

Projekt linux.hr – integracija web-a i baza podataka na siguran način

Što je linux.hr?

Na početku dobro se upoznati sa time što nudi web site linux.hr i čemu on služi. Mislim da nije potrebno detaljno objašnjavati što je web site. To su jednostavno stranice koje poduzeće ili udruga postavlja na Internet da bi omogućila svima zainteresiranima pristup podacima o toj firmi ili udruzi.

Upravo zbog tih razloga, Hrvatska Udruga Linux Korisnika (HULK) je postavila svoje web stranice na adresi linux.hr. Na njima svi zainteresirani mogu naći podatke o samoj udruzi, podatke o projektima u kojima sudjeluju članovi udruge, linkove na korisne informacije na mreži u vezi sa Linux-a i često postavljana pitanja (FAQ). Sam poslužitelj pruža još neke servise kao što su to ftp poslužitelj ftp.linux.hr i web poslužitelj faq.linux.hr na kojem se nalazi kolekcija korisnih savjeta prikupljenih sa usenet grupa, međutim njima se nećemo detaljnije baviti u ovome članku.



Slika 1: početna stranica na web poslužitelju HULK-a

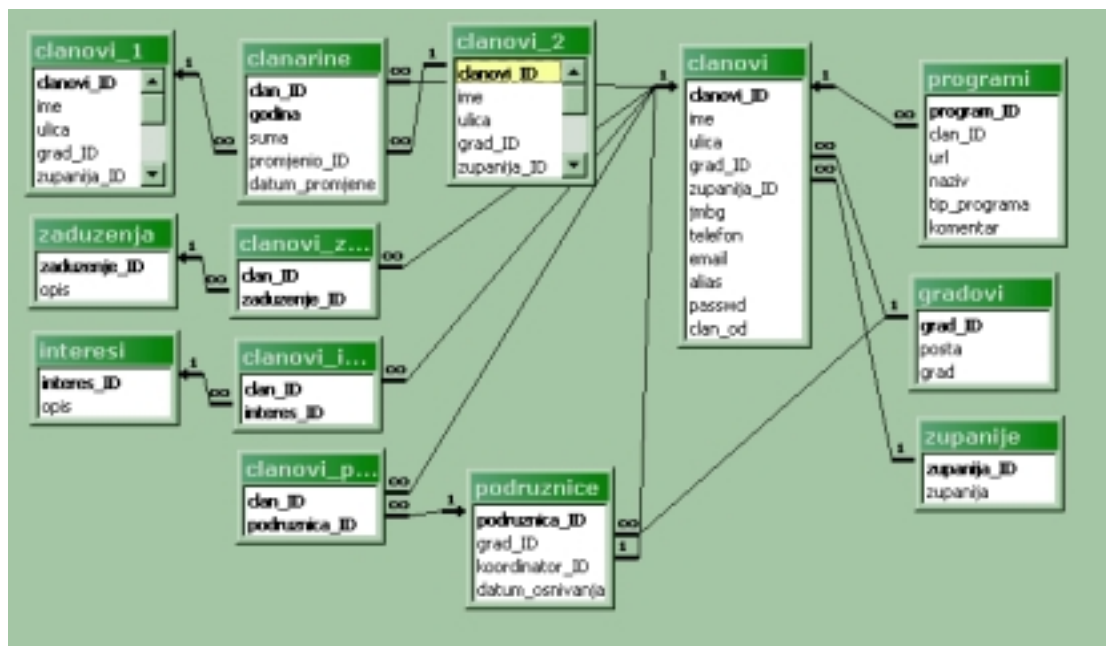
Kakve smo potrebe imali?

Razvijajući stranice HULK-a pojavila se potreba za održavanjem baze podataka članova preko web-a. Na taj način bi naša udruga mogla pružati zainteresiranima podatke o svojim članovima (jer bi članovi unosili svoje interese kao dio unosa svojih podataka), u svakom trenutku bi imali jasan popis članova i osnovnih podataka o njima (kao što je to e-mail adresa) te bi imali mogućnost održavanja naših administrativnih podataka (tko obavlja koju dužnost, plaćanje članarina itd.) sa bilo kojeg web preglednika bez obzira na fizičku poziciju (što je također veoma važno, jer u ovom trenutku predsjednik udruge većinu vremena provodi u Rijeci, dopredsjednik u Varaždinu a tajnik u Zagrebu).

Ideja vodilja od samoga početka stvaranja web stranica linux.hr bio je zahtjev za brzinom pristupa (jer većina naših korisnika dolazi putem modema), što je rezultiralo umjerenim korištenjem grafike. Drugi uvjet je bila potpuna čitljivost u web preglednicima koji mogu prikazati samo tekst (npr. lynx). To ograničenje ne izgleda na prvi pogled naročito važno, ali naša iskustva govore da je to dobra praksa (za primjer pokušajte lynx-om pronaći televizijski program na stranicama Hrvatske televizije), te ne namjeravamo uskoro odustati od toga ograničenja.

Koji alat odabrati?

Pri odabiru alata kojim ćemo se poslužiti pri izradi HULK baze podataka našli smo se pred problemima. Kako autor posjeduje znanje o relacijskim bazama podataka, struktura baze je bila modelirana prema teoretskim znanjima koja je sakupio u toku studija. Da bi se ispunile teorijske osnove relacijskih baza podataka (RDBMS-a) opis baze (tj. atributi koji su bili definirani unutar baze podataka) nisu odgovarali izgledu formi koje su se trebale pojavljivati korisnicima. U nekim tablicama je bilo "previše" atributa (tipičan primjer toga je atribut *password* kojeg, iako se nalazi u tablici *članovi*, ne treba ispisivati skoro nikada na ekran). Jedan od većih problema u ovom koraku dizajna je bio nedostatak alata za brzi razvoj aplikacija (RAD - *Rapid Application Development*) ili alata pomoću kojim se može grafički dizajnirati izgled baze. Iako takvi alati postoje uz komercijalne baze na Linux-u, mi smo se odlučili za korištenje Access-a u tu svrhu. Grafički prikaz tablica i relacija iz baze nalaze se na slijedećoj slici.



Slika 2: relacijska shema baze podataka

Upravo zbog tog ograničenja skoro odmah su otpali svi alati koji su forme definirali direktno iz fizičkog opisa baze. Autor je također poznao *perl* i DBI pristup bazama podataka iz *perl-a*, pa se *perl* pojavljivao kao logičan izbor programskog jezika za rješavanje ovih problema. Međutim, kao što će slijedeća tablica pokazati, na kraju nije bilo tako.

alat	za	protiv
html + perl cgi + dbi	<ul style="list-style-type: none"> • poznavanje alata • isprobane tehnike, iskustvo 	<ul style="list-style-type: none"> • održavanje odvojenih datoteka za html dokumente i cgi skripte • mala razina

		automatizacije stvaranja html formi <ul style="list-style-type: none"> • relativno kompliciran pristup bazi korištenjem dbi-ja
razni dodaci dbi-ju	<ul style="list-style-type: none"> • jednostavnije generiranje html koda 	<ul style="list-style-type: none"> • većinom ne koriste SQL kao upitni jezik koji autor poznaje • rješenja koja "skoro" odgovaraju, ali ne sasvim
embedded perl	<ul style="list-style-type: none"> • uklanja nedostatke održavanja odvojenih datoteka za html i perl kod • jednostavno generiranje html-a iz baze podataka 	<ul style="list-style-type: none"> • kompliciran pristup bazi zasnovan na dbi-ju
www-sql	<ul style="list-style-type: none"> • SQL upiti pišu se unutar html dokumenta • jednostavno za korištenje • jednostavno za nadogradnju 	<ul style="list-style-type: none"> • teže kreiranje formi za unos od alata koji su prilagođeni točno tome – kao neki dodaci za dbi • ograničenost – koja sili na vlastite nadogradnje, npr. nemogućnost slanja e-mail-a.

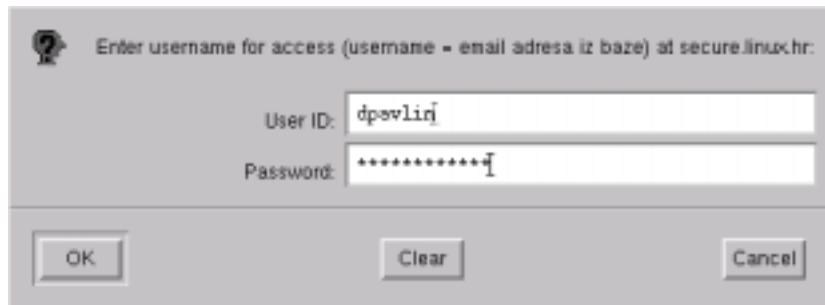
Odabir alata odužio se na skoro tri mjeseca. Naime, svaki alat je u početku pokazivao velike nade, ali se ubrzo nakon toga došlo do nekog od njegovih ograničenja zbog čega je bio napušten. Postavljanje upita u bazu programiranjem u *perl*-u (u bilo kojem obliku) preko *dbi*-ja je odbačeno jer je kod napravljen za to puno nečitljiviji od html-a, i na taj način teži za održavanje. Svi ostali alati su zahtijevali postavljanje upita na neki novi, njima logičan način. Autor međutim, preferira SQL kao upitni jezik, pa je nakon isprobavanja većeg broja takvih alata ipak zaključio da želi koristiti upravo SQL. *www-sql*, sa druge strane, je jednostavan set dodataka html-u koji omogućavaju postavljanje upita u bazu korištenjem SQL-a, jednostavno prikazivanje rezultata tih upita preko html-a na ekranu korisnika te jednostavna grananja (if-then-else tipa). Iako je bio kao stvoren za ovakve projekte, nakon nekog vremena pokazao je svoje manje nedostatke, koji su riješeni unapređivanjem samog *www-sql*-a i/ili pisanjem cgi skripti u *perl*-u. Tipičan primjer je bilo slanje e-mail-a iz baze podataka. Iako postoji mogućnost da *www-sql* pokrene vanjski program, nije bilo moguće generirati poruku koju je trebalo poslati. U tom slučaju *www-sql* stranica poziva cgi skriptu napisanu u *perl*-u koja priprema mail i šalje ga.

Drugi osnovni problem je bio problem autorizacije (tj. predstavljanja) korisnika koji posjećuju web stranice. Naime, potrebno je da korisnici imaju različite ovlasti pri radu sa podacima (obični korisnici imaju mogućnost mijenjanja svojih podataka, dok tajnik udruge ima mogućnost mijenjanja podataka o plaćenim članarinama za sve korisnike). Odabrali smo koristiti standardnu autorizaciju koja se primjenjuje na web server-ima (tzv. HTTP autorizacija) bez obzira što je njezina upotreba prilično nesigurna (tj. lozinke kod takve autorizacije se relativno jednostavno mogu prislušivati bilo gdje na mreži gdje taj promet prolazi). Problem sigurnosti smo riješili korištenjem *HTTPS* protokola koji omogućava enkripciju (zaštitu podataka pri prijenosu matematičkim metodama) svih upita i odgovora.

Korištene tehnologije

Pri izradi baze korisnika, koja je kao i svaki pravi projekt već zakasnila svoj planirani završetak između tri i šest mjeseci, koristili smo slijedeće tehnologije:

- HTTP – *HyperText Transfer Protokol* - internet protokol koji služi za prijenos datoteka između web poslužitelja (*server-a*) i web preglednika (*browsera*). Mi koristimo Apache web poslužitelj.
- HTTP autorizaciju – omogućava da se korisnik predstavi serveru korištenjem korisničkog imena i zaporke (*passworda*), dok server onda na osnovu tih podataka omogućava ili sprječava pristup određenim stranicama. Podaci iz HTTP autorizacije koriste se i kod upita u bazu, tako da samo autorizirani korisnici mogu mijenjati podatke



Slika 3: prozor za HTTP autorizaciju

- HTTPS – *Secure HTTP* – dodatak HTTP protokolu koji korištenjem SSL biblioteke štiti prenesene podatke od prisluškivanja i promjena korištenjem kriptografskih ključeva
- SSL – *Secure Sockets Library* – biblioteka funkcija koja podržava enkriptirane veze
- RDBMS – *Relational DataBase Management System* – relacijske baze podataka objašnjene su u predavanju kolege Hrvoja Vulina. Ukratko, radi se o više tablica povezanih vezama koje omogućavaju jednostavnu pohranu i pretraživanje na taj način strukturiranih podataka.
 - MySQL – relacijska baza za koju smo se mi odlučili je MySQL. To je baza čija je upotreba besplatna u nekomercijalne svrhe i podržava veći dio standarda SQL/92.
 - SQL upitni jezik – kojega podržava MySQL baza za upite prema bazi
- PERL – *Practical Extraction and Report Language* – programski jezik koji je u mnogome sličan C-u, a specijaliziran je za obradu tekstualnih datoteka, pa je zato posebno pogodan za korištenje za CGI programe. Perl se u našem projektu linux.hr pojavljuje u dva osnovna oblika:
 - CGI skripte – *Common Gateway Interface* – CGI je sučelje (API) između programa koje pokreće web poslužitelj (napisanih od strane korisnika) i samog web poslužitelja koji na sebe preuzima slanje podataka krajnjem korisniku (korištenjem standardnog HTTP protokola)
 - embeded Perl – poseban dodatak našem Apache server-u koji omogućava da se Perl programi pišu direktno u HTML kod, te da se izvršavaju kod pozivanja stranice.
- Glimpse – program za indeksiranje tekstualnih datoteka radi lakšeg pronalaženja tražene datoteke prema jednoj ili više riječi. Koristi se kod pretraživanja
- JavaScript – novotarija firme Netscape potpuno nevezana uz jezik Javu. Omogućuje izvršavanje programa kod samog korisnika. Zbog početnih uvijeta koristi se samo za dodatnu provjeru podataka kod unosa (npr. JMBG) i na taj način pomaže korisnicima kod unosa bez opterećivanja poslužitelja.

- www-sql – prema mišljenju autora, najbolji alat za izradu web stranica koje moraju imati pristup bazama podataka. Iako je u osnovi veoma jednostavan, dovoljno je moćan za skoro sve primjene. Kako dolazi u izvornom kodu, moguće ga je modificirati prema vlastitim potrebama, što je autor ovoga teksta i učinio, te poboljšanja poslao autoru da bi ih uvrstio u slijedeću verziju. Time se pokazuje da je *open source* dobar način razvoja programske podrške (www.opensource.org).

Zaključak

Ovaj projekt pokazuje da je moguće napraviti integraciju web poslužitelja i baza podataka sa poštivanjem sigurnosnih zahtjeva koji se postavljaju u slučaju da korisnici moraju pristupati takvoj bazi preko Interneta korištenjem **otvorenih sustava** kao što je to **Linux**, te korištenjem alata koji su besplatni ili na raspolaganju u izvornom kodu (**open source**).

Buduće primjene baze podataka uključuju praćenje programa i interesa članova Hrvatske Udruge Linux Korisnika (HULK-a). Praćenje svih administrativnih podataka o članovima udruge moguće je obaviti korištenjem odgovarajućeg pristupnog korisničkog imena i zaporke sa bilo kojeg web preglednika povezanog Internet-om na siguran i jednostavan način na zadovoljstvo korisnika ovog projekta.

The screenshot shows a Netscape browser window displaying a web page titled "Rad sa osobnim podacima". The page has a navigation bar with links for "HULK", "Projekti", "Linkovi", and "FAQ". The main content area contains a form with the following fields:

- Ime i prezime: Dobrica Pavlinušić
- Ulica i broj: Vrbanićeva 2
- Grad: 10000 Zagreb (with a dropdown arrow) and županija: Zagrebačka (with a dropdown arrow)
- JMBG: 0801974330019
- Telefon(i): 01/46-49-524
- e-mail: dpavlin@foi.hr
- Password: (empty field) with the note "Ovu lozinku koristite za pristup vašim podacima"

At the bottom of the form is a button labeled "Promjeni podatke". The browser's address bar shows the URL "https://secure.linux.hr/clanovi/" and the status bar shows "100%".

Slika 4: izgled forme sa podacima o članovima

Korisne adrese:

1. Stranice HULK-a: <http://linux.hr/>
2. Programski jezik Perl: <http://www.perl.com/>
3. CPAN arhiva programa i dodataka za Perl: <ftp://ftp.linux.hr/pub/CPAN/>
4. Apache web poslužitelj: <http://www.apache.org>
5. MySQL baza podataka: <http://www.mysql.com>
6. www-sql stranica: <http://www.daa.com.au/~james/www-sql/>
7. web stranica autora: <http://www.foi.hr/~dpavlin/>