki.

3.4. Kakšni so lahko **dolgoročni rezultati** vašega raziskovalnega projekta glede na zgoraj označen potencialni pomen in razvojne cilje?

Iztek tega projekta ni samo konec neke raziskave, ampak je tembolj začetek, iniciacija širše aktivnosti, ki je dobila temelj z zgoraj omenjenima spletnim pretvornikom in besedilnim repozitorijem. Pričakujemo, da se bo ob tej tehnološki kakor tudi metodološki platformi v bodoče povezala dejavnost večjega števila slovenskih humanistov in družboslovcev z raznolikimi znanstvenimi objavami v elektronski obliki. S projektom ustvarjena sredstva to možnost vsekakor odpirajo.

3.5. Kje obstaja verjetnost, da bodo vaša znanstvena spoznanja deležna zaznavnega odziva?

- a) v domačih znanstvenih krogih;
- b) v mednarodnih znanstvenih krogih;
- c) pri domačih uporabnikih;
- d) pri mednarodnih uporabnikih.

3.6. Kdo (poleg sofinancerjev) že izraža interes po vaših spoznanjih oziroma rezultatih?

Praktičen interes je že pokazala komercialna založba Mohorjeva družba Celje, ki je našo metodo elektronskega objavljanja zahtvenih znanstvenih izdaj sprejela in začela po tej metodi pripravljati lastne publikacije v sodelovanju s člani tega projekta.

3.7. Število diplomantov, magistrrov in doktorjev, ki so zaključili študij z vključenostjo v raziskovalni projekt?

Mag. Luka Vidmar, član projektne skupine in MR, je v marcu 2009 oddal doktorsko disertacijo in čaka na promocijo doktorata. Prav on je pripravil izdajo najstarejših pregledov slovstva na Slovenskem.

#### 4. Sodelovanje z tujimi partnerji:

4.1. Navedite število in obliko formalnega raziskovalnega sodelovanja s tujimi raziskovalnimi inštitucijami.

--

4.2. Kakšni so rezultati tovrstnega sodelovanja?

--

**5. Bibliografski rezultati<sup>3</sup> :**

*Za vodjo projekta in ostale raziskovalce v projektni skupini priložite bibliografske izpise za obdobje zadnjih treh let iz COBISS-a) oz. za medicinske vede iz Inštituta za biomedicinsko informatiko. Na bibliografskih izpisih označite tista dela, ki so nastala v okviru pričujočega projekta.*

**6. Druge reference<sup>4</sup> vodje projekta in ostalih raziskovalcev, ki izhajajo iz raziskovalnega projekta:**

Kot posebno referenco in dosežek štejejo tudi to, da je Slovenska akademija znanosti in umetnosti prav v času trajanja našega projekta (in pod tem vplivom) sprejela odločitev, da pripravi svojo elektronsko izdajo Slovenskega biografskega leksikona po metodi in tehnologiji, ki sta uporabljeni tudi v našem projektu.
---

<sup>3</sup> Bibliografijo raziskovalcev si lahko natisnete sami iz spletne strani:<http://www.izum.si/>

<sup>4</sup> Navedite tudi druge raziskovalne rezultate iz obdobja financiranja vašega projekta, ki niso zajeti v bibliografske izpise, zlasti pa tiste, ki se nanašajo na prenos znanja in tehnologije. Navedite tudi podatke o vseh javnih in drugih predstavitev projekta in njegovih rezultatov vključno s predstavitvami, ki so bile organizirane izključno za naročnika/naročnike projekta.

## Summary

### Electronic multimedia text centre

The project Electronic multimedia text centre had two principal goals. On the one hand, it aimed at methodological improvements in preparation of digital scholarly editions, where the method has been validated by the publication of two complex scholarly editions, publicly available on the internet. From the technological point of view the project goal was to produce digital tools which will facilitate the transformation of scholarly texts into highly structured TEI-XML files – a foundation for modern academic digital publishing.

The project research in Slovenian literary studies, history, ethnology and linguistics resulted in two digital monographs. The first one is a critical edition of Slovenian oaths and other official texts, produced by municipality authorities of Slovenian towns from 16<sup>th</sup> to early 19<sup>th</sup> centuries (*Mestna prisežna besedila v slovenskem jeziku do začetka 19. stoletja* by Boris Golec, ed. by Matija Ogrin, cf. <http://nl.ijs.si/e-zrc/prisege/>). In juridical and official administration, Slovenian language was rarely used in this period. For this reason the documents published in this critical edition are precious proofs, indicating the role and social status of Slovenian language in official structures. Most of the texts are published for the first time. In our digital edition, all the preserved texts are presented in three views: as a digital facsimile, as a diplomatically transcribed text and as a critically edited text (normalised with elements of modern orthography). The other digital monograph is a case study in Slovenian ethnology (*Folklorne pripovedi iz Bavšice* by Barbara Ivančič Kutin, ed. by Matija Ogrin, cf. <http://nl.ijs.si/e-zrc/pripovedi/>). Here, the issue was to publish digital audio recordings of folklore tales from Bavšica (district of Bovec) together with precise phonetic transcriptions on the one side and with normalised versions of the tales on the other – along with methodological and linguistic studies, pertaining to this type of folklore traditions. The third edition that will be published is a collection of the oldest surveys about Slovenian literature, ranging from Trubar (16<sup>th</sup> cent.) and Hren to Valvasor and Dolničar (18<sup>th</sup> cent.). The two monographs, published until now, embrace a wide range of methodological approaches to the textual materials: from diplomatics and palaeography to historical interpretations, from ethnological recordings to linguistic analysis of dialect, special lexica, etc. To all these methods and procedures the digital edition offered a suitable set of textual structures. We used Text Encoding Initiative (TEI) compliant XML mark-up, which offered excellent possibilities for all forms of textual research, analysis and presentation in our project.

The technological goal of the project was not only to enable the publication of these digital editions, but to produce a tool to facilitate the conversion of academic texts into TEI compliant XML and to promote it among the researchers in the Humanities and Social sciences. To this end, we

implemented a converter for text files in RTF (Word) format into TEI-XML format. The converter is publicly available on the web within the eZISS portal (<http://nl.ijs.si/e-zrc/rtf2tei/>). Besides, a repository and tracking system (based on Fedora Commons open-source software) was built behind this portal to enable storage and full-text searching in our editions. In this way, the project did not conclude only as a finished period of research, but even more as a starting point for further Humanities research and digital scholarly publishing in Slovenia.

## Povzetek

### Elektronsko besedilno središče z multimedijško komunikacijo

Namen večdisciplinarnega projekta *Elektronsko besedilno središče* je bil – z ene strani – izpopolniti metodologijo priprave elektronskih znanstvenih edicij na humanističnem področju in to metodo tudi praktično uresničiti z dvema izgotovljenima, kompleksnima znanstvenima izdajama. S tehnološke strani je bil cilj projekta izdelati elektronska orodja, ki bodo razmeroma težavno tehnološko pot do bogato strukturiranega zapisa XML–TEI kot temelja modernih elektronskih znanstvenih izdaj poenostavila in izdelavo takšnih izdaj približala širšemu krogu humanističnih znanstvenikov.

Literarnozgodovinski, historični in narodopisni cilji so bili doseženi z raziskavami, iz katerih sta nastali v projektu dve znanstveni monografiji: *Mestna prisežna besedila v slovenskem jeziku do začetka 19. stoletja* (avtor Boris Golec, ur. M. Ogrin, gl. <http://nl.ijs.si/e-zrc/prisege/>) in *Folklorne pripovedi iz Bavšice* (avtorica Barbara Ivančič Kutin, ur. M. Ogrin, gl. <http://nl.ijs.si/e-zrc/pripovedi/>). V prvi elektronski izdaji je avtor zbral vse razpoložljive oz. ohranjene uradovalne dokumente v slovenskem jeziku mestnega izvora od 16. do začetka 19. stoletja, kar je prvi tovrsten poskus doslej. Monografija uporablja celotno historiografsko metodologijo, dopolnjeno tudi s filološko metodo: od iskanja ter analize virov (rokopisov), prek dveh metodološko različnih prepisov (diplomatični, kritični prepis) do ocene historičnega pomena teh listin v originalnem socialnem okolju. Metodološko sega torej Golčeva monografija o mestnih prisegah od diplomatike in paleografije do intepretacije; s sodelovanjem jezikoslovcev je pridobila tudi pomembno historično-jezikoslovno vrednost (kritični prepis, analiza pravopisnih posebnosti, slovar historičnega besedja). Vsem tem metodičnim postopkom je koncept elektronske izdaje v zapisu TEI odgovoril z ustreznimi strukturami in jih praktično uresničil. Z drugo elektronsko monografijo – *Folklorne pripovedi iz Bavšice* – je realizirana kompleksna metodologija narodopisnega raziskovanja, tesno povezanega s slovensko dialektologijo: sega od teoretične razčlembe odnosov med živim pripovedovanjem, posnetkom in zapisom govora prek jezikoslovnega (dialektološkega) opisa govora do samega terenskega snemanja folklornih pripovedi (Bavšica nad Bovcem) in njihove večmedialne predstavitev (digitalizirani zvočni posnetki, fonetični prepis, knjižno-jezikovni zapis). Tako elektronska izdaja predstavi folklorne pripovedi v treh metodološko različnih modusih (posnetek in dva prepisa), hkrati pa v največji možni meri posreduje njihovo izvorno fakturo. Tudi tem metodološkim potrebam je elektronska izdaja lahko odgovorila z vsemi potrebnimi možnostmi. Tretja izdaja, ki predstavlja presežek programa dela – najstarejši pregledi slovenskega slovstva (Trubar, Hren, Valvasor, Dolničar) – je v uredniški proceduri in bo izšla jeseni 2009.

Tehnološka komponenta projekta Elektronsko besedilno središče nudi tem izdajam več kot le trdno funkcionalno podlago. Deluje marveč tudi kot servis, ki z javnim pretvornikom iz zapisa RTF v



XML-TEI (<http://nl.ijs.si/e-zrc/rtf2tei/>) olajša pripravo kompleksnih elektronskih znanstvenih edicij širšemu krogu humanističnih in družboslovnih raziskovalcev, s tem pa nudi kvalitetno javno (in brezplačno) storitev. Enako pomembna tehnološka sestavina je repozitorij besedil in poizvedovalni sistem Fedora Commons (nameščen za portalom eZISS – <http://nl.ijs.si/e-zrc/>), ki omogoča hitre in kvalitetne poizvedbe po obsežnih podatkovnih ali besedilnih zbirkah.

Z obema tehnološkima sistemoma se portal eZISS iz manjše zbirke znanstvenih izdaj spreminja v obsežnejši, kompakten sistem, namenjen shranjevanju raznolikih elektronskih besedil ter inteligentnemu iskanju po tekstih kot tudi njihovi uspešni javni prezentaciji ter diseminaciji.