

Slovenjenje proste programske opreme

*Andrej Vernekar,[†]Robert Ludvik,[‡]Andraž Tori

*Skupina za slovenjenje pri društvu LUGOS
LUGOS - Društvo uporabnikov Linuxa Slovenije
Predjamska 4, 1000 Ljubljana
andrej.vernekar@moj.net
(KDE)

[†]Zdravstveni dom Ljubljana
Kotnikova 28, 1000 Ljubljana
r@aufbix.org
(OpenOffice.org)

[‡]Skupina za slovenjenje pri društvu LUGOS
LUGOS - Društvo uporabnikov Linuxa Slovenije
Predjamska 4, 1000 Ljubljana
andraz.tori1@guest@arnes.si
(Smart)

Povzetek

Projekti z odprto kodo so zelo privlačni za lokalizacijo. Eden pomembnejših projektov, ki hkrati tudi zelo dobro podpira lokalizacijo je vsekakor projekt grafičnega okolja KDE, ki ga sloveni skupina v okviru društva uporabnikov Linuxa Slovenije. Novejši projekt je slovenjenje OpenOffice.org, ki se od KDE razlikuje v več podrobnostih, s pomočjo programa Smart pa lahko uporabljamo že obstoječo bazo prevodov tudi v tem projektu. V skupno bazo prevodov pa smo z njegovo pomočjo dodali še prevode iz GNOME.

1. Uvod

Z vse večjo popularnostjo Linuxa se je pričel tudi razvoj grafičnih uporabniških okolij zanj, kot je na primer okolje KDE in pisarniških paketov, kot je OpenOffice.org. Ker so ta okolja namenjena običajnim uporabnikom, je lokalizacija precej bolj pomembna kot pri specializiranih orodjih za administratorje. Poleg dostopnosti kode nam pri projektu pomaga še dejstvo, da ga razvijajo ljudje s celega sveta, tako da se že razvijalci (večinoma) zavedajo pasti, ki čakajo prevajalce. Zato tudi ne preseneča, da je eden izmed ciljev projektov lokalizacija v čimveč jezikov. V različici KDE 3.0.2 jih je bilo na primer že 43, še nekaj novih pa se bo pridružilo v različici 3.1. Poleg tega projekti niso vezani na države (primer je poskus prevoda KDE v kurdščino), ponujajo pa možnost lokalizacije tudi manjšim narodom, ki s tržnega stališča niso dovolj zanimivi.

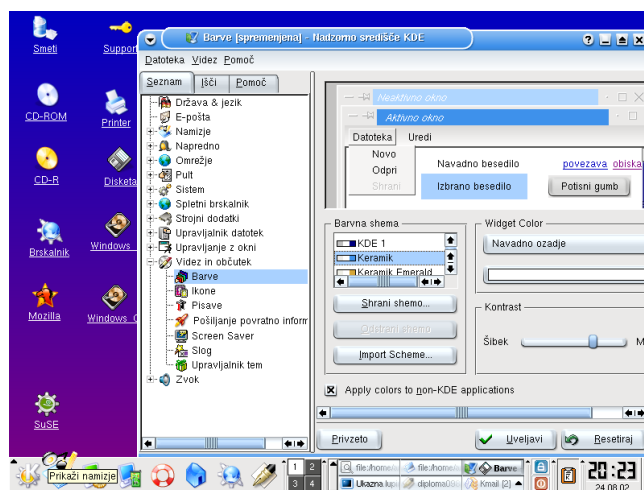
2. Namen članka

V tem članku želimo predstaviti osnove slovenjenja projektov KDE in OpenOffice.org, orodja in standarde, ki se pri tem uporabljajo, in osnove samih projektov, v okviru katerih poteka delo.

3. KDE

KDE (The K Desktop Environment) je prijazno grafično okolje za Linux/UNIX, ki je se je pričelo razvijati jeseni leta 1996. Projekt ponuja poleg grafičnega namizja še vrsto drugih programov (poštni odjemalec, spletni brskalnik, pisarniški paket ...), ki so enotnega videza in obnašanja.

Okolje ponuja tudi vrsto možnosti za lažji prehod uporabnikov z drugih operacijskih sistemov.



Slika 1: Bodoči KDE 3.1

3.1. Organizacija projekta

Projekt KDE je že v sami osnovi zasnovan tako, da omogoča čim lažjo lokalizacijo v množico jezikov, ki se razlikujejo ne le po uporabljenih pisavah ampak tudi po slogu pisanja (to velja za arabščino in hebrejščino, ki se pišeta od desne proti levi, kot tudi za vzhodnoazijske jezike, kot sta kitajščina in korejščina z uporabo pismenk).

To omogoča dosledna raba standarda Unicode (?) (ISO 10646), saj je to edini dovoljeni format za izvorno kodo, s katero imajo opravka prevajalci. Ti vse datoteke shranjujejo v kodiranju UTF-8, kar povrh vsega omogoča avtomatizacijo številnih opravil, ki bi jih sicer morali izvajati člani prevajalskih skupin: samodejna ekstrakcija novih sporočil, ki poteka enkrat dnevno, in zlivanje ustvarjenih predlog z obstoječimi katalogi, gradnja paketov in njihova objava, izdelava kataloga vseh prevodov, pretvorba med formatoma XML DocBook in PO in podobno.

Prevajalci tako lahko vsak dan naložijo kataloge sporočil in prevedejo nova sporočila oziroma popravijo stara, kar omogoča delo „v živo“. To pa hkrati pomeni, da lahko hkrati z angleško različico izidejo tudi vse ostale, ki izpolnjujejo pogoje.

3.1.1. Upravljanje projekta

Pogoji, ki jih mora jezik izpolnjevati, so naslednji: popolnoma prevedena paketa kdelibs.po in desktop.po ter 90% prevedenega paketa kdebase. Ker to še ne pomeni v celoti prevedenega okolja, sledijo še nove izdaje, ki so namenjene dokončanju prevodov in popravkom hroščev. Tako sta sledili različici KDE 3.0 še različici 3.0.1 in 3.0.2. V te različice se lahko uvrstijo tudi jeziki, ki niso uspeli priti v prvotno in so vmes dopolnili prevode.

Norme določa in nadzoruje korodinator kde-18n-doc, ki skrbi tudi za komunikacijo z razvijalci in sodeluje z upraviteljem izdaje.

Da se prevajalcem omogoči usladitev prevodov, zadnja dva tedna pred izdajo razvijalci ne smejo več spreminjati grafičnega vmesnika, za nujne popravke pa rabijo dovoljenje kde-i18n-doc.

3.2. Standardi in orodja

Poleg že omenjenega standarda Unicode uporablja KDE še vrsto drugih odprtih rešitev. Jeziki so tako označeni na primer po standardu ISO 639(?), za zapis pomoči se uporablja XML DocBook, za ekstrakcijo sporočil pa skrbi orodje gettext.

Ker je KDE poleg GNOME eden prvih tako obsežnih projektov, ki uporablja gettext, je naletel tudi na omejitve, ki pa so jih rešili nekoliko po svoje. Tak je na primer zapis množinskih oblik, ki jih je gettext rešil šele potem, ko je KDE uporabil svojo.

Še ena rešitev, ki je posebnost KDE, pa je pretvorba formata XML DocBook v format PO (gettext). Razlog za to je dejstvo, da je za format PO na voljo obsežna infrastruktura, ki skrbi za samodejno zlivanje različic, ki je za XML DocBook preprosto še ni.

Tako se format oblike

```
<TITLE>Options</TITLE>
<PARA>Some text to be translated</PARA>
```

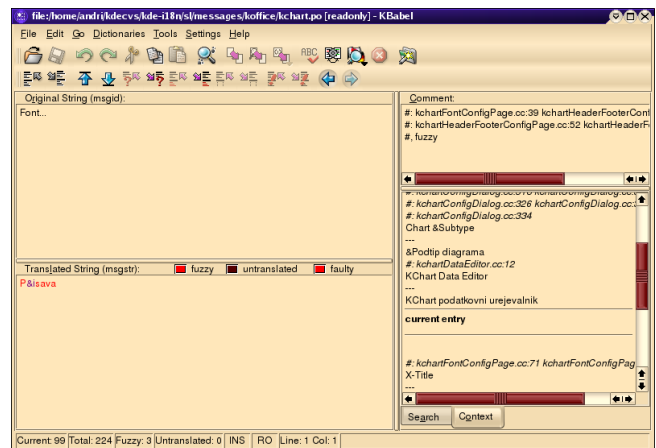
razbije v naslednjo PO datoteko:

```
msgid "Options"
msgstr ""
```

```
msgid "Some text to be translated"
msgstr ""
```

3.2.1. Orodja

Projekt ponuja tudi kup orodij, ki olajšajo delo prevajalcem. Poleg teh, ki tečejo samodejno na strežniku KDE (so pa na voljo tudi v paketu kdesdk, primer je xml2pot) je poglavitno orodje prevajalcev KBabel. Ta ima izvrsten upravljalnik katalogov sporočil, samodejno zna preverjati pravilnost datotek (tudi na primer število množinskih oblik!), skrbi za vnos podatkov o spremembah in avtorjih in prevajalcem onemogoča, da bi „pokvarili“ datoteke, kot se lahko hitro zgodi pri urejevalnikih besedil.



Slika 2: Urejevalnik KBabel

3.3. KDE in slovenščina

Med jeziki, v katere je preveden KDE je že od različice 1.1.2 naprej tudi slovenščina. Podprte so seveda tudi vse lokalne nastavitve, kot je tipkovnica, zapis datumov in števil ter valute.

3.3.1. Obseg prevodov

Uporabniški vmesnik (GUI) je preveden skoraj v celoti (manjkajo le nekateri bolj eksotični program iz novih paketov), preveden pa je tudi del dokumentacije, predvsem osnove uporabe okolja (Uporabniški priročnik, pomoč za datotečni upravljalnik in spletni brskalnik in za nadzorno središče).

Skupaj s projektom je rasla tudi baza prevodov, ki vsebuje že okoli 43 000 fraz. Tu velja poudariti, da je podvajanja fraz zelo malo, saj so takšne, ki se pojavljajo na več mestih večinoma predstavljene v eno samo datoteko - kdelibs.po (ki je zato tudi pogoj za uvrstitev jezika). Sama koda prevodov tako zavzema preko 11MB podatkov in če odštejemo originalna angleška sporočila in dodatne podatke v datotekah to še vedno predstavlja okoli 4 milijone znakov.

3.4. Proces prevajanja

Prevajalec s strežnika cvs.kde.org pobere zadnjo različico prevodov in jo v KBablu prevede. Za pomoč lahko za samodejno prevajanje ¹ uporabi bazo, ki jo ustvari KBabel ali pa program Smart, ki ga je napisal Andraž Tori. Ko je datoteka končana, se uvrsti nazaj v sistem za nadzor različic (CVS), tako da imajo do nje dostop tudi ostali

¹Ovisno od vrste datoteke je lahko rezultat med 5 do 30%

prevajalci. Ker projekt KDE ne omogoča neomejenega dostopa za pisanje v CVS (v skupini ga imajo trije člani) morajo ostali prevajalci poslati datoteke nekemu, ki ga ima, da jih uvrsti zanje.

3.4.1. Posebnosti

Oglejmo si rešitev, ki jo je projekt KDE uporabil za množinske oblike.

```
msgid ""
"_n: \%n directory\n"
"%n directories"
msgstr ""
"%n imenikov\n"
"%n imenik\n"
"%n imenika\n"
"%n imeniki"
```

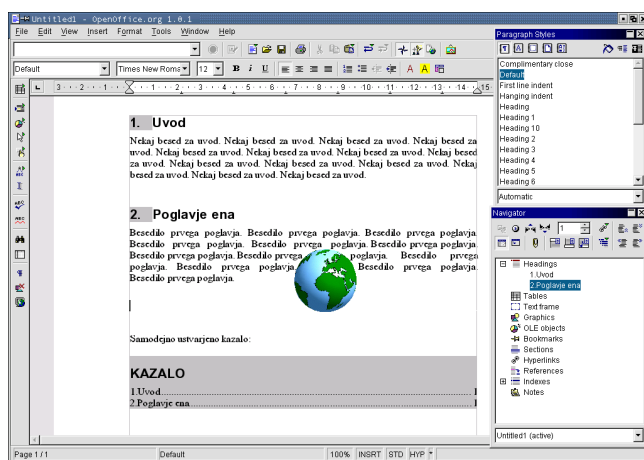
_n na začetku sporočila nam pove, da gre za množinske oblike, ki jih moramo oblikovati ustrezno jeziku, v katerega prevajamo (za slovenščino so to 4 oblike).

3.4.2. Prevajanje dokumentacije

Od KDE 2.1 dalje poteka proces pravzaprav skoraj povsem enako, kot prevajanje grafičnega okolja. Edino, na kar je treba paziti so oznake XML DocBooka. Osnovne (chapter, title itd.) pretvorba iz XML v PO skriva, nekatere vgnezdene pa lahko še vedno srečamo in jih seveda ne smemo pokvariti, saj lahko ena sama napaka prepreči generacijo celotnega dokumenta, zato se skladnost datotek preveri še z orodjem updateXML, ki generiran nazaj datoteko XML in jo tudi preveri, ali ustreza opisu (definiciji).

4. Prevajanje OpenOffice.org

OpenOffice.org je projekt prostega pisarniškega paketa, katerega koda je osnova za Star Office. Nastal je z objavo kode, ki jo pridobilo podjetje Sun z nakupom podjetja Star-Division. OpenOffice.org 1.0 je celovit pisarniški paket, ki vsebuje urejevalnik besedil, preglednico, predstavitevni in risarski program, deluje pa v okoljih Windows, Linux in Solaris.



Slika 3: Urejevalnik Writer iz pisarniškega paketa OpenOffice.org

4.1. Osnove projekta

Do danes je bil programski paket OpenOffice.org uspešno preveden v 27 jezikov, trenutno pa se aktivno prevaja še v vrsto drugih, med njimi od februarja 2002 tudi v slovenščino. Zaenkrat se lahko prevede le uporabniški vmesnik (GUI). Vsebine pomoči paketa se še ne prevaja, ker ob izdaji izvorne kode paketa ni bila last podjetja Sun Microsystems. Razvijalci obljublajo, da bo v doglednem času na voljo tudi prevajanje sporočil pomoči, ki naj bi jih združili v XML format.

Sporočila za prevajanje izvlečemo iz izvorne kode paketa s priloženim orodjem, ki nam iz izvornih datotek generira nekaj več kot 21.000 vrstic v enem od že prevedenih jezikov. Za prevajanje v slovenščino smo izbrali angleški jezik.

Generirana datoteka je navadna besedilna datoteka, katere posamezna polja so ločena s tabulatorji. Iz te datoteke je potrebno izluščiti vsebino treh polj, sporočila prevesti in jih uvrstiti na enako mesto v besedilni datoteki. Z istim orodjem se prevedena sporočila zlije nazaj v izvorne datoteke in paket je pripravljen za „gradnjo“ (build).

4.2. Oblika sporočil in orodja za prevajanje

Na začetku prevajanja smo naleteli na problem, katero orodje naj uporabimo za prevajanje. Razvijalci projekta so izdelali orodje za izvoz/zlivanje sporočil v/iz izvornih datotek, ne pa tudi za samo prevajanje sporočil. Ta korak je prepuščen prevajalcem. Na koncu smo se odločili za prevajanje z uporabo sistema gettext. Napisali smo skripte (BASH in Perl), ki nam poberejo sporočila iz besedilne datoteke in generirajo 28 datotek POT (po 28 imenih modulov v paketu). Za vsako sporočilo dobimo dve vrstici datoteki POT - v prvi je vsebina sporočila (msgid), v drugi pa prazen prostor za prevedeno sporočilo (msgstr). Posamezne POT datoteke prevajalec najprej delno samodejno prevede s pomočjo spletnega vmesnika Smart, preostanek sporočil pa je potrebno prevesti ročno v enem od urejevalnikov besedil. Odstotek samodejno prevedenih sporočil je odvisen od posamezne POT datoteke, giblje pa se med 5 in 40 %.

Več manjših datotek POT namesto ene velike nam omogoča, da si prevajalci delo razdelimo. Z istimi skripti se prevedene datoteke POT združi v prevedeno besedilno datoteko pravnega formata za zlivanje nazaj v izvorne datoteke.

4.3. Stanje prevoda

Prvi korak pri prevajanju OpenOffice.org je uspešno preveden slovar pogosto uporabljenih besed/fraz v programskem paketu OpenOffice.org. Slovar zajema približno 7000 vrstic. Ta prevod je zagotovilo, da se je posameznik oziroma skupina resno lotila lokalizacije projekta, po drugi strani pa prevajalcem zagotavlja konsistentnost prevodov.

Julija 2002 je bilo v celoti prevedenih 23 datotek, 5 pa jih je bilo še v fazi prevajanja. Ocenjujemo, da bodo POT datoteke v celoti prevedene do septembra 2002. Sledi še lektoriranje vseh prevedenih datotek in prva priprava paketov (build), predvidoma ob koncu leta 2002. Pred končno prevedeno različico bo potrebno testiranje celotnega paketa, med drugim tudi zato, ker v posameznih primerih ni

jasno razviden kontekst sporočila in se bo zagotovo pojavilo nekaj nesmiselnih prevodov. V času testiranja se bodo sporočila o napakah zbirala in se po končanem testiranju vključila v novo različico prevoda. Pri sami izdelavi paketa nismo vezani na nikogar, saj vsa stvar poteka pri nas.

Poleg samega prevajanja grafičnega vmesnika je bil prilagojen tudi črkovalnik. Temelji na prosto dostopnem črkovalniku Ispell, katerega format .aff datoteke je bilo potrebno prirediti za potrebe OpenOffice.org. Črkovalnik trenutno prepozna okrog 180.000 besed, do konca leta 2002 pa naj bi se število besed precej povečalo.

4.4. Z OpenOffice.org povezane dejavnosti/projekti

Postavljena je bila spletna stran <http://openoffice.lugos.si> z informacijami o stanju prevoda in novicami. Odprt forum za vprašanja, predloge, pripombe in drugo na temo prevajanja in uporabe OpenOffice.org najdete na naslovu <http://groups.yahoo.com/group/slooo>.

4.5. Novi projekti

Po prevedenem grafičnem vmesniku bo prvi korak vzpostavitev prevajalske skupine za prevajanje trenutno dostopne dokumentacije, ki obsega navodila za namestitev paketa, vodiče za začetek dela s programi v paketu, pogosto zastavljena vprašanja (FAQ) in nekaj predlog za posamezne programe v paketu. Poleg tega je na voljo prevajanje vsebine spletne strani projekta (<http://lang.openoffice.org>), kjer se lahko vse prevedene stvari tudi zbirajo.

4.6. Sodelavci in zahvale

Kdor aktivno sodeluje pri prevajanju OpenOffice.org, je naveden na spletni strani <http://openoffice.lugos.si> v rubriki „Trenutno aktivni prevajalci“. Zahvala gre Andražu Toriju za izvrstno orodje Smart in hitro realizacijo želja in pripomb, celotnemu dopisnemu seznamu lugos-slo@lugos.si za nasvete in predloge pri prevajanju, organizacijskemu odboru Lugos za vsesplošno pomoč, podjetju Agenda Open Systems za pomoč pri prevajanju ob pravem času in za celotno lektoriranje, MID za podporo pri projektu prilagoditve črkovalnika in vsem, ki kakorkoli podpirajo prevajanje paketa in njegovo uveljavitev.

5. Smart

Pri prevajanju programov z odprto kodo je še posebej privlačna možnost, da lahko uporabimo že obstoječe prevode sorodnih programov ali okolij. Žal pa so te pogosto v različnih formatih. Uporaba baze KDE (ki je najobsežnejša) v GNOME zaradi majhnih razlik ni bila dovolj preprosto izvedljiva, zato je nastalo orodje Smart.

Gre za program v perlu, ki shranjuje prevode v podatkovno bazo (PostgreSQL), hkrati pa omogoča uporabo le te brez specifičnih dodatkov posameznih okolij oziroma njihovo ustrezno pretvorbo. Tako lahko zamenja bližnjice iz KDE (&) s tistimi iz GNOME, briše komentarje, prilagaja velike in male črke in podobno.

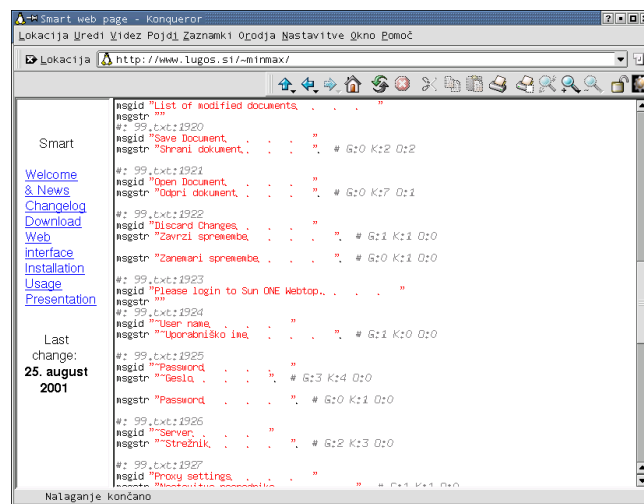
Program omogoča uvažanje baz tako iz projekta KDE, GNOME kot tudi iz OpenOffice.org (ki ima še nekaj svojih posebnosti). Z zadnjo različico ob uspešno prevedenih

sporočilih prikaže tudi, iz katerega dela podatkovne baze prevodov je bilo prevedeno sporočilo vzeto („G“ za GNOME, „K“ za KDE in „O“ za OpenOffice.org).

Poleg tega program omogoča pregled baze, kjer lahko iščemo podvojena sporočila (kar je uporabno za usklajevanje izrazja).

5.1. Spletni vmesnik

V okviru projekta je bil izdelan tudi spletni vmesnik, ki omogoča uporabo programa tudi iz okolij, v katerih ga iz različnih razlogov (različice knjižnic, starejše distribucije) nimamo nameščenega. Vmesnik najdemo na strani <http://www.lugos.si/minmax>. Podamo mu le ustrezno neprevedeno datoteko in dobimo delno prevedeno.



Slika 4: Spletni vmesnik programa Smart

6. Zaključek

Prosta programska oprema omogoča zainteresiranim posameznikom vključitev v že obstoječe ali ustvarjanje novih projektov lokalizacije. Rezultat tega je poleg končnega prevedenega programskega paketa tudi pojavitev orodij, ki prevajalcem olajšajo delo in prihranijo čas.

V zadnjem letu je skupina za slovenjenje proste programske opreme pri društvu LUGOS pridobila nekaj novih članov, ki so poleg pomoči pri obstoječih projektih začeli prevajati spletni brskalnik Mozilla.

7. Literatura

The KDE Translation HOWTO

<http://i18n.kde.org/translation-howto/index.php>

Slovenjenje namizja KDE

<http://www.lugos.si/delo/slo/KDE-s1/>

OpenOffice.org

<http://l10.openoffice.org>

OpenOffice.org; Universal i18n framework;

<http://l10n.openoffice.org/>

Universal_i18n_framework.pdf

Standard UNICODE

<http://www.unicode.org>

Poimenovanja jezikov po standardu ISO 639

<http://www.oasis-open.org/cover/iso639a.html>

N. Walsh, L. Muelner: Docbook: The Definitive Guide;

O'reily& Associates, 2002