

GNUSL: PROSTO PROGRAMJE IN SLOVENŠČINA

Aleš Košir,¹ Primož Peterlin,² Tomaž Erjavec³

(1) Hermes SoftLab,
Litijška 51, 1000 Ljubljana,
ales.kosir@hermes.si

(2) Inštitut za biofiziko MF, Univerza v Ljubljani
Lipičeva 2, 1000 Ljubljana
primoz.peterlin@biofiz.mf.uni-lj.si

(3) Odsek za inteligentne sisteme, Institut Jožef Stefan
Jamova 39, 1000 Ljubljana
tomaz.erjavec@ijs.si

POVZETEK

GNUsl na <http://nl.ijs.si/GNUsl/> je ime za zbirko po internetu prosto dostopnih virov in programov, ki izkoriščajo znanje o slovenskem jeziku, državno uzakonjeni standardizaciji in kulturnih značilnostih, s čimer omogočajo rabo ali olajšujejo uporabo računalnikov v slovenskem kulturnem prostoru.

Zbirko programov spremljajo navodila, ki vsebujejo nasvete s področja internacionalizacije in lokalizacije. Nasveti s primeri se praviloma nanašajo na prosto programje in na bolj razširjene komercialno dostopne izdelke, pri čemer so rešitve sistematično ponujene predvsem za različne operacijske sisteme in za bolj razširjene urejevalnike besedil. Poleg tega navodila vsebujejo tudi informacije, specifične za slovenski kulturni prostor, na primer o pravilnem pisanju števil in decimalno vejico in o tem, kako so predpisani prazniki v Republiki Sloveniji. Navodila so pripravljena tako v obliki dokumentov, dostopnih na internetu, in v obliki elektronskih brošur, primernih za natis.

V ponudbi GNUsl je tudi Trubar, internetni servis za preverjanje pravilnosti črkovanja v mnogih jezikih, med katerimi je tudi slovenski. Uporabljeni črkovalnik GNU *ispell* je v duhu GNUsl prosto dostopen z izvorno kodo in s seznama uporabljenih besednih oblik vred.

Predstavljeno gradivo na GNUsl skuša sistematično zajeti prosto dostopno znanje s področja jezikovnih tehnologij in ga predstaviti na način, ki bo vsakomur razumljiv. Navodila so opremljena z značilnimi primeri.

ABSTRACT

GNUsl, hosted on <http://nl.ijs.si/GNUsl/>, is the name for a collection of free documentation and software that uses information about the Slovene language, Slovene standards, and cultural specifics and helps people to use computers in the Slovene cultural environment.

The software collection is accompanied by guides, containing instructions on internationalization and localization. Instructions are focused on free software and

on popular commercial software. Solutions are systematically provided for various operating systems and for prevalent text editors. Guidelines include notes that are specific for the Slovene cultural environment, such as the rule for comma as decimal separator and legislation on holidays in Slovenia. Instructions are available through Internet as hyperlinked active documents or in ready-to-print form.

GNUsl includes Trubar, an Internet spelling checker supporting many languages, including Slovene. Trubar is built on top of GNU *ispell*, a free spelling checker for which both program sources and dictionaries are freely available.

Materials presented on GNUsl compile most freely available knowledge from the field of language technologies for the Slovene language. They are presented in the form of understandable instructions, which are illustrated with examples.

1 UVOD

Računalniška tehnologija je tradicionalno zelo angleško naravnana in do pred kratkim še ni sistematično upoštevala lokalnih značilnosti v kulturnih okoljih, v katerih se je uporabljala. V minulih treh letih je opaziti spremembe, ki kažejo, da se tudi pri nas skuša to pomanjkljivost vse bolj načrtno popraviti. Veliki proizvajalci programske opreme zaradi pritiska uporabnikov včasih popustijo in se sprijaznijo z mednarodnimi standardi, ki so se razvili iz potrebe po smotnejšem in bolj metodičnem razvoju programske opreme za medsebojno silno različna kulturna okolja, četudi se največji in najmočnejši med njimi namesto tega oklepajo lastnih industrijskih standardov in tako varujejo svoja vlaganja v razvoj programske opreme. Ozadje te težnje je seveda zniževanje stroškov razvoja programske opreme, ki mora v različnih okoljih zadoščati zelo različnim zahtevam. Žal takšno vsiljevanje povzroča uporabnikom težave, brž ko se srečajo različne uporabljene tehnologije, kar gotovo ne olajšuje dela z računalniškimi orodji in komunikacije med uporabniki.

V tem navzkrižju interesov imajo uporabniki pravico zahtevati uporabo mednarodnih standardov. Primer velike raznolikosti je kar najosnovnejše opravilo, ki ga mora izvajati urejevalnik besedil. Univerzalni urejevalnik mora vsaj upoštevati, da odvisno od okolja ljudje pišejo v mnogih pisavah, v različnih smereh, pri čemer uporabljajo specifične načine vnosa znakov.

Takšna prizadevanja po prilagojevanju proste programske opreme so tudi v Sloveniji pokazala nekaj rezultatov [1-2]. Da bi jih zbrali na enem mestu in omogočili prost dostop do njih, je skupina za raziskovanje naravnih jezikov na Odseku za inteligentne sisteme na Institutu Jožef Stefan omogočila uporabo strežnika nl.ijs.si s programsko opremo Apache. Na njem je izključno s prostovoljnimi delom nastala domača stran z imenom GNUsl.

2 GNU

GNUsl skuša v duhu GNU ponuditi tiste tehnologije, ki so prosto dostopne in vezane na slovenski kulturni prostor. GNU je ozko povezan z idejo prostih programov, zato najprej pojasnimo ta pojem [6]. Uporaba prostih programov ne terja plačila. Vendar pa tudi Microsoftov Internet Explorer lahko uporabljamo, ne da bi bilo potrebno za to plačati. Microsoft ponuja neokrnjeno različico programa na svoji spletni strani in vsak, ki to želi, jo lahko presname na svoj disk in uporablja. Pa Internet Explorer zato vseeno ni prost program. Uporabnik dobi namreč samo prevedeno kodo, izvorne pa ne. Zato je ne glede na to, ali to zna ali ne, prikrajšan, da bi odpravil napake, dodal kaj novega k programu, ali ga priredil tako, da bi tekel na drugem operacijskem sistemu, in prikrajšan za to, da bi tako izboljšano izvedbo programa delil z ostalimi. Še več, pravzaprav niti prevedene kode ne sme deliti, saj mu kot končnemu uporabniku pogodba dovoljuje le, da ga shrani na disk in uporablja. Prostost predvsem ni stvar cene. Nič nenavadnega ni, da moramo za CD-ROM s prostimi programi plačati. Del plačila bo pri tem ostal v trgovini, del ga založnik porabi za kritje svojih stroškov, del ga običajno nameni podpori avtorjem programja na plošči. Dostopnost izvorne kode je nujen predpogoj, da je program prost, sama po sebi pa še ne zagotavlja prostosti programov. Na kratko: prosti programi uporabniku dovoljujejo uporabo, razmnoževanje, razširjanje, razumevanje, spreminjanje in izboljševanje programa. Natančneje svoboščine razdelimo na tri ravni:

1. svoboda branja izvorne kode, razumevanja delovanja programa in po potrebi prilagojevanja lastnim zahtevam;
2. svoboda razširjanja kopij programa;
3. svoboda izboljševanja programa, vključno s svobodo razširjanja svojih sprememb in dopolnitev.

Vodilo pri projektu GNU je bilo zagotoviti, da čisto vsi uporabniki prejmejo programje pod enakimi pogoji. Koncept se imenuje Copyleft in je besedna igra z

angleškim izrazom za zaščito avtorskih pravic Copyright. Idejo strnemo z besedami: vsakdo, ki razširja programje v spremenjeni ali nespremenjeni obliki, mora z njim vred razširjati tudi vse pravice in svoboščine, ki jih je sam prejel, vključno s pravico do nadaljnega razširjanja in spreminjanja.

Copyleft je splošni koncept. Njegova izvedba v obliki dovoljenja za razširjanje in razmnoževanje je GNU General Public License, ali na kratko GPL, katere izvod je priložen vsakemu programu iz projekta GNU.

Sistem copyleft ima daljnosežne posledice. Zagotavlja namreč, da bo program, izdan pod temi pogoji, vselej ostal dostopen pod enakimi pogoji. Konkretno, če ste na primer v GNU Emacs dodali možnost vključevanja slik v besedilo, tako spremenjenega programa ne morete prodajati, niti ne morete ponujati samo prevedene kode. Lahko ga le ponudite naprej pod enakimi pogoji, kot ste ga sami dobili, z izvorno kodo vred. Rezultat je seveda natančno to, za kar si prizadeva projekt GNU: širjenje baze prostega programja. Nekateri avtorji svoje sicer prosto programje izdajajo pod ohlapnejšimi pogoji, ki dopuščajo tudi komercialno izrabo.

Med prosto programje so uvrščeni programi v javni lasti, programi, izdani pod pogoji copylefta, med katerimi so vsi pod pogoji GNU GPL izdani programi, in prosti programi, izdani pod drugačnimi pogoji. Skupina programov, ki ne ustreza prej zapisanim trem pogojem za prosto programje, sodi v skupino zakonsko zaščitenega (angl. proprietary) programja. Mednje spada večina komercialnega programja ter programov za pokušino (angl. shareware). Slednji sicer navadno dovoljujejo prosto razširjanje, praviloma pa niso dostopni v izvorni kodi, in avtorji za njihovo uporabo zahtevajo plačilo.

3 GNUSL

Stran GNUsl na naslovu <http://nl.ijs.si/GNUSL/> je bila postavljena v letu 1996 in od takrat deluje neprekinjeno. Doslej so jo obiskali uporabniki interneta iz več kot 100 domen, povprečno je na dan prek 70 dostopov [7]. Od začetka leta 1997 do sredine leta 1998 je bilo zabeleženih skupno 75.000 dostopov.

Prevladujejo uporabniki iz domene .si, ki jim sledijo sorazmerno pogosto uporabniki iz pretežno ameriške domene .com in nato iz domene .net. Zelo visok je odstotek uporabnikov iz vzhodnoevropskih držav, predvsem tistih, kjer govorijo slovanske jezike in se srečujejo s podobnimi težavami kot Slovenci.

Tako na devetem mestu zasledimo domeno Estonije s prek 1300 dostopi in 2% količine prenešenih podatkov, na štirinajstem mestu Češko s prek 2500 dostopi in 1,4% prenešenih podatkov, na štiriindvajsetem mestu Jugoslavijo s 2400 dostopi in 0,4%. Hrvaška je na osemindvajsetem mestu z 1700 dostopi in 0,3%.

Doslej so na GNUsl našli svoje mesto črkovalnik Trubar, zbirka navodil "Slovenščina in računalniki" in besedilo o

slovenskih katalogih sporočil v okviru projekta GNU. Črkovalnikom je posvečena posebna pozornost, na GNUsl je dostopen tudi pregled zgodovine črkovalnikov za slovenski jezik. Z GNUsl se odpira pogled na zbirko sorodnih strani, od katerih je najbližja domača stran projekta MULTEXT-East, ki se je ukvarjal z večjezičnimi besedilnimi orodji in korpusi v jezikih srednje in vzhodne Evrope.

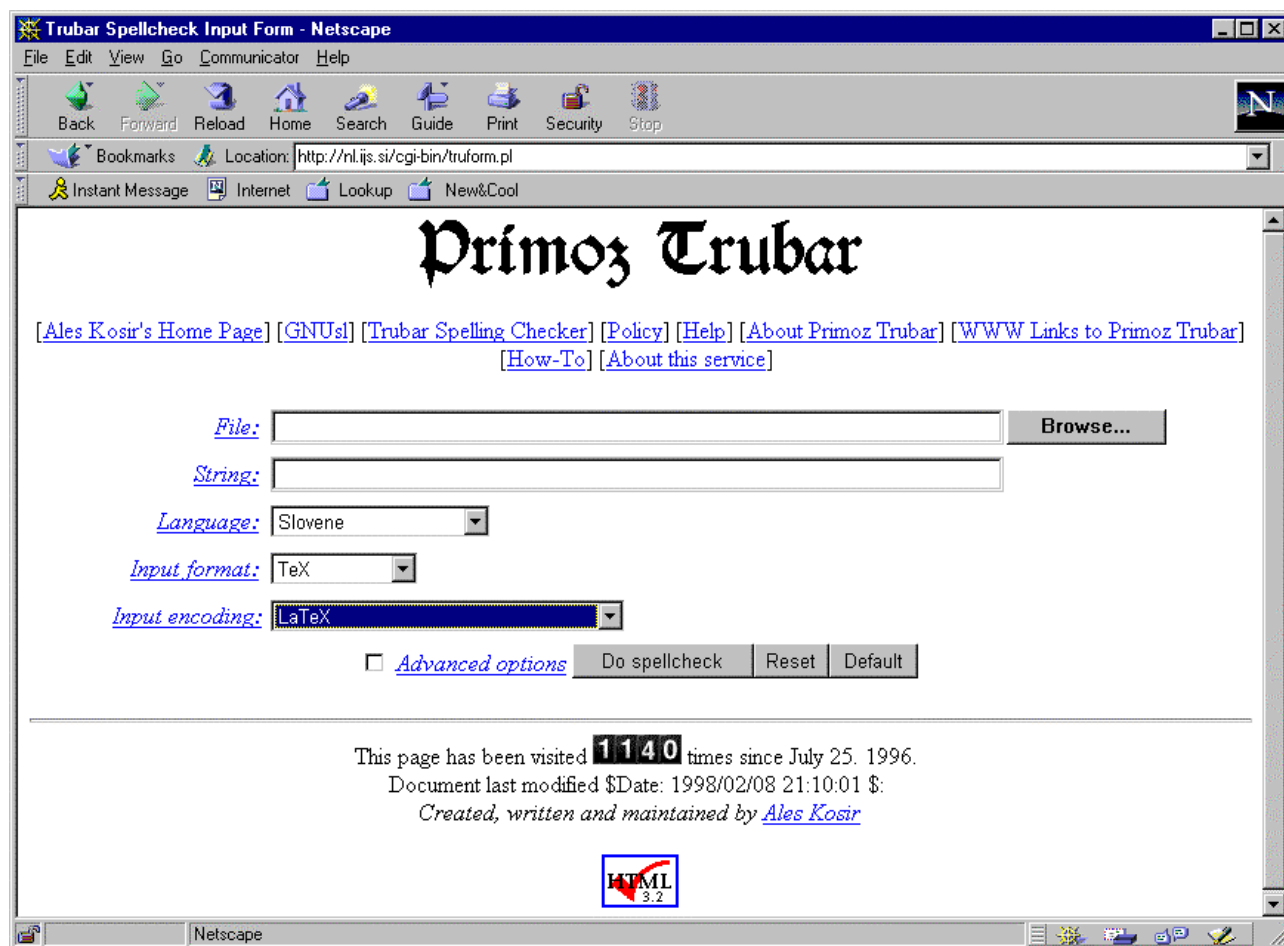
4 TRUBAR: INTERNETNI ČRKOVALNIŠKI SERVIS

Trubar je javno dostopni servis za preverjanje pravilnosti črkovanja [10]. Dostopen je na internetnem naslovu <http://nl.ijs.si/cgi-bin/truform.pl>. Uradno deluje neprekinjeno od 8. februarja 1996. Doslej ga je uporabilo nad tisoč uporabnikov. Trenutna strojna oprema, na kateri teče Trubar, žal ne dopušča množične uporabe tega servisa.

Trubar, prikazan na sliki 1, je izvorni grafični vmesnik, pod katerim teče program *ispell*, in je bil napisan za

potrebe GNUsl. *ispell* je zaradi preproste uporabe in primernosti za številne jezike razširjen črkovalnik, ki je dostopen pod licenco GNU. Zlahka ga uporabljamo tudi s slovenščino in omogoča nam, da imamo svoje lastne slovarje besed, pri čemer pa je poglobljena težava programa *ispell* v slovenščini prav pomanjkanje kvalitetnih zbirk besednih oblik. Doslej je bilo njihovo zbiranje omejeno na tiste kvalitetne elektronske vire, ki jim je zaupati, da vsebujejo le knjižno pravilne besedne oblike [3]. Taka zbirka besednih oblik je nastala v okviru projekta MULTEXT-East [5] in obsega približno 200.000 besednih oblik, kar je sicer majhen delež od približno 3.300.000 besednih oblik [4], generiranih iz korenov SSKJ. K tej se je v času delovanja pridružilo še približno 60.000 besednih oblik.

Za potrebe slovenskega jezika je bilo potrebno programu *ispell* navesti, kako so kodirane pisave, s katerimi so zapisani slovenski znaki, in katera priponska pravila naj uporablja, da baza z besednimi oblikami ne bi bila prevelika.



Slika 1: Uporabniški vmesnik prosto dostopnega internetnega servisa za preverjanje pravilnosti črkovanja.

Trubar podpira naslednje jezike:

- ameriško angleščino,
- britansko angleščino,
- nizozemščino,
- francoščino,
- nemščino,
- portugalsščino in
- slovenščino.

Da bi preverili pravilnost črkovanja besed v izbranem dokumentu, prek vmesnika, ki je predstavljen na sliki 1, navedemo njegovo ime in izberemo jezik, format datoteke in način kodiranja. Pri slovenskem jeziku imamo na voljo več baz besednih oblik: bazo z ročno preverjenimi knjižnimi besedami in bazo z nepreverjenimi besednimi oblikami. Če želimo preveriti pravilnost le ene besede, potem ni potrebno podati imena datoteke, pač pa jo zapišemo v polje String. Izpis vsebuje abecedni seznam neznanih besed in za vsako neznano seznam podobnih znanih besed. Neznane besede se na strežniku shranjujejo in so po ročnem pregledu dodane v osnovno bazo besed.

Zahtevnejši uporabniki lahko izbirajo dodatne možnosti, med njimi način urejenosti besed v izpisu, dolžino besed, ki so vselej pravilne, možnost, da so sestavljene besede vselej pravilne in možnost preproste statistične obdelave besedila.

5 SLOVENŠČINA IN RAČUNALNIKI

GNUsl ponuja tudi obsežno zbirko navodil, kako prilagodimo nekatera računalniška orodja slovenskemu jeziku. Navodila so pod naslovom Slovenščina in računalniki dostopna na slovenski spletni strani <http://nl.ijs.si/GNUsl/tex/tslovene/slolang/slolang.html> [9].

Uvodoma so naštetih standardizirani in nekateri specifični kodni nabori, ki se uporabljajo v povezavi s slovenskim jezikom. Svetuje se, da v medsebojnem pisnem sporazumevanju v elektronski obliki, kjerkoli je to le mogoče, uporabljamo kodni nabor ISO-8859-2, ki ob času pričujočega pisanja velja za priporočeno izbiro.

Kako na raznovrstnih operacijskih sistemih uporabljamo slovenske znake, je navedeno za večino razširjenih sodobnih sistemov: MS-DOS, Windows 3.x, Windows 95, Windows 95 OEM Service Release 2, Windows CE, Windows NT, IBM OS/2 Warp, Linux, HP-UX 9.05, HP-UX 10.01, SUN Solaris 2.x, SunOS 4.1.x, Apple Macintosh in VAX/VMS. Kjer je le mogoče, so nasveti ponazorjeni s primeri.

Širše je pozornost posvečena jezikovnim problemom, vezanim na urejevalnike besedil. Tu ponujena rešitev za vnos šumnikov s sestavljanjem znaka za dvojni narekovaj in sičnika za ustrezní šumnik je preskušena in

podrobno opisana za raznoliko programsko opremo. Rešitev je priporočena povsod tam, kjer uporabniki pri delu uporabljajo tipkovnico z neslovenskim razporedom. Kot alternativni predlog je ponujena rešitev za vnos šumnikov s tipkovnico, na kateri ni posebnih tipk za šumnike in te vnašamo s sočasnim pritiskom na tipko ALT in nato na tipko za ustrezní sičnik. Pri tem običajno na račun vnosa šumnikov izgubimo kakšno drugo funkcionalnost in je zato ta način vnosa šumnikov manj priljubljen.

Znaten del pojasnil je namenjen urejevalniku GNU emacs, ki je eden izmed osnovnih komponent projekta GNU. Ta urejevalnik je že prerasel ozke okvire osnovnega namena, za katerega je bil napisan, in vse bolj predstavlja integrirano okolje, v katerem uporabnik preživi večino časa pred zaslonom. Emacs je vsekakor sodoben sistem, ki omogoča preprosto uporabo različnih pisav in jezikovnih orodij, prevsem pa ga je mogoče poljubno prilagojevati, spreminjati in razširjati, kar cenijo raziskovalci s posebnimi zahtevami. Kako raznovrstne pisave hkrati podpira, je prikazano na sliki 2.

Sistem za oblikovanje besedil LaTeX je v Sloveniji od leta 1985 zelo razširjen in močno priljubljen in je v nekaterih nišah našel svoje stalno mesto. Z njim v zvezi spregovorimo o slovenskih delilnih vzorcih in o orodjih za avtomatsko indeksiranje besed. Nadalje se izčrpno obravnava podpora za preverjanje pravilnega črkovanja, kjer je opisano, kako deluje Trubar in kako si na osnovi prosto dostopnega paketa *ispell* po potrebi sami zgradimo črkovalnik po svoji meri ali pa uporabimo že vnaprej pripravljene baze besednih oblik.

Z nabori znakov je seveda tesno povezano vprašanje pretvorb med njimi. Prikazani in ovrednoteni so različni načini, kako pretvarjamo med nabori.

V duhu časa je dodano poglavje o internetu in o standardu HTML, ki se nadaljuje v poglavje o programih za izmenjevanje elektronskih sporočil. Prav tu opažamo, da se tako v uradnih elektronskih pismih državnih institucij kot v osebnih pismih in na elektronskih straneh, objavljenih bodisi na inter- ali intranetu, ne držimo priporočil in standardov, kar prav gotovo ne olajšuje in pospešuje uporabe teh tehnologij.

Nadalje so tu zbrani drobci informacij, za katere domnevamo, da bi koristile v tem kulturnem in geografskem območju. Na kratko so predstavljene nastavitve, s katerimi preklapljammo med nabori pisav na tukaj razširjenih tiskalnikih. Sledijo pravila, kako so v Sloveniji uzakonjeni prazniki in dela prosti dnevi ter kako program *gcal*, ponovno iz projekta GNU, pripravimo, da jih upošteva.



Slika 2: Urejevalnik emacs je sodoben izdelek, ki je v marsikaterem tehnološkem pogledu pred komercialnimi izdelki najbolj znanih proizvajalcev.

Zbirka navodil se zaključuje z informacijami o nekaterih geografskih točkah v Sloveniji, kar koristi, če želimo izračunavati od zemljepisnega položaja odvisne dogodke, na primer astronomske pojave, in pri satelitski sinhronizaciji računalnikov.

Ob koncu prispevka je še razprava o nekaterih s stališča uporabnikov najbolj ključnih potrebah in tehnologijah, ki bi olajšale delo z računalnikom. V stopnjujočem redu zahtevnosti si te želje po prosti dostopnosti sledijo:

- program, ki za poljubno slovensko besedo določi njeno osnovno obliko,
- program, ki pri podani lemi tvori vse njene morfološke oblike,
- orodje za avtomatsko indeksiranje slovenskih besedil,
- obsežen in verificiran morfološko-sintaktični slovar,
- obsežen in verificiran slovar vseh slovenskih besednih oblik,
- Slovar slovenskega knjižnega jezika (SSKJ) v elektronski obliki,
- tezaver slovenskega jezika v elektronski obliki,
- program za preverjanje skladišne pravilnosti,
- večjezične slovarje za slovenski jezik v elektronski obliki,
- program za razpoznavanje govora,
- slovensko govoreči vmesnik med računalnikom in človekom.

6 PROJEKT GNU IN SLOVENSKI KATALOGI SPOROČIL

V projektu GNU so začeli z internacionalizacijo in posledično lokalizacijo programja v letu 1995 [11]. V času pisanja tega prispevka je podprta lokalizacija v projektu GNU na naslednjih nivojih:

- Sistemska knjižnica GNU za programski jezik C, `glibc`, podpira lokalizacijo na način, določen s priporočilom POSIX.2 [12-13]. To zajema vseh šest kategorij lokalizacije: `LC_CTYPE`, ki določa izbrani kodirani nabor znakov, `LC_COLLATE`, ki določa abecedni red znakov v naboru, `LC_TIME`, ki določa izpis datuma in ure, `LC_NUMERIC` in `LC_MONETARY`, ki določata izpis zneskov, ter `LC_MESSAGES`, ki določa dialoge programa z uporabnikom.

- Zbirka orodij za pomoč pri prevajanju katalogov sporočil v programih, pisanih v programskem jeziku C, GNU `gettext`.
- Program za pretvorbo med različnimi kodiranimi nabori znakov GNU `recode`.

Prvih pet kategorij lokalizacije je odvisnih le od izbranega kulturnega okolja, v našem primeru slovenskega, in jih opravimo enkrat za vselej. Za primer slovenščine je bilo to delo že opravljeno v Laboratoriju za odprte sisteme in mreže na Institutu Jožef Stefan. Zadnja kategorija, `LC_MESSAGES`, pa določa način dialoga med računalnikom in uporabnikom. Program med delom izpisuje sporočila, npr. „Warning: All data on non-removable disk will be destroyed. Proceed? (Y/N)“ in pričakuje odgovor uporabnika, v tem primeru pritisk na tipko Y ali N. V programu je navadno od nekaj deset do nekaj sto tovrstnih sporočil, ki so za uporabo z mehanizmom `gettext` zbrana v takoimenovane kataloge sporočil za posamezni programski paket. Vsako od sporočil v katalogu je treba prevesti v vsakega od podprtih jezikov. Trenutno stanje podprtosti ponazarja tabela 1.

Slovenski prevodi so, enako kot prevodi v druge jezike, delo prostovoljnih skupin prevajalcev, ki jih koordinira François Pinard z Univerze v Montrealu. Stran GNUsl služi obveščanju in izmenjavi izkušenj med člani slovenske prevajalske skupine.

7 ZAKLJUČEK

GNUsl predstavlja poizkus, kako bi pregledali, združili in ocenili delo, ki je bilo doslej opravljeno na področju jezikovnih tehnologij za podporo slovenskemu jeziku in katerega rezultati so prosto dostopni, ter kako bi smotrno porazdelili napore, da se vloženo delo ne bi po nepotrebnem pomnoževalo. Velika odmevnost servisa kaže, da so obravnavana vprašanja zanimiva in da dosedanje ljubiteljsko delo ni bilo neopaženo.

8 ZAHVALA

Avtorji se zahvaljujejo za nesebično pomoč vsem, ki so s prostovoljnim delom pomagali GNUsl, da je oživel in da živi.

	cs	da	de	en	es	fi	fr	it	ja	ko	nl	no	pl	pt	ru	sl	sv	
bash																	3	
bison																	3	
clisp																	4	
cpio																	6	
diffutils																	5	
enscript																	6	
fileutils																	11	
findutils																	9	
flex																	4	
gcal																	5	
gettext																	12	
grep																	10	
hello																	12	
id-utils																	3	
indent																	5	
libc																	7	
m4																	6	
make																	6	
music																	2	
ptx																	8	
recode																	9	
sh-utils																	8	
sharutils																	6	
tar																	11	
texinfo																	4	
textutils																	9	
wdiff																	8	
wget																	3	
17 jezikov	cs	da	de	en	es	fi	fr	it	ja	ko	nl	no	pl	pt	ru	sl	sv	
28 paketov	7	4	26	1	18	1	26	3	1	12	21	10	19	7	5	7	17	185

Tabela 1: Tabela podprtosti za programske pakete iz projekta GNU. Navpično so izpisana imena programskih paketov, vodoravno pa okrajšave za jezike, skladno s standardom ISO 639:1988. Slovenščino tako najdemo v stolpcu, označenem s sl.

9 VIRI

- [1] A. Košir, P. Peterlin, „Slovenski črkovalniki in international Ispell“, 824. seminar v okviru seminarja za numerično in računalniško matematiko na Inštitutu za matematiko, fiziko in mehaniko, 12. feb. 1997.
- [2] A. Košir, „(La)TeX in slovenščina“, 824. seminar v okviru seminarja za numerično in računalniško matematiko na Inštitutu za matematiko, fiziko in mehaniko, 15. jan. 1997.
- [3] M. Tavzes (ur.), „Računalniška besedišča: material za okroglo mizo, Ljubljana, 14. 6. 1991“, Ljubljana, oktober 1991.
- [4] P. Jakopin, A. Bizjak, „O oblikoslovnem označevanju slovenskega besedila“, Slavistična revija, Vol. 45/3-4, 1997.
- [5] „East meets West: A Compendium of Multilingual Language Resources Vol. II“, spletna stran <http://nl.ijs.si/ME/CD/mte-home.html>
- [6] „The GNU General Public License“, spletna stran <http://www.fsf.org/copyleft/gpl.html>
- [7] T. Erjavec, A. Košir, P. Peterlin, „GNUsl“, spletna stran <http://nl.ijs.si/GNUsl/>
- [8] R. Stallman, „GNU“, spletna stran <http://www.fsf.org/>
- [9] A. Košir, „Slovenščina in računalniki“, spletna stran <http://nl.ijs.si/gnusl/tex/tslovene/slolang/>
- [10] A. Košir, „Trubar“, spletna stran <http://nl.ijs.si/cgi-bin/truform.pl>
- [11] U. Drepper, „Internationalization in the GNU project“, v „Proceedings of the First Conference on Freely Distributable Software, February 3-5, 1996“, Cambridge, Massachusetts, 1996.
- [12] ISO/IEC 9945-1:1996 Information technology - Portable Operating System Interface (POSIX) - Part 1: System Application Program Interface (API) [C Language]
- [13] ISO/IEC 9945-2:1993 Information technology - Portable Operating System Interface (POSIX) - Part 2: Shell and Utilities