

Pretraživanje pod Linuxom

Beagle search

Desktop pretraživanje napokon je stiglo i na Linux. Želimo li sve dokumente, poštu, web stranice, slike i glazbu imati indeksirane, pogledat ćemo što nam nudi Beagle

Piše: Ivan Capan

U posljednjih se nekoliko mjeseci podigla velika prašina oko nove, dotad malo zastupljene kategorije softvera - lokalnih, tj. desktop pretraživača. Budući da su kod većine korisnika višegodišnje nesustavno stvaranje sadržaja i nebriga o pospremanju rezultirali neorganiziranošću direktorija i datoteka, programi te vrste su nužnost. Za Windowse ih postoji nekoliko, no kakva je situacija pod Linuxom? Elementarnom pretraživanju može se pristupiti standardnim Linux alatima, poput grepa (pretražuje datoteke u potrazi za nizom znakova), finda (traži datoteke po raznim kriterijima, poput imena, veličine, vremena pristupa, dozvola i sl.) i locatea (traži po imenu i to trenutno, zbog baze svih imena). Iako se i na spomenute načine ponekad može naći tražena datoteka, navedeni alati se pokreću iz naredbenog retka, pa to obično znači i pisanje skripti, a i predstavlja opterećenje za računalo koje to treba pregledati. Nama treba neki program koji ima podršku za velik broj danas najčešćih dokumenata, sprema bazu na disk radi bržeg pretraživanja i, naravno, sve to pakira u dopadljivo sučelje (složili bismo se i s pristupom iz *browsera*).

Pretraživač Beagle

Program koji ispunjava naše zahtjeve, no još uvijek je relativno nepoznat, naziva se Beagle (www.gnome.org/projects/beagle). Njegova slaba zastupljenost može se objasniti i time da kao projekt postoji tek od ljeta prošle godine, a dogurao je tek do verzije 0.0.6.1. No, koliko možemo vidjeti, na razvoju radi nekoliko programera iz Gnome svijeta (na čijem se sučelju program primarno i bazira) i stvari se odvijaju brzim tempom. Na stranicama programa već je izgrađena cijela infrastruktura za ugodniji razvoj, a razvila se i zajednica korisnika - nadajući se kako će ovaj projekt prerasti u nešto vrijedno pozornosti.

Iako je funkcionalnost samog programa formirana (pogotovo uz oslanjanje na dodatne filtre za pojedini sadržaj), sami razvijatelji naglašavaju kako je on još uvijek u eksperimentalnoj fazi te ga ne preporučuju za svakodnevnu upotrebu. No, budući da smo mi od faze prikupljanja informacija do dovršenja članka doživjeli čak dvije nove verzije, program će vjerojatno u skorije vrijeme dostići fazu stabilnosti. Zasad ni sama instalacija nije jednostavna. Broj potrebnih biblioteka, dodatnih programa pa i krpanja jezgre zavidno je upravo zato što je pisano u novim tehnologijama (Mono). Mi smo ga uspjeli upogoniti tek na Ubuntu Linuxu, no korisnici Susea imaju već gotove pakete. Upute se nalaze na www.beaglewiki.org/index.php/Installing_Beagle, otkud slijede veze za pojedine distribucije (zasad samo Ubuntu i Gentoo).

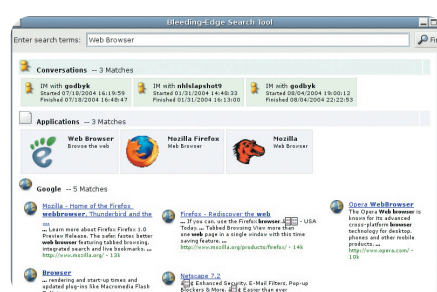
ožujak 2005.



S obzirom na razvojnu platformu, očekujemo da ga u nekom trenutku vidimo i na Windowsima. Već se i u ovoj fazi stvaraju izvršne datoteke s jednim, za Linux vrlo neobičajenim nastavkom EXE. To je izravni rezultat razvoja u C# programskom jeziku uz pomoć razvojne platforme Mono, bazirane na Microsoftovoj .NET arhitekturi. Njena je prednost lako prenošenje programa između sustava. Ovom zanimljivom temom svakako ćemo se jednom prilikom pozabaviti dublje. Onima s manje avanturizma preporučujemo da s Beagleom pričekaju neko vrijeme (vjerujemo da će nekoliko mjeseci biti dovoljno) i tek onda krenu u instalaciju. Mi smo ga ipak odlučili opisati jer vjerujemo u njegov potencijal i entuzijazam programera. Pogledajmo što sve nudi.

Mogućnosti

Prva stvar što će nas zanimati svakako su podržani formati koji će se indeksirati. Osim uredskih formata (MS Office i OpenOffice), koji su svakako nužnost, indeksira se i RTF, PDF, običan tekst, HTML stranice, *e-mailovi* i izvorni kod. Osim formata s tekstualnim sadržajima, moguće je arhivirati slike (JPG i PNG) i glazbu, prilikom čega se spremaju njihovi metapodaci, odnosno ID3 *tagovi*. Također se nazdize i povijest surfanja Firefoxom i Mozillom kao i konverzicije putem IM programa i IRC-a. Razlog što smo morali mijenjati jezgru je intofity, pod-sustav koji javlja Beagleu svaku promjenu ili stvaranje datoteke na sustavu. Ona će zato automatski biti indeksirana, bez primjetnih usporjenja. Beagle će nakon pokretanja pozadinskog programa beagled sam proći kroz naš direktorij i pregledati datoteke te ih dodati u bazu. Uz besplatnu registraciju na



▲ Uz datoteke na lokalnom disku, Beagle zna posjetiti i Google

api.google.com/createkey moguće je pretraživati i Internet. Podaci se spremaju unutar vlastitog direktorija, pa do curenja privatnih informacija ne bi trebalo doći (osim ako ne postavimo pogrešne dozvole).

Postoji nekoliko alata u naredbenom retku, no najčešće ćemo pokretati best, grafičko sučelje za pretraživanje. Best se integrira s Gnomeom, postavljajući ikonicu u *System Tray*, odakle je dovoljan jedan klik da počnemo pretraživati. Program brzo vraća rezultate, i to sve istom prozoru. Pojavi li se nova datoteka koja zadovoljava kriterij nakon što smo ga zadali, u sučelju će se njeno ime odmah pojaviti, bez potrebe za pokretanjem nove pretrage. Nismo uspjeli osposobiti pretraživanje putem *browsera*. Osim nestabilnosti i neprepoznavanja većeg broja nominalno podržanih formata, uglavnom smo zadovoljni vidjenim. No, ipak ćemo Beagleu dati malo vremena prije aktivnijeg korištenja. Demonstracija rada može se naći na stranicama <http://nat.org/demos>.