



INSTITUT ZA HRVATSKI JEZIK I JEZIKOSLOVLJE

Institute of Croatian Language and Linguistics

Ulica Republike Austrije 16, HR-10000 Zagreb, tel.: ++385 1 3783 833, faks: ++385 1 3783 803,
e-pošta: ured@ihjj.hr, www.ihjj.hr, IBAN: HR6923900011100012967, OIB: 12268324202

URBROJ: 64-01-14/3-1758-2

Zagreb, 4. ožujka 2014.

Agencija za znanost i visoko obrazovanje
Donje Svetice 38/5
HR-10000 Zagreb

PREDMET: Očitovanje Instituta za hrvatski jezik i jezikoslovlje na Reakreditaciju

Poštovani,

u prilogu Vam dostavljamo Očitovanje Instituta za hrvatski jezik i jezikoslovlje na Reakreditaciju Instituta, *završno izvješće*, od 10. veljače 2014., KLASA: 640-02/13-02/0004, URBROJ: 355-06-02-14-6, zaprimljeno 21. veljače 2014.

Očitovanje Instituta za hrvatski jezik i jezikoslovlje o reakreditaciji dostavljamo u traženom roku.

U Očitovanju smo naveli niz činjeničnih podataka koje Stručno povjerenstvo ili nije uzelo u obzir ili ih je pogrešno interpretiralo, a držimo ih iznimno važnima kako bi ocjena sadašnje uprave Instituta pa i Instituta u cjelini bila objektivnija i povoljnija.

S poštovanjem,

ravnatelj Instituta za hrvatski jezik i jezikoslovlje

dr. sc. Željko Jozić



REPUBLIKA HRVATSKA
355 - AGENCIJA ZA ZNANOST I VISOKO
OBRAZOVANJE

Primljeno:	6.3.2014. 15:59:31		
Klasifikacijska oznaka	Org. jed.		
640-02/13-02/0004	02-05		
Urudžbeni broj:	Pril.	Vrij.	
380-14-0007	0	0	





INSTITUT ZA HRVATSKI JEZIK I JEZIKOSLOVLJE

Institute of Croatian Language and Linguistics

Ulica Republike Austrije 16, HR-10000 Zagreb, tel.: ++385 1 3783 833, faks: ++385 1 3783 803,
e-pošta: ured@ihjj.hr, www.ihjj.hr, IBAN: HR6923900011100012967, OIB: 12268324202

Očitovanje Instituta za hrvatski jezik i jezikoslovlje o Završnome izvješću Stručnoga povjerenstva za reakreditaciju Instituta za hrvatski jezik i jezikoslovlje

Zahvaljujemo Stručnomu povjerenstvu na Izvješću te osobito na preporukama za poboljšanje kvalitete rada naše institucije, koje ćemo svakako nastojati što prije provesti u djelo. No, najsrdanije molimo Stručno povjerenstvo da uzme u obzir i neke dodatne podatke koji bi mogli bitno utjecati na ocjenu Instituta, a osobito na ocjenu sadašnje uprave Instituta.

Naime, velik dio podataka u dokumentu samoanalize, a osobito podatci o znanstvenoj produktivnosti, međunarodnoj suradnji, te što je najvažnije – financiranju, odnosili su se na razdoblje od početka 2008. do kraja 2012., jer je tako bilo predviđeno obrascem. U tom razdoblju sadašnja uprava nije upravljala Institutom. S obzirom na to da se u Izvješću donosi i ocjena sadašnje uprave Instituta, a na temelju pokazatelja za koje uglavnom nije ni zaslužna ni odgovorna, molimo da se pri ocjeni uzmu u obzir i podatci iz 2013. godine, za koje je jedine sadašnja uprava odgovorna i zaslužna.

Ovdje ćemo se osvrnuti na nekoliko navoda iz Izvješća stručnoga povjerenstva, za koje molimo da se revidiraju u skladu s podacima koje navodimo.

„Ograničen odjek spomenutih djelatnosti vidljiv je i iz proračuna Instituta. Manje od jedan posto prihoda dolazi iz „drugih izvora“ (koji uključuju prodaju knjiga), iako se danas uobičajilo od drugih europskih znanstvenih instituta usporedive veličine i misije zahtijevati da sami zarade od 15 do 20 % vlastitog proračuna.“

Manje od jedan posto prihoda iz drugih izvora odnosi se na razdoblje koje završava s 2012. godinom, kada je na čelu Instituta bila druga uprava. U 2013. godini taj se postotak **znatno uvećao**, te za 2013. **iznosi više od 10 %**, a to je ostvareno samo u posljednja dva mjeseca 2013. godine, za što je izravno zaslužna nova uprava Instituta. Pokretanjem internetske prodaje institutskih izdanja (knjige.ihjj.hr), koju je također pokrenuo sadašnji ravnatelj, prodaja svih ostalih institutskih izdanja iznimno se povećala.

„Nadalje, IHJJ svojim glavnim zadatkom ne smatra zaradu od komercijalnih ugovora s industrijom, nego osiguravanje dostupnosti svojih izdanja javnosti. Jasan je pokazatelj toga činjenica da je Hrvatski pravopis slobodno dostupan na mrežnim stranicama, umjesto da se, kao što je uobičajeno, prvo tiskao (s elektroničkom inačicom ili bez nje) i tek nakon nekog vremena postao slobodno dostupan.“

Ravnateljeva je ideja bila objaviti ponajprije Hrvatski pravopis na internetu pa ga tek nakon toga tiskati, što je suprotno od uobičajene procedure, koja je ravnatelju itekako bila poznata, ali je zbog specifičnosti hrvatskih prilika (u kojima su na tržištu bila već četiri različita hrvatska pravopisa koja su se borila za dominaciju i prevlast) odlučeno krenuti obrnutim smjerom. Ta se odluka pokazala iznimno dobrom i uspješnom. Besplatnost je pravopisu dala mnogo medijskoga prostora u obliku intervjua i besplatnih reklama te mnogo poklonika, a time je popularnost *Hrvatskoga pravopisa* Instituta za hrvatski jezik i jezikoslovlje bila golema i prije njegova tiska. Takva je reklama zaslužna



INSTITUT ZA HRVATSKI JEZIK I JEZIKOSLOVLJE

Institute of Croatian Language and Linguistics

Ulica Republike Austrije 16, HR-10000 Zagreb, tel.: ++385 1 3783 833, faks: ++385 1 3783 803,
e-pošta: ured@ihjj.hr, www.ihjj.hr, IBAN: HR6923900011100012967, OIB: 12268324202

što se, jednom kad je tiskan, pravopis rasprodao u rekordnom roku i donio Institutu znatnu financijsku dobit te znatno pridonio društvenoj ulozi Instituta. Nije nevažno spomenuti da je tiskano izdanje Hrvatskoga pravopisa, unatoč činjenici da je na tržištu bio samo u posljednja dva mjeseca 2013. godine, najprodavanija knjiga u Republici Hrvatskoj na svim sajmovima te godine, ali i uopće: riječ je o najprodavanijoj knjizi u Hrvatskoj za 2013. godinu, a prodaja ne jenjava ni 2014. godine. Njegova je raširenost, a koja se postigla upravo njegovom besplatnom dostupnošću, bila i temelj proglašavanju toga pravopisa službenim, što je dakako, također djelovalo na prodaju knjige.

„Čini se da je ravnatelj prilično pasivan i nedovoljno proaktivan kad je riječ o uspostavi međunarodne mreže, privlačenju financiranja i suradnje na području Europske unije, suradnji unutar zemlje ili rješavanju problema uzrokovanih manjkom osoblja stručnog u računalnim tehnologijama.“

U 2013. godini Institut je postao članom dvaju projekata Europske unije: ravnatelj je postao članom Upravljačkoga vijeća COST akcije *TextLink: Structuring Discourse in Multilingual Europe*, a pomoćnica ravnatelja članica je upravljačkoga vijeća COST akcije *European Network on e-Lexicography*.

U 2013. godini Institut je prijavio **tri** projekta za financiranje Hrvatske zaklade za znanost (od kojih su **sva** tri ocijenjena u prvome krugu najvišim ocjenama) i jedan projekt za FP7, a u 2014. prijavit će se još četiri projekta za financiranje. Uprava je prošla niz seminara i predavanja o fondovima Europske unije, o Obzoru 20/20 i o pisanju uspješnih projekata. O tome je uprava održala i nekoliko predavanja za zaposlenike Instituta te time, a i vlastitim primjerom, iznimno poticala zaposlenike da se prijavljuju na natječaje i u tome im pomagala. Prijava tolikoga broja projekata velik je iskorak za ovu instituciju.

Uprava je iznimno proaktivna i kad je u pitanju suradnja s drugim institucijama unutar zemlje (v. sljedeću točku).

Osobito se aktivnost sadašnjega ravnatelja očituje u popularizaciji znanosti (u posljednjih godinu dana dao je više od 150 intervjua i imao više od 50 gostovanja u relevantnim kulturnim i znanstvenim emisijama) te je na temelju toga ostvarena i suradnja s izdavačima na polju iskorištavanja resursa koje Institut ima, a osobito za obrazovne svrhe, pa je posebice intenzivirana suradnja s Učiteljskim fakultetom u Zagrebu, nakon čega je pokrenuta izrada metodičkih priručnika za hrvatski jezik, koji će također povećati institutske prihode i učvrstiti vodeću institutsku ulogu u hrvatskome jezikoslovlju.

Manjak osoblja stručnoga u računalnim tehnologijama uglavnom je riješen zahvaljujući pokretanju suradnje s institucijama koje takve stručnjake imaju te povećanjem vlastitih prihoda koji su omogućili ugovorno zapošljavanje računalnih lingvisti.

„Konačno, čini se da Institut smatra nepotrebnim dijeliti vlastite resurse s drugim hrvatskim institutima i sveučilištima. Povjerenstvo žali što Institut, čini se, ne shvaća ozbiljno prednosti suradnje, bilo da je riječ o internacionalizaciji ili o predstavljanju vlastitih resursa modernoga ili povijesnoga hrvatskog jezika. Stoga nema ni planova za suradnju, povezivanje ili čak objedinjavanje korpusa raznih ustanova, iako je IHJJ u idealnom položaju da pokrene takvo objedinjavanje resursa.“



INSTITUT ZA HRVATSKI JEZIK I JEZIKOSLOVLJE

Institute of Croatian Language and Linguistics

Ulica Republike Austrije 16, HR-10000 Zagreb, tel.: ++385 1 3783 833, faks: ++385 1 3783 803,
e-pošta: ured@ihjj.hr, www.ihjj.hr, IBAN: HR6923900011100012967, OIB: 12268324202

U 2013. godini Institut je pokrenuo nacionalno važan projekt CroLLOD – povezani hrvatski jezikoslovni podatci, koji je projekt suradnje između Instituta za hrvatski jezik i jezikoslovlje, Fakulteta strojarstva i brodogradnje, Filozofskoga fakulteta u Rijeci, Filozofskoga fakulteta u Osijeku, i Matematičkoga fakulteta u Osijeku i CARNeta. U projekt će se uključiti i nastavnici hrvatskih osnovnih i srednjih škola. Taj je projekt upravo zamišljen kao projekt objedinjavanja korpusa i drugih jezičnih resursa raznih ustanova, a takvo je objedinjavanje inicirao upravo Institut. Prilažemo i kratak opis projekta:

Temelj ovoga računalnoga pristupa hrvatskomu jezikoslovlju nalazi se u tekstu/korpusu. Iz njega se dohvaća i(ili) provjerava morfologija, sintaksa i semantika. S računalne strane radi se o mrežnom sustavu za hrvatski jezik i jezikoslovlje s mogućnošću proširenja na širu kulturnu baštinu, a sastoji se od sljedećih modula/računalnih cjelina:

A) KORPUS [Modul služi za spremanje i dohvaćanje dokumenata (tekstova, digitaliziranih rukopisa/slika, tonskih zapisa, filmova, tj. svih vrsta digitaliziranih podataka) s preciznim postavljanjem u prostor (*Google-map*) i vrijeme (*timeline*), te razvrstavanjem po autorstvu, nazivu, kategorijama, mediju i sl., s jezikoslovnim alatima koji uključuju lematizaciju, konkordanciju, tvorbu rječnika i sl. U izgradnji je također praćenje povijesti (nastanka, života i nestajanja) pojedine hrvatske riječi ili skupa hrvatskih riječi u korpusu. Odabrane dokumente korpusa moguće je vezati s modulom za oblikoslovlje C) ili CroLLOD sustav G).]

B) OZNAČIVANJE dokumenata za djelotvornu računalnu obradbu [Za označivanje dokumenata razvili smo dva programa *DocMark* i *TEIMark*. *DocMark* Prvi služi za označivanje netranskribiranih dokumenata, npr. stari rukopisi, kojim se vizualne oznake postavljaju nad slikom dokumenta, u slojevima (pri čemu se original ne oštećuje, niti mijenja). Program *TEIMark* služi za vizualno označivanje digitaliziranog/unesenog teksta jednog ili više dokumenata, s osobnim ili TEI (*Text Encoding Initiative*) oznakama. I u ovom slučaju moguće je suradničko označivanje dokumenta, rad u nezavisnim slojevima i njihova obradba. Modul za označivanje može se direktno povezati s modulima D), E), F) i G).

C) MORFOLOGIJA s uključenim glasovnim promjenama [Program ima mogućnost predlaganja rješenja koje korisnik izabire za dotad nepoznatu riječ (ako ne postoji u bazi riječi) ili izborom glasovnih promjena (kojim stručnjak) stvara jednoznačno rješenje. Program radi nad bazom koja ima oko milijun označenih/tagiranih riječi (uključujući sve gramatičke oblike riječi). Program je na testiranju za sve vrste promjenljivih riječi.]

D) SINTAKSA rečenice po rečeničnim dijelovima [Sintaksa rečenice ostvarena je trima modulima: Poluautomatiziranoj zamjeni (**PAZ**) riječi s pripadnim oznakama (eng. *PoS – part of speech*); Univerzalnom rečeničnom grafu (**URG**) – pokušaj stvaranja svih gramatičkih struktura; Generaliziranom dohvaćanju strukturne i pojmovne informacije – s pomoću regularnih izraza.]

E) DOHVAĆANJE informacije/smisla iz konteksta dokumenata [Dohvaćanje ili 'čupanje' informacije/smisla temelji se na algoritmu koji objedinjuje znakovni i gramatički pogled/filtar, uz generaliziranu konkordanciju konteksta. Program se jednako dobro može koristiti u sintaktičkoj kao i semantičkoj primjeni. Radi se samo o vještini dobro i stručno promišljene definicije uzoraka (apstraktnih riječi) i njihovih kombinacija/izraza (temeljenih na regularnim izrazima/eng. *regular expressions*). Na semantičkoj razini program se lako proširuje izrazima za dohvaćanje n-grama, kolokacija, fraza i metafora...Dohvaćanje (nestrukturirane) informacije u tekstu postignuto je izgradnjom Python programa prilagođenih za hrvatske potrebe i različitosti]

F) TVORBA temeljnih informacijskih trojaca [Problemi goleme količine mrežne informacije i dohvaćanje neželjene ili netočne informacije klasičnih tražilica (*Google, Yahoo*) usmjerili su suvremeni razvoj interneta na tvorbu povezanih podataka (eng. *linked data*) svih, pa i digitalnih jezičnih tehnologija. Povezivanjem trojaca osnovne informacije



INSTITUT ZA HRVATSKI JEZIK I JEZIKOSLOVLJE

Institute of Croatian Language and Linguistics

Ulica Republike Austrije 16, HR-10000 Zagreb, tel.: ++385 1 3783 833, faks: ++385 1 3783 803,
e-pošta: ured@ihjj.hr, www.ihjj.hr, IBAN: HR6923900011100012967, OIB: 12268324202

(subjekt- predikat-objekt) stvaraju se grafovi znanja, koji se mogu lakše, brže i preciznije dohvatiti.]

G) CroLLOD ontologije - sustav povezanih jezikoslovnih podataka [Načinjeni su i razvijaju se svi nužni preduvjeti da se u globalni jezikoslovni oblak ugrade i hrvatski lingvistički povezani podaci (**LLOD** – *Linguistic Linked Open Data*) kako su to učinili mnogi drugi narodi sa svojim jezicima.

Svi moduli zamišljeni su tako da se međusobno nadopunjuju i isprepliću u mrežnom sustavu kako bi bili što korisniji u istraživačkom radu i školskoj/akademskoj poduci. Tako na primjer, ako se u online skladnji rečenice pojavi neka riječ koja nije do tada bila morfološki obrađena, program otvara oblikoslovni modul u kojem se to načini. Jednom načinjen i provjeren uradak više se ne treba ponavljati, jer se sve nalazi i događa na istom mjestu - Internetu (unutar hrvatske akademske i istraživačke mreže *CARNet*). Sustav nije samo zamišljen, on je većim dijelom (65 %) do danas i načinjen, a do konca 2014. nadamo se završiti ga (u trenutačno zamišljenu opsegu A – E).

Držimo da su ovi podatci iznimno važni i da bi trebali znatno utjecati na ocjenu sadašnje uprave Instituta. Najsrdacnije molimo Stručno povjerenstvo da ih uzme u obzir.

S osobitim poštovanjem,

ravnatelj Instituta za hrvatski jezik i jezikoslovlje

dr. sc. Željko Jozić

