

Web2.0 - a gdje su podaci?

- prijedlog predavanja -- DRAFT *

te=0x05dafe0 [0, 0]

Contents: [Dobrica PavlinuÅ¡iÄ 's random unstructured stuff]

- Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (Web2.0 - a gdje su podaci?)
- Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (SaÅ¼etak)
 - ◆ Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (1.Da li je Web 2.0 dÅ¼ungla podataka?)
 - ◆ Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (2.Primjer integracije razliÄ itih izvora podataka)
 - ◆ Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (3.IRC arhiva)
 - ◆ Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (4.Kako organizirati poruke? Tagovi!)
 - ◆ Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (5.A svi ostali izvori podataka?)
 - ◆ Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (6.Svaka prava komunikacija je dvosmjerna!)
 - ◆ Dobrica PavlinuÅ¡iÄ 's random unstructured stuff (7.Sinergijski efekt povezivanja)

SaÅ¼etak

U danaÅ¡nje doba Interneta problem dostupnosti informacija napokon viÅ¡e nije zanimljiva tema svakodnevnog razgovora. BaÅ¡i napravotiv, tolika dostupnost stvorila je sasvim novi problem viÅ¡ka informacija koje nas napadaju sa svim strana i zahtijevaju od nas dodatan napor u odabiru samo onih koje su nam stvarno korisne.

Zbog toga Ä emo u ovom radu pokuÅ¡ati pogledati koje sve tipove informacija susreÄ emo i prikazati jedan od moguÄ ih naÄ ina za organiziranje razliÄ itih izvora informacija u korisnu cjelinu koristeÄ i IRC, RSS, tagove i mnogo maÅ¡ite.

1.Da li je Web 2.0 dÅ¼ungla podataka?

Da bi lakÅ¡e pogledali sa kojim se sve Web2.0 podacima svakodnevno susreÄ emo , podijelit Ä emo ih prema naÄ inu primanja: oni mogu dolaziti do nas bez naÅ¡e intervencije (push), ili moramo po podatke otiÄ i na neku adresu na mreÅ¼i (pull). Da bi bilo jasnije, pogledajmo nekoliko najpopularnijih mreÅ¼nih servisa prema ovoj podjeli:

- pull (korisnik "odlazi" po podatke na mreÅ¼u)
 - ◆ Web (HTTP)
 - ◆ RSS (Atom, podcasts...)
- push (dolaze bez intervencije korisnika)
 - ◆ e-mail (STMP)
 - ◆ Instant Messaging (IRC, IRQ, Jabber...)

Ako pogledamo ovu podjelu od vrha prena dnu, primjetiti Ä emo da podaci imaju razliÄ itu "svjeÅ¼inu" u zavisnosti od toga kojim naÄ inom dolaze do nas. Ä toviÅ¡e, mi zapravo koristimo razliÄ ite naÄ ine komunikacije sa istim ljudima u zavisnosti od toga koliko nam brzo treba njihov odgovor.

Već ina spomenutih načina komunikacije je asinhrona i ne zahtijeva da oba korisnika budu spojena na mrežu u isto vrijeme. Iznimka je instant messaging za koji je potrebno da obje strane budu istovremeno spojene na mrežu.

Zanimljivo je primjetiti da mnogi korisnici interneta koriste više različitih davatelja usluga istog servisa. Uobičajno je imati nekoliko različitih e-mail adresa u zavisnosti sa kojim krugom ljudi komuniciramo (npr. poslovnu i privatnu). Taj problem davno je riješen mogućnošću u prosječanjanja e-mail poruka na jedno centralno mjesto, a sličan pristup koristimo i za RSS izvore koje zapravo sakupljamo na jednom mjestu da bi ih pročitali.

Upravo to okupljanje podataka na jednom mjestu i njihovo filtriranje je princip koji pokušavamo iskoristiti ne bi li na jednostavan način mogli zavladatai tom gomilom podataka kojom nas mreža zatrpava.

Sa jedne strane, standardni protokoli za razmjenu podataka (HTTP, SMTP, Jabber, IRC...), i načina zapisa (HTML, XML, e-mail poruke) omogućavaju nam da takav mashup napravimo, ali sa druge strane način i alat koji točno zadovoljava naše potrebe (i ima podršku za sve tipove izvora podataka koji su nam zanimljivi) nije lako.

2. Primjer integracije različitih izvora podataka

Naši primjer je grupa računarskih zanesenjaka koja se okuplja u stvarnom svijetu i virtualno na IRC kanalu. Osim IRC kanala koji je centralni komunikacijski medij, koristi se također i wiki te mailing lista, ali većina komunikacije (i najveći izazov) je upravo bio IRC kanal.

Prva logična ideja je bila napraviti web arhivu koja bi omogućavala da na jednostavan način pogledamo što se dešavalo na irc kanalu kada nismo bili spojeni. Tako je nastao projekt irc-logger (pomalo otkivanog imena) koji je prerastao u zanimljivu kombinaciju ulazno-izlaznih protokola a sve u svrhu objedinjavanja zanimljivih informacija na jednom mjestu.

3. IRC arhiva

Da bi stvorili arhivu poruka na IRC kanalu, irc-logger je spojen kao običan korisnik (bot) na IRC server. Na taj način on vidi sve poruke na kanalu i može stvoriti arhivu, te je prikazati na web-u (po datumu) ili pretražiti po bilo kojoj riječi iz sadržaja.

Zanimljiva je mogućnost pretraživanja arhive korištenjem IRC klijenta. Slanjem poruke

```
/msg irc-logger search proba
```

dobiti ćemo (u obliku poruka) sve zapise u arhivi koji su sadržavali riječ proba. Na taj način zapravo imamo mogućnost korištenja arhive i pretraživanja na dva načina: preko web-a i preko IRC poruka.

4. Kako organizirati poruke? Tagovi!

Da bi omogućili jednostavnije organiziranje poruka, odlučili smo uvesti dogovor za tagove u poruke na IRC-u.

Bilo koja riječ koja završava sa dva znaka kroz (//) na taj način postaje tag. Tagovi se prikazuju u obliku oblaka korištenja u web arhivi, ali i omogućavaju jednostavno pozivanje svih poruka koje

koriste isti tag.

5.A svi ostali izvori podataka?

Mnogo potencijalno zanimljivih podataka za članove naše grupe već se nalazi na nekim servisima na mreži. Najbolji primjer su možda linkovi na del.icio.us-u ili filtrirane poruke iz raznih RSS ita. Na sreću, većina tih servisa omogućava generiranje RSS feeda sa tim podacima.

Zbog toga, irc-logger može pratiti proizvoljan broj RSS feedova i slati promjene u njima na sam kanal. Prikazivanje promjena stranica na wiki-ju (preko RSS feeda promjena) na IRC kanalu pokazao se kao zgodan mehanizam međusobnog pritiska kada treba napisati neku najavu: netko napíše poetak, svi dobiju obavijest na IRC kanal i ubrzo je cijela obavijest gotova!

6.Svaka prava komunikacija je dvosmjerna!

Osim mogućnost da prima podatke preko RSS-a, irc-logger može i stvoriti RSS feedove sa oblakom korištenih tagova ili zadnjim porukama koje su u sebi imate neki od tagova. Taj RSS feed je onda koristan za uključivanje u druge servise, a mi ga koristimo unutar wiki-ja koji na taj način postaje uvijek živ, čak i kada ga nitko ne editira.

7.Sinergijski efekt povezivanja

Većina mogućnosti irc-loggera na prvi pogled izgleda prilično uobičajno i ne naročito zanimljivo. Ali, kada pogledamo njihovu kombinaciju dolazimo do zaključka da kombinacijom standardnih protokola i malo možete napraviti zanimljiv mashup koji se u našem slučaju pokazao kao vrlo koristan članovima naše grupe.