



UNIVERZA V LJUBLJANI

FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Ljiljana Perkovič

Primerjava Web 2.0 in Web 3.0

DIPLOMSKO DELO

NA VISOKOŠOLSKEM STROKOVNEM ŠTUDIJU

Mentor: doc. dr. Rok Rupnik

Ljubljana, 2011



Št. naloge: 00056/2010

Datum: 06.12.2010

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **LJILJANA PERKOVIĆ**

Naslov: **PRIMERJAVA WEB 2.0 IN WEB 3.0**
THE COMPARISON OF WEB 2.0 AND WEB 3.0

Vrsta naloge: Diplomsko delo visokošolskega strokovnega študija prve stopnje

Tematika naloge:

Proučite standarda Web 2.0 in Web 3.0. Na podlagi tega primerjejte oba standarda, poudarek pri primerjavi dajte predvsem področju iskalnikov in načinih obravnavanja semantike.

Mentor:

doc. dr. Rok Rupnik

Dekan:

prof. dr. Nikolaj Zimic



Rezultati diplomskega dela so intelektualna lastnina Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljane ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje Fakultete za računalništvo in informatiko ter mentorja.

IZJAVA O AVTORSTVU

diplomskega dela

Spodaj podpisani/-a _____,
z vpisno številko _____,

sem avtor/-ica diplomskega dela z naslovom:

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal/-a samostojno pod mentorstvom (naziv, ime in priimek)

in somentorstvom (naziv, ime in priimek)

- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki »Dela FRI«.

V Ljubljani, dne _____ Podpis avtorja/-ice: _____

Zahvala

Zahvaljujem se predvsem staršem, ki so mi omogočili študij v Sloveniji, ter so mi bili zelo velika podpora za čas študija na Fakulteti za računalništvo in informatiko. Zahvalila bi se tudi mentorju s katerim sem lepo sodelovala, prijateljem ter fantu, ki je imel veliko razumevanja in mi je bil podpora za dokončanje študija.

Kazalo

POVZETEK.....	1
ABSTRACT	2
1. UVOD	3
1.1 KAJ JE SPLET?	3
1.2 KAKO DELUJE SPLET?	3
1.3 RAZVOJ SPLET-A	4
2. WEB 2.0.....	6
2.1 DEFINICIJA	6
2.2 RAZVOJ	6
2.3 ZNAČILNOSTI.....	7
2.3.1 Sestavni deli Web 2.0	8
2.4 TEHNOLOGIJE	9
2.4.1 Principi Web 2.0 tehnologije.....	11
3. WEB 3.0.....	14
3.1 DEFINICIJA	15
3.2 RAZVOJ	15
3.3 ZNAČILNOSTI.....	16
3.4 TEHNOLOGIJE	18
4. PRIMERJAVA WEB-A 2.0 Z WEB-OM 3.0.....	22
4.1 KRITERIJI ZA PRIMERJANJE	22
4.2 ISKALNIKI DANES IN JUTRI.....	23
4.3 MIKROFORMATI IN RDFa.....	29
4.3.1 Mikroformati	30
4.3.2 RDFa.....	33
4.3.3 Mikroformati vs. RDFa	37
5. SKLEP	43
PRILOGE	44
6. LITERATURA	45

Seznam uporabljenih kratic in simbolov

Web - kratica za WWW

WWW - world wide web ali svetovni splet ali samo splet

W3C - World Wide Web Consortium ureja telo spletnih standardov in protokolov

Blog - pomeni spletni dnevnik; to je spletna stran, ki periodično prikazuje besedila, slike in druge elemente, ki jih njihovi avtorji sproti dodajajo

Peer to peer – P2P omrežja omogočajo izmenjavo datotek (npr. glasbe, filmov, računalniške opreme) preko Interneta

Podcasting - je skupek tehnologij za prenos zvočnih ali video datotek iz interneta po naročilu.

Wiki – spletna aplikacija, ki omogoča enostavno urejanje in dodajanje vsebin (spletnih strani) vsem registriranim uporabnikom

Metapodatki - vsebujejo podatke o podatkih; obsegajo podatke, ki se nanašajo na vsebino, strukturo, kvaliteto, lastništvo, distribucijo, tehnologijo, namen, uporabnost in druge elemente, ki so pomembni za pravilno interpretacijo oziroma uporabo podatkov.

Ontologija - Predstavlja koncepte, lastnosti, hierarhična in ne-hierarhična razmerja med koncepti in posamezniki v strojno obdelanem načinu

Spamming - nadležna pošta (tudi vsiljena, nezaželena ali nenaročena elektronska pošta) je pošiljanje enakih ali podobnih sporočil na veliko število (tisoče, milijone) naslovov

Mashup - v spletnem razvoju, Mashup je spletna stran ali program, ki uporablja in združuje podatke, predstavitev ali funkcionalnost iz dveh ali več virov za ustvarjanje novih storitev.

Folksonomija – sistem razvrščanja s pomočjo značk (tagov), ki nastane v procesu sodelovanja uporabnikov pri ustvarjanju vsebine

Parsing – v računalništvu, parser je ena izmed sestavin tolmača ali prevajalnika, ki preverja pravilno sintakso in gradi strukture podatkov (pogosto neke vrste strukture dokumenta, abstraktno sintakso drevesa ali druge hierarhične strukture), ki posredno izhaja iz vhodnega tokena

Imenski prostor – angl. namespace je edinstveno ime, ki identificira organizacijo, ki je razvila XML shemo. Imenski prostor je opredeljen z Uniform Resource Identifier (URL ali URI). Uporaba imenskih prostorov omogoča opredelitev elementa, ki se nedvoumno identificira s URI, čeprav se lahko oznaka "title" sama pojavi v številnih nizih metapodatkov

Povzetek

Danes je Internet nekaj, brez česa si ne moremo zamisliti življenja. Ko se soočimo z nekakšnim problemom, nam je vedno prva asocijacija iskanje rešitve na Internetu. To misel nam je omogočil razvoj Web 2.0. Danes, ko iščemo rešitev, se nam lahko iskanje zavleče tudi na nekaj ur, zato omogoča Web 3.0 standard uporabnikom predvsem hitro in učinkovito iskanje rešitve na Internetu. Prav tako pa omogoča računalnikom razumevanje človeškega govora, da bi bilo iskanje tako bolj učinkovito.

V diplomski nalogi sem se obrnila na razlago Web 2.0 in Web 3.0 standardov ter njihovega medsebojnega sodelovanja. Najprej sem pokazala kako delujejo iskalniki obdobja Web 2.0 in kakšen je načrt delovanja v obdobju Web 3.0. Potem sem se obrnila na tehnologije teh dveh standardov, ki z medsebojnim sodelovanjem omogočajo veliko. RDFa in Mikroformati sta dva najbolj pogosta načina za vgradnjo semantike v spletne dokumente.

Web 2.0 uporablja splet za branje in pisanje, bloge, spletne aplikacije, bogate medije, virusne medije, označevanje ali folksonomijo med izmenjavo vsebin, in osredotočanje na skupnosti.

Web 3.0 standard uporablja semantični splet, mashup-e, pripomočke, obnašanje uporabnikov ali Jaz-onomijo, oglase, uporabniško angažiranost, združuje dinamično vsebino in se osredotoča na posameznike.

Predlagana širitev Web 3.0 tehnologije vključuje tudi naslednje:

- Mobilne naprave, vsenavzočo povezljivost, mobilni dostop do interneta s širokopasovnim sprejetjem
- Odperte tehnologije, odprte formate podatkov in podatkovno dovoljenje, odprte API-je in protokole
- Odprto identiteto, odprt ugled, osebne podatke in prenosno gostovanje identitete
- Inteligentne spletne tehnologije, semantične tehnike kot so OWL, RDF, SPARQL, SWRL, in izjavo na osnovi podatkov prodajalnah s semantičnimi aplikacijami
- Inteligentne aplikacije, strojno učenje, avtonomne agente, strojno sklepanje in obdelavo naravnega jezika.

Ključne besede: Web 2.0, Web 3.0, spletne tehnologije, iskalniki, mikroformati, RDFa metapodatki, povezovanje

Abstract

Today we can't imagine our lives without the Internet. When we are faced with some kind of a problem, there is always a first association of finding a solution on the Internet. This idea has enabled us Web 2.0 development. Today, as we seek informations, we can search for solutions even for a couple of hours. Therefore, allows Web 3.0 standard users particular fast and efficient search solutions on the Internet and computer understanding of human speech, in order for the serch to be more effective.

In this thesis I turned to the interpretation of Web 2.0 and Web 3.0 standards and their mutual cooperation. First, I showed how the search engines of the period of Web 2.0 work and what is the plan of working for a period of Web 3.0. Then I turned to the technologies of these two standards, which enable a lot with working together. RDFa and microformats are two of the most common ways for embedding semantics into web documents.

Web 2.0 uses the read write web, blogs, web applications, rich media, viral media, tagging or Folksonomy while sharing content and focusing on communities.

The Web 3.0 standard uses semantic web, drag n drop mash ups, widgets, user behavior or Me-onomy, advertisement, user engagement, consolidates dynamic content and focuses on individuals.

The proposed expansion of the web 3.0 technology includes the following:

- Mobile devices, Ubiquitous connectivity, mobile internet access with broadband adoption.
- Open technologies, open data formats & data license, open API's and protocols.
- Open identity, open reputation, personal data and roaming portable identity.
- Intelligent web technologies, semantic techniques like OWL, RDF, SPARQL, SWRL, and statement based data stores with semantic application platforms.
- Intelligent applications, machine learning, autonomous agents, machine reasoning and natural language processing.

Key words: Web 2.0, Web 3.0, web technologies, search engines, microformats, RDFa, metadata, integration

1. Uvod

1.1 Kaj je Splet?

World Wide Web, kratica WWW in splošno znan kot Splet, je sistem medsebojno povezanih dokumentov hiperteksta dostopnih prek interneta. S spletnim brskalnikom je mogoče pregledovanje spletnih strani, ki lahko vsebujejo besedilo, slike, video posnetke in druge multimedije, ter navigacijo med njimi prek hiperpovezav. Izraza Internet in Svetovni splet se pogosto uporabljata v vsakdanjem govoru brez velike razlike. Vendar Internet in Splet nista eno in isto. Internet je globalni sistem medsebojno povezanih računalniških omrežij. V nasprotju s tem, je Splet ena od storitev, ki poteka na internetu. To je zbirka med seboj povezanih dokumentov in drugih virov, povezanih s hiperpovezavami in URL-ji¹. Skratka, Splet je aplikacija, ki teče na Internetu in je nedvomno postala ena od najpomembnejših konceptov 21. stoletja. Danes je večina medijev, vključno s časopisi, televizijo, radijem, zabavo, glasbo ali komunikacijskimi sredstvi, kot so elektronska pošta, chat ali telefon, postala del spleta. Razvoj Splet-a je tesno povezan s pojmom Internet in ima vsak dan vse večjo priljubljenost.

1.2 Kako deluje Splet?

Delovanje Spleta sloni na treh standardih[1]:

- URL – določa za vsak dokument v Spletu enolični »naslov«, na katerem ga lahko najdemo
- HTTP² – določa način, kako se sporazumevata spletni strežnik in brskalnik
- HTML³ – določa skladno označevanje metabesedilnih elementov (naslovov, slik itd., predvsem pa povezav na druge dokumente) v besedilu

Pregled spletne strani na svetovnem Spletu se običajno začne bodisi tako, da vnesemo URL strani v spletni brskalnik ali pa sledimo hiperpovezavo na to stran ali vir. Spletni brskalnik nato sproži vrsto komunikacijskih sporočil, v zakulisju, da jih ujame in prikaže. Najprej

¹ Uniform Resource Locator - enolični kazalec virov

² Hyper-Text Transfer Protocol - protokol za prenos hiperteksta

³ Hyper-Text Markup Language - hipertekstovni označevalni jezik

brskalnik prevaja strežnikov del URL-ja v naslov IP⁴ z globalno, porazdeljeno Internetno podatkovno zbirko DNS⁵. Ta IP naslov mora stopiti v stik s spletnim strežnikom. Brskalnik nato zahteva vir, ki ga pošlje z zahtevo HTTP na spletni strežnik, pri tem določenem naslovu. V primeru tipične spletne strani je prva zahteva besedilo HTML in je takoj razčlenjeno s spletnim brskalnikom, ki naredi potem dodatne zahteve za slike in druge datoteke, ki izpolnjujejo celotno stran. Medtem ko se sprejemajo datoteke iz spletnega strežnika, lahko brskalnik postopoma postaja stran na zaslonu, taka kot je določena s svojim HTML, CSS⁶ ali drugim sestavnim jezikom te strani. Vse slike in drugi viri so pripravljeni na zaslonu spletne strani, ki je uporabnik vidi. Večina spletnih strani vsebuje povezave do drugih strani in sicer do datotek, izvornih dokumentov, definicij in drugih spletnih virov. Taka zbirka uporabnih, povezanih virov, med seboj povezanih preko hipertekstovnih povezav, je poimenovana Splet informacij.

1.3 Razvoj Splet-a

Spletni konzorcij (W3C) je ustanovil Tim Berners-Lee oktobra 1994. Splet omogoča širjenje informacij prek interneta z enostavno uporabo in prilagodljivim formatom, ter se od svojega nastanka dalje močno spreminja. Razvoj novih tehnologij omogoča nove aplikacije. Na podlagi teh tehnologij lahko njegov razvoj razdelimo na tri stopnje: Web 1.0, Web 2.0, v zadnjem času pa se čedalje več govori o Web 3.0. Gre za obdobje prevladovanja treh sklopov različnih tehnologij, ki se med sabo časovno močno prepletajo.

1. Web 1.0

Obdobje Web 1.0 se je začelo na začetku devetdesetih let, ko je internet postal dostopen širši javnosti. V tem obdobju so bila orodja in tehnologija za objavljanje vsebin na spletu draga, za tovrstne operacije pa so bili potrebni strokovnjaki. Zaradi tega so v začetkih interneta na spletu objavljala podjetja, vlade in univerze. Tipičen primer Web 1.0 so novice o različnih dogodkih napisane s strani profesionalnih novinarjev.

⁴ Internet protocol - številka, ki natančno določa računalnik v omrežju Internet

⁵ Domain name server - distributivno ponujeno ime za omrežja s prijaznim imenom za računalnike in uporabnike

⁶ Cascading Style Sheets - so predloge, ki določajo izgled spletnih strani

2. Web 2.0

Web 2.0 se je začel pojavljati v začetku tega stoletja. Omogočilo ga je padanje cen objavljavanja na spletu in poenostavitev postopkov objave. Tako je lahko praktično vsakdo postal avtor vsebin na spletu. Značilnost Web 2.0 je participacija, sodelovanje pri tvorjenju različnih vsebin s strani samih uporabnikov interneta. Najpomembnejši mejniki tega obdobja so družabne mreže, blogi, '*peer to peer*' sistemi za izmenjavo vsebin, in podobno. Tipičen primer tega obdobja je poročanje o različnih dogodkih na blogih, družabnih mrežah in podobnih kanalih sporočanja.

3. Web 3.0

Web 3.0 tehnologije dajejo pomen medsebojni komunikaciji različnih naprav, ki uporabniku omogočajo dostop do informacij, oziroma posredovanje le-teh s čim manj truda, na različne prenosne naprave z dostopom do interneta. Web 3.0 se še ni čisto zakoreninil, vendar lahko pričakujemo hitrejši razvoj tovrstnih tehnologij v prihodnjih letih. Web 3.0 na nek način združuje vsebine Web 1.0 in 2.0. Informacije, ki jih ponuja, so hkrati na lokalni in globalni ravni, bistvo je, da so zanimive za uporabnika spleta. Gre za predstavitev dogodkov na klasičen način (novinarsko poročanje), predstavitev s strani uporabnikov (blogi), hkrati pa je dodana še neka širša dimenzija obveščanja, ki zajame čim večji krog ljudi.

2. Web 2.0

Web 2.0 temelji na "interaktivni" uporabi spleta. Ljudje lahko izvajajo interakcije z drugimi in si lahko med seboj pošiljajo misli in ustvarjalne izraze, prav tako pa obstaja prostor za druge, da pustijo svoje komentarje. Če je 21. stoletje razvoj človeških virov in odnosov, medsebojno kot tudi mednarodno, Web 2.0 močno prispeva k temu razvoju. Globalni projekti, sodelovanja in interakcije so na voljo in so dokaj enostavni - čeprav je lahko časovni pas dejavnikovih interakcij "v živo" zastrašujoč. Na srečo na spletu obstajajo email-i, blog-i, tweets-i, itd., drugi pa jih lahko najdejo takrat kadar hočejo.

2.1 Definicija

Definicija pojma Web 2.0 se glasi[2]:

„Filozofija vzajemnega zvišanja kolektivne inteligence in dodane vrednosti za vsakega udeleženca z dinamičnim ustvarjanjem in deljenjem informacij.“

Web 2.0 lekcija: „Urediti stvari tako, da samousluga kupca in algoritmsko upravljanje s podatki doseže celotni splet, vse do robov, ne samo do centra, do dolgega repa, ne samo do glave.“

Sam pojem postane značajan po prvi Web 2.0 Konferenci O'Reilly Media⁷ leta 2004. Ustanovitelj podjetja Tim O'Reilly pojem definira kot: „Web 2.0 je poslovna revolucija v računalniški industriji, ki je nastala s tretiranjem mreže kot platforme in s potrebo, da se dojamejo pravila uspeha na tej novi platformi“ [3].

2.2 Razvoj

Web 2.0 pred vsem označuje novo in izboljšano generacijo Splet-a, čigar razvoj se je začel v 90-ih letih. Pojem »Web 2.0« je prvič uporabila leta 1999 Darcy DiNucci. Njena uporaba tega izraza se največ ukvarja s spletnim dizajnom, estetiko in povezovanjem vsakdanjih objektov z Internetom. Zatem pojem sploh ni bil omenjen do leta 2002, ko sta Scott Dietzen in John Robb začela delati na konceptih, ki so bili v tem obdobju usklajeni s pojmom Web 2.0. V letu 2003 je začela rasti priljubljenost pojma, ko sta O'Reilly Media in

⁷ O'Reilly Media - <http://oreilly.com/>

MediaLive gostila prvo Web 2.0 konferenco. V svojem uvodnem govoru sta John Battelle in Tim O'Reilly predstavila svojo opredelitev "Splet kot platforma", kjer so programske aplikacije zgrajene tako, da se izvajajo na Splet-u in ne na namizju. Za tem se je pojem spet začel uporabljati leta 2004, ko sta si duo Tim O'Reilly in Dale Dougherty izmislila ime za spremembe, ki so se začele dogajati na spletu. Ni se zgodilo nič tehnološko revolucionarnega, vendar je bilo potrebno ime za konceptualne spremembe, ki so se začele dogajati pri oblikovanju spletnih strani. Glavna sprememba je bila v njihovi interakciji - obiskovalci niso bili več le bralci, ki se jim servira gotova in nespremenljiva informacija, ampak so postali sodelavci in jim je bila dana možnost, da sodelujejo pri ustvarjanju vsebine ter so tako lahko delili svoje mnenje in vsebino z drugimi obiskovalci. Vloga lastnika spletne strani je na ta način zmanjšana in teko se je delo urejanja vsebine preneslo na končne uporabnike. Ker se je število obiskovalcev in sodelavcev začelo višati, so spletne strani postajale vse bolj zanimive in uporabne. O'Reilly-jeve Web 2.0 konference so potekale vsako leto od leta 2003 in so pritegnile podjetnike, velika podjetja in tehnološke novinarje.

2.3 Značilnosti

Web 2.0 je zmnožek spletnih tehnologij in inovacij uporabe, ki dopolnjujejo dotedanjo uporabo in razumevanje spleta in njegovega ozadja, ki ga pogosto predstavljajo kot poslovno revolucijo s premikom k Internetu, kot osnovi za uspeh. Bistvo Web-a 2.0 je sodelovanje uporabnikov Interneta pri dopolnjevanju obstoječih in oblikovanju novih vsebin, medtem ko so uporabniki na Web-u 1.0 uporabljali le obstoječe vsebine. Posameznik prispeva k vsebini, zato nudi Web 2.0 informacije vseh možnih tem in področij. Sistem prinaša kategorizacijo in urejenost vsebine s povezano arhitekturo. Sodelovanje pa hkrati zahteva vedno hitrejša povezave in izboljšane tehnologije. Podatki na Web 2.0 se pojavljajo v obliki in na lokaciji po želji uporabnika, ki jih lahko poljubno uredi ali pa preuredi v popolnoma drugačen kontekst, v povezavi z drugimi informacijami. Za primer navedimo kuharski blog novih receptov in ocene restavracij. Primer: Avtor lahko ob stran bloga doda povezave do produktov, ki bi morda zanimali bralce bloga, kot je npr. nov kuhinjski lonec Zepter. Avtor bloga bo pri tem služil od števila klikov drugih obiskovalcev na ponudbe. Pri Web 2.0 so spletne vsebine pod slabim nadzorom lastnikov spleta in bližje Tim Berners-Lee-jevi zamisli spleta, ki naj bo demokratičen, osebni (angl.: Do-It-Yourself) medij komunikacije. Tok vsebin ne poteka več izključno po e-pošti, ampak preko različnih družbenih in osebnih mrež. Web 2.0 predstavlja

odmik od statičnih spletnih strani in z uporabo iskalnikov stremi k dinamičnemu, interaktivnemu svetovnemu sodelovanju na spletu.

Osnovne značilnosti koncepta Web 2.0, od katerih je potrebno posebej izpostaviti naslednje:

- Internet kot platforma – končni uporabniki uporabljajo aplikacijo v celoti s pomočjo spletnega brskalnika
- uporabniki ustvarjajo vsebine in imajo nadzor nad njimi
- vsebine pripadajo uporabnikom
- družbena komponenta
- izboljššan grafični vmesnik

Ena od glavnih prednosti je možnost uporabnika, da ustvarja vsebino. Vendar pa ustvarjanje vsebin vključuje tudi objavo, zato se je razvila široka paleta spletnih platform, ki se uporablja za njihovo distribucijo. Večina od njih se poslužuje storitve hostinga, ki uporabnikom omogoča zadostno količino spletnega prostora za objavo svojih del. Vsaka oblika vsebine je običajno povezana z ustrezno distribucijsko platformo, tako da se na primer besedilne vsebine pogosto pojavljajo v blogih in podobno. Najbolj priljubljene distribucijske platforme so:

- blogi,
- wiki sistemi,
- podcast-i
- družabna omrežja in
- druge podobne storitve.

2.3.1 Sestavni deli Web 2.0

Socialno mreženje je postalo sinonim za pojem Web 2.0. Kar pomeni aktivno sodelovanje v virtualnih skupnostih, t.j. skupine uporabnikov enakih potreb zbranih okoli ene Internet storitve (blogi, forumi, itd.). Eni od najbolj priljubljenih družbenih spletov (družbeni programi) so Facebook, Twitter, MySpace in še veliko drugih.

Blog je izraz, ki se nanaša na osebni dnevnik napisan na spletu v obratnem kronološkem vrstnem redu. Na pojem blog se navezuje blogosfera - skupnost uporabnikov interneta, ki sodelujejo pri ustvarjanju blogov, blogger - avtor bloga in bloganje - pogosto pisanje bloga in

komentiranje drugih blogov. V blogih in mobilni blogih so pisanje in branje urejeni preko mobilnih telefonov, ročnih računalnikov in podobnega.

Pomembno mesto v socialni interakciji zasedajo tudi forumi (javno razpravljanje o posameznih temah prek interneta), kot tudi takojšnje sporočanje in klepet (lastnosti - izmenjava sporočil v realnem času).

Folksonomija ali kolaborativno označevanje je kolaboracijska kategorizacija vsebin z uporabo oznak (ključnih besed, da bi se opisal vaš blog, profil, spletna stran itd.) Najbolj znan tako kategoriziran splet namenjen za splošno enciklopedijo je Wikipedija.

Flickr.com je kombinacija internetnih storitev za objavo digitalnih fotografij in spletne družbe.

YouTube je podobna storitev za objavo, pregledovanje in izmenjavo ter komentiranje video posnetkov.

2.4 Tehnologije



Slika 1: Prikaz sestavin Web 2.0

Tehnologije, ki se uporabljajo s spletnim brskalnikom (na strani odjemalca) pri Web 2.0 so Asinhroni JavaScript⁸ in XML⁹(Ajax), Adobe Flash in Adobe Flex¹⁰ ogrodje ter

⁸ JavaScript je objektni skriptni programski jezik

⁹ Extensible Markup Language - je preprost računalniški jezik podoben HTML-ju, ki nam omogoča format za

Javascript/Ajax ogrodja. Podatki, ki so določeni na zahtevo Ajax-a so ponavadi oblikovani v XML ali JSON¹¹ formatu, dveh pogosteje uporabljenih formatih strukturiranih podatkov. Ker sta oba formata avtomatično razumljiva s strani JavaScript-a, jih lahko programerji enostavno uporabljajo za prenos strukturiranih podatkov v spletno aplikacijo. Ko so ti podatki prejeti prek Ajax-a, JavaScript potem uporablja Document Object Model (DOM) za dinamično posodobitev spletne strani na podlagi novih podatkov, ki omogočajo hitro in interaktivno uporabniško izkušnjo. Skratka, z uporabo teh tehnik lahko spletni oblikovalci naredijo, da njihove strani delujejo kot namizne aplikacije. Adobe Flex je še ena tehnologija, ki se pogosto uporablja pri Web 2.0 aplikacijah. V primerjavi z JavaScript knjižnicami, kot je jQuery, Flex omogoča programerjem, da lažje zapolnijo velike mreže podatkov, grafikone in druge težke interakcije uporabnikov. Aplikacije programirane v Flex-u so prevedene in prikazane kot tehnologija Flash v brskalniku. Kot splošno dostopen plugin neodvisen od W3C standardov, lahko Flash počne veliko stvari, ki niso bile mogoče pred HTML 5 jezikom, ki se uporablja za izdelavo spletnih strani. Glede na številne možnosti Flash-a, je najbolj pogosto uporabljena v Web 2.0 njegova sposobnost za predvajanje avdio in video datotek. To je omogočilo oblikovanje Web 2.0 strani, kjer je video medij popolnoma povezan s standardom HTML. Poleg Flash-a in Ajax-a, so postala JavaScript/Ajax ogrodja v zadnjem času zelo priljubljeno sredstvo pri ustvarjanju Web 2.0 strani. Pri njihovi osnovi, ta ogrodja ne uporabljajo tehnologije drugačne od JavaScript-a, Ajax-a, in DOM-a. To kar ogrodja storijo je nemoteno čez neskladnosti med spletnimi brskalniki in razširitvijo funkcionalnosti, ki je na voljo razvijalcem. Veliko njih prihaja tudi s spremenljivimi, montažnimi "widgeti", ki izpolnjujejo skupne naloge, kot določanje datuma na koledarju, prikazovanje podatkov grafikona, ali urejanje plošče z zavihki.

Na strani strežnika uporablja Web 2.0 številne enake tehnologije kot Web 1.0. Nove programske jezike, kot so PHP, Ruby, ColdFusion, Perl, Python, JSP in ASP uporabljajo razvijalci za dinamičen izhod podatkov z uporabo informacij iz datotek in podatkovnih baz. Kar pa se je začelo spreminjati v Web 2.0, je način kako se ti podatki uporabljajo. V zgodnjem razvoju interneta ni bilo velike potrebe po različnih spletnih straneh za komunikacijo med seboj in izmenjavo podatkov. V novem "participativnem spletu" je izmenjava podatkov med lokacijami postala bistveno bolj izvedljiva. Če želite deliti svoje

opisovanje strukturiranih podatkov ali arhitektura za prenos podatkov in njihovo izmenjavo med več omrežji
¹⁰ Adobe Flash in Adobe Flex - zasnovani na odprtih tehnologijah ter sodelovanju skupnosti in prinašajo inovacije v brskalniki na računalnikih in mobilnih napravah

¹¹ JavaScript Object Notation - lahki format podatkov za izmenjavo

podatke z drugimi spletnimi stranmi, mora spletna stran imeti možnost, da ustvarja izhodno stran v strojno berljivi obliki, kot so XML, RSS¹² ali JSON. Ko so podatki ene strani na voljo v eni od teh oblik, jo lahko druga spletna stran uporabi tako, da vključi del funkcionalnosti tiste spletne strani pri sebi na način, da naredi medsebojno hiperpovezavo. Ko se ta vzorec oblikovanja uveljavi, prihaja do podatkov, ki jih je tako lažje najti in bolj natančno opredeliti, kot znak filozofije operacije Web 2.0. Iz tega sledi da je AJAX ključna tehnologija za izgradnjo Web 2.0, saj zagotavlja bogato uporabniško izkušnjo in deluje z vsemi brskalniki, ali je to Firefox, Internet Explorer ali katerikoli drugi priljubljen brskalniki. Potem je treba uporabiti jezik z zelo dobro podporo spletnih storitev, da se zgradi Web 2.0 aplikacija. Poleg tega je treba uporabiti jezik, ki bo iterativnega pomena, da lahko pomaga, da se enostavno in hitro dodajajo in uporabljajo funkcije.

2.4.1 Principi Web 2.0 tehnologije

Tim O'Reill je prvi poudaril šest idej, ki pomagajo razumeti močan vpliv Web 2.0 tehnologije. To so ideje ustvarjanja, več kot zmore le globalni informacijski prostor, in imajo družabni poudarek. Sodelovanje, prispevanje in skupna last so osnova Web 2.0. Imamo občutek, kot da se pred nami gradi nova struktura družbe. Ideje so neposreden ali posreden odziv moči omrežja, katerega posledica topologije na mikro in makro ravneh ustvarja milijarde internetnih uporabnikov[4].

Individualna in uporabniško ustvarjena vsebina

S klikom na miško naloži uporabnik na svoj spletni prostor video, tekst, sliko ali film, ga tekstovno etiketira - tagira, in ga izmenjuje s prijatelji na spletu. Veliko ljudi piše bloga ali ustvarja nove informacije s pomočjo wiki-jev. Ta orodja so tako odstranila oviro vstopa na internet. Mediji so začeli uporabljati uporabniško ustvarjeno vsebino v svojih prispevkih. Uporabniki interneta so s prispevanimi gradivi za tv mreže, časopise in spletne strani postali soustvarjalci informacij. Zaradi vse cenejših in kakovostnejših video kamer postajajo uporabniki spleta novinarji na kraju dogodka.

¹² RSS - je družina XML datotečnih oblik za spletno zlaganje, ki ga uporabljajo spletne strani, ki nudijo novice, ter tako imenovani spletni dnevniki

Izkoriščanja znanja množic

Modrost množic predvideva hipotezo o zbiranju informacij v skupinah, ker je v mnogih primerih skupinska ocena natančnejša od odločitve strokovnjaka posameznika. Primer uporabe modrosti množice na Web 2.0 je programsko orodje Cloudmark¹³, ki deluje kot filter nezaželenih pošte in zbira podatke uporabnikov e-pošte in njihove odločitve, kot nezaželeno ali zaželeno. Na podlagi teh podatkov Cloudmark potem ugotavlja, v katero kategorijo spada prispela pošta. Sistem je doslednejši od osebne ocene vsebine sporočila. V isti sklop sodi tudi Crowdsourcing. To je spletno obveščanje izbranih multimedijskih vsebin drugim uporabnikom (angl.: outsourcing). Primer so multimedijske spletne strani Youtube, Flickr in podobne, ki so uporabniško ustvarjene vsebine za uporabo plačila.

Podatki v ogromnih količinah

Z informacijsko tehnologijo ustvarjamo in uporabljamo trajno naraščajočo količino informacij. Za podjetja, kot je npr. Google, je pomembnost podatkov neprecenljiva. Ključni del poslovanja teh podjetij je upravljanje in urejanje ogromnih količin podatkov v podatkovnih bazah. Google meri obsežnost podatkov svoje podatkovne baze v nekaj sto petabajtih, ki se vsak dan poveča za terabajt novih. Brez podjetij, kot je Google, si življenja ne predstavljamo več, njihove storitve nam lajšajo življenje. Pojavi pa se vprašanje lastnine in varnosti podatkov. Podatki so vir sredstev, ki se lahko ponovno uporabijo, pri tem pa nastopi še vprašanje zasebnosti.

Arhitektura sodelovanja

Je ozadje, ki izboljša aplikacijo oziroma storitev. Sistem zajame sodelovanje uporabnika in izkoristi izkušnje izboljšanja sistema; tak primer je Google search. Netipični predstavnik Web 2.0 je program BitTorrent¹⁴, ki ima decentralizirano delovanje. Pri njem ugotovimo, da prav mreža uporabnikov - nalagalnikov naredi sistem tako učinkovit, ker nudi vsak uporabnik svojo pasovno širino in podatke drugim uporabnikom. Več ko je uporabnikov, več virov in sredstev je na voljo. BitTorrent potrjuje ključno Web 2.0 ugotovitev: več kot je uporabnikov, boljša je storitev.

¹³ Cloudmark - <http://www.cloudmark.com/>

¹⁴ BitTorrent - <http://www.bittorrent.com/>

Učinek omrežja

Obstajata dva koncepta vključevanja Web 2.0 tehnologije. Prvi je učinek omrežja, ki temelji na obsežnosti omrežja, ali ekonomični in socialni vpletenosti novih uporabnikov storitev na spletu. Drug koncept je moč zakonodaje in njena vključenost v omrežje. Oboje nas popelje v razpravo o dolgem repu. Osredotočili se bomo le na učinek omrežja. Učinek omrežja je gospodarski izraz, ki opisuje povečano korist obstoječemu uporabniku storitve, ko se poveča število novih uporabnikov zaradi možnega medsebojnega delovanja - interakcije. Več kot je uporabnikov, večji je učinek omrežja. Izraz se največkrat uporablja pri opisovanju obsega povečane uporabnosti telefonskega omrežja, s povečanim številom uporabnikov. Za primer navedimo vključitev novega uporabnika v telefonsko omrežje. Z vključitvijo ne pridobi le nov uporabnik, ampak imajo neposredno korist že obstoječi uporabniki, ki lahko sedaj pokličejo nekoga, ki ga prej niso mogli. Obstaja očitna povezava z razvojem družabne tehnologije, kot je npr. MySpace. Ko se pridruži nov uporabnik k družabni mreži, imajo korist od te pridružitve že obstoječi uporabniki. Ko učinek omrežja narašča in postanejo uporabniki pozorni na povečan obseg storitev, se uporaba storitev na trgu neizmerno poveča.

Odprtost

Vzporedno z razvojem spleta so se razvila tudi pravna, politična in kulturna pravila za pravice dostopa do digitalnih vsebin. Vendar velja na spletu tradicija sodelovanja na odprtosti, ki jo neizprosno podpira Web 2.0 tehnologija. Delo z odprtimi standardi, uporaba odprtokodnih programskih orodij, izkoriščanje prosto dostopne vsebine, ponovna uporaba vsebine in delovanje v namenu odprte inovacije so lastnosti, ki se neposredno nanašajo na Web 2.0. Pomembna tehnologija pri razvoju Web 2.0 je odprta koda brskalnika Firefox in njen sistem raztegljivih dodatkov, ki dovoljujejo eksperimentiranje. Web 2.0 poudarja uporabo informacij v veliki podatkovni bazi, ki jo polnijo razne storitve.

3. Web 3.0

Ker spletna tehnologija še ni prišla na svoje mesto, nas Web 3.0 uvaja v nekatere poglede, da je "persona" spletnih brskalnikov. Internet Explorer, Firefox, itd. bodo sposobni slediti in sintetizirati informacije o zgodovini uporabnikov, ter iz njih predstaviti predloge. Tudi z boljšo sintezo podatkov, lahko oseba vtipka vprašanje samo v dveh ali treh besedah in pri iskanju se poišče pomensko vse skupaj. Pri tem, Web 3.0 ne vpliva neposredno na interakcije med ljudmi toliko, kot na "uporabo" zmogljivosti računalnika za razumevanje. Na nek način vodi Web 3.0 nazaj proti Web 1.0 in omogoča uporabniku bolj učinkovit in uspešen način uporabe interneta. Internet skupnost se kot celota nagiba k tem, da opredeli dva izraza "Semantičen splet" in "Web 3.0", ki sta sinonima v konceptu, če nista v celoti drugačna. Definicija je še vedno odvisna od tega o katerem se govori. Veliko soglasje je, da je Web 3.0 gotovo "naslednja velika stvar", vendar pa obstajajo samo špekulacije, le kaj bi to bilo. To bo izboljšanje v zvezi s tem, da še vedno vsebuje Web 2.0 lastnosti, medtem ko naprej prispeva k svojemu vedno širšemu leksikonu in knjižnici vlog. Obstajajo nekateri, ki trdijo, da bo Web 3.0 bolj temeljen na aplikacijah ter, da bo usmeril središče dela na doseganje bolj grafično sposobnih okolj in na še bolj primerno uporabo in razvoj umetne inteligence. Na primer, Conrad Wolfram trdi, da Web 3.0 izgleda tako, kot da "ustvarja računalnik nove informacije" namesto da bi jih ljudje [5]. Nekateri preprosto izražajo svoje prepričanje, da bo Web 3.0 v prvi vrsti osredotočen na bistveno izboljšanje funkcionalnosti in uporabnosti iskalnikov. Obstaja pomemben dejavnik, ki ga morajo imeti uporabniki še vedno v mislih in sicer, da je prehod na Web 2.0 iz Web 1.0 trajal približno deset let. Glede na enako časovno obdobje ta naslednji prehod razvoja ne bo končan do približno leta 2015.

Sir Tim Berners-Lee, soizumitelj svetovnega spleta, direktor W3C konzorcija in oče semantičnega spleta[6]:

»Sanjam o spletu, v katerem bodo računalniki sposobni analizirati vse podatke, ki krožijo po internetu: vsebino, povezave in interakcije med ljudmi in računalniki. Semantični splet, ki bi to omogočal, še mora nastati, a ko bo prišel na dan, bodo v vsakodnevnem življenju, prometu, birokraciji, stroji komunicirali s stroji.«

3.1 Definicija

Definicija Web 3.0 ali Semantičnega spleta se glasi[7]:

Semantični splet je "mreža podatkov", ki omogoča strojem razumeti semantiko, ali pomen informacij na svetovnem spletu. Razširja mreže hiperpovezav berljivih spletnih strani z vstavitvijo strojno berljivih metapodatkov o strani in kako so povezane med seboj, kar omogoča avtomatskim agentom dostop do spleta bolj inteligentno in izvajanje nalog v imenu uporabnikov.

Web 3.0 (Semantični splet) je baza podatkov oz. podatkovni svetovni splet. Podatki so distribuirani, razdrobljeni, vseprisotni na spletu. To je splet za ljudi in končno tudi za inteligentne stroje (agente), ki znajo te podatke razumeti, analizirati in ponuditi ljudem.

3.2 Razvoj

Izraz Web 3.0 se je prvič pojavil v začetku leta 2006 v članku Jeffrey Zeldman-a v kontekstu kritik do Web 2.0 in z njim povezanih tehnologij, kot je AJAX.

V maju 2006, [Tim Berners-Lee] je rekel[8]:

"Ljudje se sprašujejo, kaj je Web 3.0? Mislim, ko imaš prosojnico z raztegljivo vektorsko grafiko SVG¹⁵ – vse žubori in je prepogibno in zgleda megleno – na Web 2.0 in dostop do semantičnega spleta je integriran v ogromen prostor podatkov, imeli boste dostop do neverjetnega vira podatkov ." (Tim Berners-Lee, "Web" bolj revolucionarno")

Na vrhunskem srečanju Technet v novembru 2006, je Jerry Yang, ustanovitelj in predsednik Yahoo-a, povedal naslednje: [8]:

"Web 2.0 je zelo dokumentiran in obravnavan. Pristojnost Interneta je dosegla kritično maso s potencialom razvijalcev interneta, ne samo v strojni opremi, kot so igralne konzole in mobilne naprave, ampak tudi z arhitekturo programske opreme. Vam ni treba biti ne vem kdo, da ustvarite računalniški program. Opažamo kaj se dogaja v Web 2.0 in Web 3.0 bo veliko nadaljevanje tega, dejansko pomeni skupno... razlika med profesionalcem, polprofesionalcem in potrošnikom bo ustvarila nejasen učinek na poslovne aplikacije in omrežja."

Na istem Technet vrhunskem srečanju, je Reed Hastings, ustanovitelj in direktor podjetja

¹⁵ Scalable Vector Graphics - je označevalni jezik za opis dvorazsežne statične in risane vektorske grafike v XML

Netflix, povzel preprosto enačbo, ki določa različne faze Web-a: [8]:

"Web 1.0 je bil dial-up 50Kb povprečne pasovne širine, Web 2.0 je povprečno 1 megabit pasovne širine in Web 3.0 bo 10 megabitov pasovne širine ves čas, ter bo na ta način popoln video Web in tako se bo čutil Web 3.0"

Avgusta 2007 je brazilska digitalna agencija CUBE opredelila Web 3.0 kot možnost strankam, da komunicirajo s podjetji, bodisi neposredno z uporabo spletnih dnevnikov in drugih Web 2.0 aplikacij, ali posredno, kot da bi mi bili imetniki psihografskih podatkov analiziranih s semantičnim spletom in drugimi orodji za trženje, kot je microtargeting/Silent Marketing. V letu 2009 vidimo že več proizvodov, ki temeljijo na odprtih, strukturiranih podatkih, kot je na primer WolframAlpha¹⁶. Izraz Web 3.0 je postal in postaja predmet večjega zanimanja in razprave od konca leta 2006 do danes.

3.3 Značilnosti

Veliko je razprav o pomenu Web 3.0 in kaj bi bila njegova ustrezna opredelitev. Web 3.0 ali semantični splet ni ločen splet, ampak nadgradnja obstoječega. Semantičen splet ne opredeljuje samo spletne strani in povezave, temveč odnose med podatki. To je skupek tehnoloških standardov, ki omogoča strojem razumeti semantične dokumente in podatke. Glavni namen Semantičnega spleta, ki spodbuja razvoj sedanjega spleta, je podati možnost uporabnikom, da ga uporabljajo za ves svoj potencial, ter da jim omogoča lažje iskanje, izmenjavo in združevanje podatkov. Ljudje imamo sposobnost spletnega izvajanja nalog, kot so npr. iskanje tuje besede za "mapo", rezervacija knjige v knjižnici ali iskanje nizke cene DVD-ja¹⁷. Vendar stroji ne morejo izpolniti vseh teh nalog brez, da bi jih ljudi usmerili, saj so spletne strani namenjene branju ljudem, ne pa strojem. Semantični splet je vizija informacij, ki jih je mogoče razložiti s stroji tako, da lahko stroji opravljajo več mučnega dela pri iskanju, ki združuje in deluje na podlagi informacij na spletu.

Semantični splet predstavlja rešitev na višjem nivoju. Gre za predstavitev v jezikih, ki so izdelani posebej za podatke: Resource Description Framework¹⁸, Web Ontology Language¹⁹ in Extensible Markup Language (v nadaljevanju XML). HTML opisuje dokumente in povezave med njimi. RDF, OWL in XML v nasprotju, lahko opišejo poljubno stvari, kot so ljudje,

¹⁶ WolframAlpha - <http://www.wolframalpha.com/>

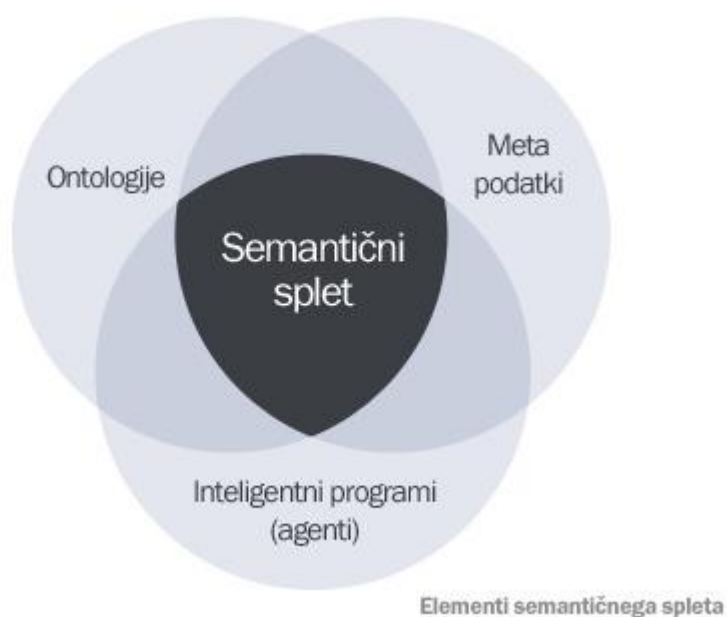
¹⁷ Digital Versatile Disc - je digitalni pomnilniški medij

¹⁸ RDF - je družina W3C specifikacij prvotno zasnovana kot metapodatkovni model

¹⁹ OWL - je družina jezikov za predstavitev znanja in avtorskih ontologij

sestanki ali dele letala[7]. Te tehnologije so združene z namenom, da zagotovijo opise, ki dopolnjujejo ali nadomeščajo vsebine spletnih dokumentov. Tako se lahko vsebina odčituje samo kot opis podatkov, shranjen v spletu dostopnih baz podatkov, ali kot markup v dokumentih (predvsem v XHTML²⁰ mjestimice z XML, ali pogosteje zgolj v XML s postavitvijo in onemogočanjem prikazov, ki se skladiščijo ločeno). Strojno-berljivi opisi omogočajo vsebinskim menedžerjem dodati pomen vsebini, torej da opišejo strukturo znanja, ki jo imamo o tej vsebini. Na ta način lahko stroj samostojno procesira znanje, namesto besedila, z uporabo postopkov, ki so podobni človeški deduktivni obrazložitvi in sklepanju tako, da se pridobi več smiselnih rezultatov in pomaga računalnikom opraviti avtomatsko zbiranje podatkov in raziskave.

Srce semantičnega spleta predstavlja tehnologija, ki omogoča uporabnikom iskanje zelenih informacij ter korelacij med njimi, ne glede na to, kje na spletu se iskani podatek nahaja. Gre za pomensko in ne sintakso označevanje in povezovanje podatkov, s čemer lahko postanejo le-ti uporabni za strojno obdelavo.



Slika 2: Elementi Semantičnega spleta

Elementi[6]:

1. Ontologije so slovarji znanja o določenih področjih. Vsebujejo definicije pojmov: razrede, lastnosti, odnose med njimi; ustvarjajo jih lahko vsi in objavijo na spletu.

²⁰ Extensible HTML - je označevalni jezik, ki ima enak namen kot HTML, vendar je usklajen s sintakso XML

Primer ontologije:

Razred LCD_zaslon, vsebuje lastnosti:

- diagonala
- cena
- prodajalec (tip: Trgovec)

Razred Trgovec, vsebuje lastnosti:

- naziv
- telefonska številka

2. Metapodatki so podatki, ki opisujejo druge podatke, so reference do pojmov, ki so definirani v ontologijah in tako ustvarjajo globalni pomen meta-podatkov.

Primer:

Moj_LCD_zaslon je tipa LCD_zaslon in ima naslednje vrednosti:

- diagonala: 24"
- cena: 800 EUR
- prodajalec: Matrix_Trade

Matrix_Trade je tipa Trgovec in ima naslednje vrednosti:

- naziv: Matrix Trade, imaginarna družba za prodajo, d.o.o.
- telefonska številka: +38615550011

3. Inteligentni programi (agenti) analizirajo metapodatke in uporabljajo ontologije, ki so dostopne na spletu, da bi te metapodatke razumeli.

Primer inteligentnega iskalnika:

Uporabnikov zahtevek: Najdi mi telefonsko številko med vsemi Trgovci, ki ponujajo LCD_zaslon, ki ima diagonalo 24" in ceno manjšo kot 1.000 EUR.

Odgovor iskalnika: +38625550011

3.4 Tehnologije

Tehnologije Semantičnega spleta omogočajo ljudem, da ustvarjajo shrambe podatkov na spletu, gradijo slovarje in pišejo pravila za ravnanje z znanjem. Glavna predstavitev znanja

formalizma je ontologija. Predstavlja koncepte, lastnosti, hierarhična in ne-hierarhična razmerja med koncepti in posamezniki v strojno obdelanem načinu. Bolj formalno je ontologija prikazana kot $i = \langle C_i, T_i, R_i, I_i, A_i \rangle$, kjer je C_i množica konceptov za konceptualizacijo domene, T_i je množica taksonomskih razmerij, R_i je množica razmerij, ki niso taksonomska, I_i je množica posameznikov in A_i je množica aksiomov, ki opredeljujejo dodatna pravila in omejitve[9]. Ontologija poskuša doseči model za področje predstavitve znanja in njegove utemeljitve.

Semantični splet predstavljajo standardi in orodja XML, XML Schema, RDF, RDF Schema in OWL, ki so organizirani v Skladu semantičnega spleta. OWL Web Ontology Language Pregled opisuje delovanje in razmerje med vsako od teh sestavin semantičnega spleta:

- XML zagotavlja elementarno sintakso za strukturo vsebine v dokumentih, vendar ne povezuje nobene semantike s pomenom vsebine, ki jo vsebuje. XML danes v večini primerov ni nujna sestavina Semantičnih spletnih tehnologij, saj obstaja alternativna sintaksa, kot je npr. Turtle. Turtle je de facto standard, vendar ni šel skozi formalni postopek standardizacije.
- XML Schema je jezik za zagotavljanje in omejevanje strukture in vsebine elementov iz XML dokumentov.
- RDF je preprost jezik za izražanje podatkovnih modelov, ki se nanašajo na objekte ("sredstva") in njihova razmerja. RDF osnovni model je lahko predstavljen v sintaksi XML.
- RDF Schema razširja RDF in je slovar za opis lastnosti in razredov virov na osnovi RDF-ja s semantiko za generalizirano-hierarhijo takšnih lastnosti in razredov.
- OWL izgleda bolj kot slovar za opis lastnosti in razredov: med drugimi tudi odnosov med razredi (npr. disjunktnost), kardinalnosti (npr. "natanko en«), enakosti, bogatejši tipizaciji lastnosti, značilnosti lastnosti (npr. simetrija) in naštetim razredom.
- SPARQL je protokol in poizvedbeni jezik za semantični splet podatkovnih virov.

Trenutne trajajoče standardizacije vključujejo:

- Rule Interchange Format²¹ kot Rule Layer Stack semantičnega spleta

Plasti, ki še niso izkoriščene v celoti so:

²¹ RIF - del infrastrukture za semantični splet, skupaj z (predvsem) RDF in OWL

- Združene Logic in Proof plasti, ki doživljajo aktivno proučevanje.

Namen je povečati uporabnost in koristnost spleta in njegovih povezanih virov preko:

- Strežnikov, ki izpostavljajo obstoječe podatkovne sisteme, ki uporabljajo RDF in SPARQL standarde. Veliko pretvornikov v RDF obstaja iz različnih aplikacij. Relacijske podatkovne baze so pomemben vir. Semantični spletni strežnik se veže na obstoječi sistem, ampak ne vpliva na njegovo delovanje.
- Dokumentov "označenih" s semantično informacijo (razširitev HTML oznake <meta>, ki se uporablja v današnjih spletnih straneh za pridobivanje informacije za spletne iskalnike s pomočjo spletnih robotih). To je lahko strojno razumljiva informacija o človeško razumljivi vsebini dokumenta (npr. avtor, naslov, opis, itd. dokumenta) ali so pa lahko zgolj metapodatki, ki predstavljajo celoto dejanj (kot so sredstva in storitve nekje drugje po straneh). (Upoštevajte, da je vse, kar je mogoče opredeliti z Uniform Resource Identifier-jom²², mogoče opisati, tako da lahko semantični splet skrbi za živali, ljudi, kraje, ideje, itd.) Semantično označevanje je pogosto ustvarjeno samodejno, ne pa ročno.
- Slovarjev skupnih metapodatkov (ontologijah) in kart med slovarji, ki omogočajo ustvarjalcem dokumentov vedeti kako označiti svoje dokumente tako, da lahko agenti uporabijo informacijo v metapodatkih (tako da se Avtor v smislu "Avtor strani" ne sme zamenjati z Avtor v smislu knjige, ki je predmet recenzije knjige).
- Avtomatiziranih agentih za opravljanje nalog za uporabnike semantičnega spleta z uporabo teh podatkov.
- Spletnih storitev (pogosto s svojimi lastnimi agenti), da priskrbijo informacije še posebej agentom (na primer storitev Trust, ki bi lahko vprašala agenta, če nekatera spletna trgovina ima v preteklosti slabšo storitev ali spamming).

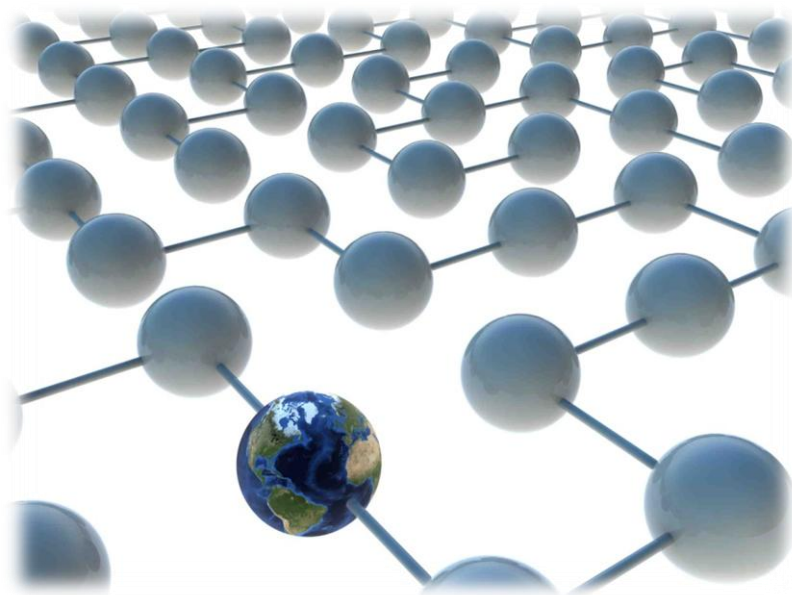
Web 3.0 tehnologije na kratko vključujejo[10]:

1. Umetno inteligenco
2. Avtomatizirano obrazložitev
3. Kognitivno arhitekturo
4. Sestavljene aplikacije
5. Porazdeljeno procesiranje

²² URI - v računalništvu, Uniform Resource Identifier je niz znakov, ki se uporabljajo za identifikacijo imena ali vira na internetu

6. Predstavitev znanja
7. Ontologijo
8. Rekombinantno besedilo
9. Razširljivo vektorsko grafiko
10. Semantični splet
11. Semantičen Wiki
12. Programske agente

4. Primerjava Web-a 2.0 z Web-om 3.0



Slika 3: Prikaz delovanja Web 3.0

Razvijajoči Web 3.0 ali Semantični splet je za razliko od spleta 1.0 in 2.0, ki sta povezani množici dokumentov, ogromna baza podatkov, distribuiranih na spletu. Govorimo o spletu podatkov. To je splet tako za ljudi (kar je bilo značilno tudi za splet 1.0 in 2.0) kot za računalnike oziroma za inteligentne agente, ki znajo podatke razumeti, analizirati in ponuditi ljudem. Semantični splet je stopnja v razvoju svetovnega spleta, v katerem je opredeljen pomen (semantika) informacij in storitev na spletu, kar mu omogoča "razumevanje" in izpolnjevanje zahtev ljudi in strojev. Na spletu imamo digitalne podatke in informacije o skoraj vsakem vidiku našega življenja.

Pri semantičnem spletu gre za dve stvari:

- za standarde, ki omogočajo integracijo in kombinacijo podatkov iz različnih virov;
- za jezik za zapisovanje povezav med podatki.

4.1 Kriteriji za primerjanje

Eden od vidikov Web-a 2.0 je poleg zanimivih novih vmesnikov in uporabe skupnih inteligenc, da potiska inteligenco in aktivne agente strežnikov do odjemalcev, natančneje brskalnikov. Razvoj aktivnih aplikacij na strani odjemalca pomeni tudi, da uporabljajo te aplikacije vse vrste podatkov. Podatke, ki so kjerkoli na spletu ali podatke, ki so vgrajeni v stran, čeprav niso nujno vidni na zaslonu. Primer je označevanje strani, kot je tip označevanja z mikroformati, koledar podatkov na spletu, označevanje slik ali povezave shranjene na

spletni strani, itd. Ta vidik Web 2.0 je takšen, da temeljijo aplikacije na kombiniranih različnih vrstah podatkov ("mashing up" podatkov), ki so razpršene po vsem spletu in se ujemajo z bistvom Semantičnega spleta. To kar Semantični splet zagotavlja, je bolj skladen model in orodje za opredelitev in uporabo kvalificiranih povezav med podatki na spletu. V bistvu se obe tehnologiji usmerjajo na inteligentno izmenjavo podatkov. Veliko število značilnih Web 2.0 demonstracij in aplikacij, ki se pojavljajo v ozadju, uporabljajo Semantična spletna orodja v povezavi z AJAX-om in drugimi zanimivi pristopi uporabniškega vmesnika. V številnih primerih se z uporabo elementov RDF tehnike naredi mashing up proces lažji. Predvsem takrat, ko so podatki zbrani z eno aplikacijo lahko ponovno uporabljeni nekje drugje na poti. Splošne značilnosti RDF-ja naredijo "mashup veriženje" enostavno, kar pa pogosto ne velja za enostavnejše Web 2.0 aplikacije. Ta dva pristopa poskušamo predstaviti kot alternativo, ali celo kot zahtevano folksonomijo, da bi zgledala višja od pristopa Semantičnega spleta. To je bila tema pogovora na blogih in na različnih publikacijah že nekaj časa, vendar pa sta obe skupnosti seznanjene s tem, da se tisti dve tehniki dopolnjujeta in nista konkurenca ena drugi. Torej, ko se lotimo primerjave teh dveh tehnologij moramo biti seznanjeni s tem, da to nista dve popolnoma različni stvari. V bistvu je ena povezana z drugo, tako da jo na neki način izboljšuje kot je tudi za pričakovati v tem svetu hitrega razvoja tehnologij. Medtem ko Web 2.0 uporablja Internet, da bi naredil povezavo med ljudmi, Web 3.0 bo uporablja Internet za povezovanje informacij.

Za primerjavo teh dveh standardov, smo izbrali dva primera v katerih bomo podrobneje predstavili razlike med Web 2.0 in Web 3.0:

1. Iskalniki
2. Mikroformati in RDFa.

Iskalnike bomo uporabili kot primer, zato ker so povsod navedeni, kot pomembna razlika med Web 2.0 in Web 3.0. Zatem bomo malo razložili kako delujejo danes in kako je planirano njihovo delovanje v prihodnosti. Mikroformati in RDFa sta tehnologiji, ki pri svoji uporabi prikazujeta razliko med Web 2.0 in Web 3.0 na enostaven način. Mikroformati so se začeli uporabljati v obdobju Web 2.0 in so skupaj z RDFa uporabljeni v prihodnjem Web 3.0.

4.2 Iskalniki danes in jutri

Sedanja generacija iskalnikov je zelo omejena v svojem razumevanju potreb uporabnikov in vsebine Spleta ter posledično pri prilagajanju na potrebe po informacijah z veliko ponudbo virov na spletu. Končni cilj semantičnega spleta oz. spleta povezanih podatkov je omogočiti

računalnikom oz. inteligentnim agentom, narediti več koristnega dela pri iskanju informacij in razviti sisteme za zaupanja vredne interakcije prek omrežja. Semantične tehnologije (kot so RDF, SPARQL, OWL...) omogočajo ljudem kreirati na spletu zbirke podatkov, zgraditi slovarje in napisati pravila za obdelavo podatkov.

Ko se je začelo obdobje Web 2.0, so spletni razvijalci odkrili pomen uporabnikov in so se odločili, da jih bodo uporabili tako, da prispevajo informacije s tega področja, torej obseg informacij se je zelo povečal. Začelo se je Wiki obdobje in blogger spletne strani so postale pravo mesto za iskanje informacij specializiranih uporabnikov. To je najpomembnejše iskanje in se je začelo tudi s pomočjo socialnega mreženja. Potem so se začeli RSS viri, ki so vstopili na »sceno«. Vse to so bili samo različni načini za uporabnika, da pridobiva informacije. Splet je postal zelo inteligenen. Če ste imeli problem in ste ga poiskali na Internet iskalniku, bi bili takoj usmerjeni na točno to stran, kjer lahko najdete informacije ali rešitev. Če ne bi nobena od spletnih strani odgovorila na vaše vprašanje, bi bil nekje forum in bi neki drugi uporabnik sedel na kakem drugem koncu sveta ter bi odgovoril na vaše vprašanje. Tako lahko rečemo, da je uvedba Web 2.0 povsem spremenila podobo Interneta. Še ena pomembna stvar je, da strani niso bile več izdelane zgolj v HTML kodi. Jeziki kot so PHP, Javascript, Ajax, itd, ki so omogočili obvladovati dinamične strani, so prišli v uporabo.

Iskalniki bodo postali bolj in bolj specifični, to je verjetno največja razlika med Web 3.0 in Web 2.0. Zdaj, ko iščemo nekaj na internetu, je informacija, ki se nam prikaže, ista kot pri drugem uporabniku, ki išče isto stvar na drugem terminalu. Web 3.0 iskalniki bodo vedno bolj specifični do uporabnika in proizvajali bodo natančne rezultate iskanja. Zmedeni? No, oglejmo si stanje malo bolj podrobno. Iskalne naprave Web 3.0 ere naj bi se ravnale po kompleksni poizvedbi, le-te pa so sprogramirane zelo podobno našem naravnem govoru. Podobno tistemu, kar uporabnik utipka kot poizvedbo: "Selim se iz Hrvaške v Ljubljano in iščem namestitev. Sem poročen in imam sina in hčerko. Kakšni so življenjski stroški v Ljubljani?". Iskalnik bo pridobil podatke iz različnih lokacij in bo podal pravo stran z rezultati v delčku sekunde. Torej domnevamo, da bodo spletne strani med seboj bolj komunikativne. Ni dvoma, da so spletne strani komunicirale tudi v Web 2.0 obdobju. Z enim klikom na naslov URL na eni strani vam je omogočeno potovanje navzdol vse do nove spletne strani. V Web 3.0 fazi pa se bodo izmenjavale informacije med seboj, da se dobijo prav rezultati, ki jih uporabnik želi. Drug primer je, da nam bo ena prijava omogočala, da posodobimo naš status na Facebook-u, Twitter-u ali MySpace-u naenkrat.

Toliko vsega za natančen del iskanja. Zdaj pa, kaj mislimo, ko govorimo o podatkih, ki so

specifični za enega uporabnika. Pojmi, ki so uporabljeni kot pri iGoogle²³, bodo postali popularni in izboljšani. Rezultati iskanja se bodo razlikovali pri vsakem uporabniku posebej. Iskalniki bodo sledili rezultatom za katere se določeni uporabnik zanima, in bodo podajali različne rezultate iskanja za različne uporabnike. Tudi oglasi, ki si jih uporabnik ogleda bodo drugačni od tistih, katere si ogleda druga oseba z isto iskalno poizvedbo. To bo vse uporaba umetne inteligence. Web 3.0 aplikacije bodo narejene tako, čeprav ne enako inteligentne kot človeški možgani, ampak kot nadaljnja pot iz vašega urejevalnika besedil s sposobnostjo napovedovanja besed. Nekateri menijo, da je to vpogled v njihovo zasebnost, vendar se iskalniki ne morejo drugače pokazati. Ampak v to razpravo ne bomo šli globlje.

Tradicionalni iskalniki temeljijo na tehnologijah pridobivanja informacij. Izvajajo postopke kot je logična poizvedba, približno iskanje, pomen besedila in analiza povezav (Primeri: Google, Yahoo, MSN, Ask).

Web 2.0 iskanje na podlagi ključnih besed povzroča preobremenjenost z informacijami. Naslednji atributi, ki bodo del Web 3.0 so: vsebinsko iskanje, posebno oblikovano iskanje, prilagojeno iskanje, razvoj 3D Spleta in dedukcija[11].

Najlažje je prikazati koncept novejšega iskanja s primerom. Recimo, da razmišljate o dopustu. Rad bi šli nekam v tople tropske kraje. Prihranjenih imamo 2500eur za naše potovanje. Želimo prijeten kraj za bivanje vendar ne želimo, da bo drago. Prav tako si želimo precej možnosti za izbiro letala. S spletno tehnologijo, ki nam je trenutno na voljo smo morali narediti veliko raziskav, da bi poiskali najboljše možnosti za počitnice. Morali smo raziskovati potencialne destinacije in se odločiti, katera je ta prava za nas. Morda smo morali obiskati dve ali tri cenejše potovalne strani in primerjati cene letov in hotelskih sob. Zgubili smo veliko svojega časa za iskanje rezultatov po različnih iskalnikih. Celoten proces je lahko trajal tudi več ur. Po mnenju nekaterih Internet strokovnjakov, se bomo z Web 3.0 sede prepustili, da Internet naredi vso delo za nas. Lahko uporabimo storitev iskanja in ozke parametre iskanja. Brskalnik potem združuje, analizira in prikazuje podatke na način, ki omogoča primerjavo posnetka. To lahko stori, saj bo Web 3.0 sposoben razumeti informacije na spletu. Pri Web 2.0 standardu nas iskalnik med uporabo ni mogel zares razumeti. Videti je, da vsebujejo spletne strani ključne besede, ki so v našem iskalnem pojmu. Iskalnik ne more povedati, če je spletna stran dejansko pomembna za našo poizvedbo. Lahko samo pove, da je ključna beseda na spletni strani. Na primer, če iščemo izraz "Saturn", lahko končamo z rezultati na spletni strani o planetu ali na strani proizvajalca avtomobilov. Web 3.0 iskalnik

²³ iGoogle je vaša prilagojena Googlova stran. Vaši strani dodajate novice, fotografije, vreme in različne elemente s celotnega spleta

lahko najde ne samo ključne besede iskanja, ampak tudi razlago v zvezi z našo zahtevo. Vrnil bo ustrezne rezultate in predlagal drugo vsebino povezano z iskanimi besedami. Če bi v našem primeru počitnic, vnesli "destinacija tropske počitnice pod 2500eur" kot iskalno zahtevo, lahko Web 3.0 brskalnik vrne rezultat skupaj s seznamom zabavnih in odličnih restavracij, ki so v zvezi z rezultatom iskanja. To naj bi obravnaval celoten Internet, kot veliko bazo podatkov, na voljo za kakršno koli vprašanje.

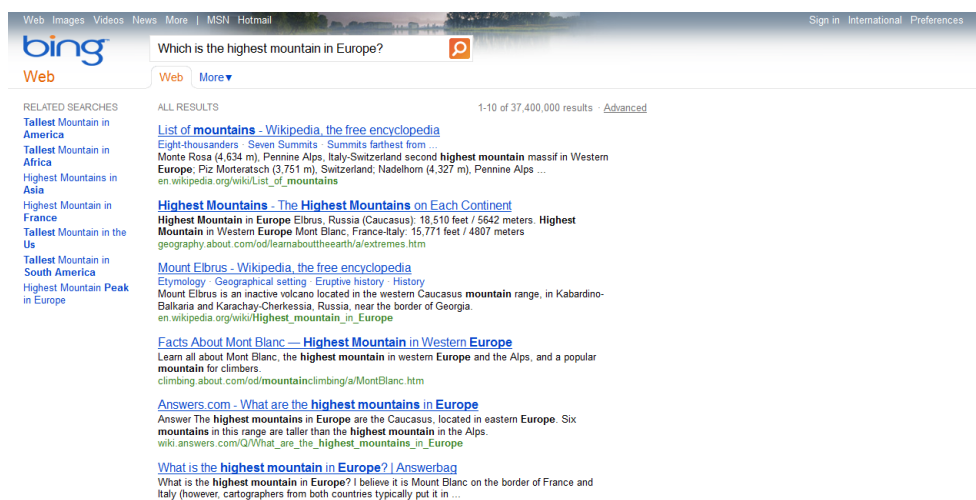
Primeri nekaterih semantičnih iskalnikov:

1. Bing



Slika 4: Semantični iskalnik Bing

Bing (prej Live Search, Windows Live Search in MSN Search) je Microsoftov spletni iskalnik. Bing je prvič predstavil Microsoft CEO Steve Ballmer 28. maja 2009 na konferenci »All Things Digital« v San Diegu. Nato se je v celoti pokazal na spletu 3. junija 2009. Pomembne spremembe vključujejo seznam predlogov za iskanje, kot so vnesene poizvedbe in seznam povezanih poizvedb (imenovan "Explore pane"), ki temelji na semantični tehnologiji Powerset, ki jo je Microsoft kupil leta 2008. 29. julija 2009 sta Microsoft in Yahoo napovedala dogovor v katerem bi bil Bing močnejši kot Yahoo Search. Prehod na samo Yahoo! Search naj bi stranke po svetu in partnerji naredili v začetku leta 2012.



Slika 5: Prikaz iskanja v Bing iskalniku

Powerset je delal na izgradnji iskanja na principu naravnega jezika, da lahko najde ciljno usmerjene odgovore na vprašanja uporabnikov (v nasprotju z iskanjem na podlagi ključnih besed). Na primer, ko vzamemo vprašanje kot je "Katera je najvišja gora v Evropi?", običajni iskalniki ignorirajo frazo vprašanja in namesto tega iščejo na podlagi ključnih besed "najvišja", "gora" in "Evropa". Powerset na drugi strani poskuša uporabiti obdelavo naravnega jezika, da se razume narava vprašanja in vrne strani, ki vsebujejo odgovor.

Ene od značilnosti Bing iskalnika[12]:

- Značilnosti vmesnika (npr. Stran s slikami/Images page) prikazuje glavno sliko iz tega dne in štiri iskanja, ki se nanašajo na to sliko s tremi predogledi slik po iskalnem izrazu)
- Značilnosti medijev (npr. Iskanje slik z neprekinjeno premikajočimi slikami na strani z rezultati, ki ima možnost nastavitve velikosti, oblike, barve, stila in ljudi)
- Takojšnji odgovori (npr. Napredni izračuni. Z uporabo WolframAlpha računskega iskalnika, lahko poda Bing tudi rezultate za napredne matematične probleme (npr. $\lim x/2x$ za $x \rightarrow 2$) in druge WolframAlpha sorodne poizvedbe (npr. kalorije v pizzii))
- Lokalne informacije (npr. Treutne prometne informacije)
- Integracija z Hotmail-om (npr. S Hotmailovo "Quick Add" funkcijo, uporabniki lahko vstavijo rezultate iskanja pri Bing iskalniku neposredno v svojo e-pošto)

sporočilo, npr. kot so ocene restavracij, filmski spored, sliko, video posnetek in zemljevid)

- Integracija z Facebook-om (npr. Uporabniki imajo možnost do pošiljanja sporočil svojim prijateljem v rezultatih iskanja)

2. Swoogle



Slika 6: Semantični iskanik Swoogle

Swoogle je iskalnik Semantičnega spleta in sicer njegovih ontologij, dokumentov, pogojev in podatkov, ki so objavljeni na spletu. Swoogle zaposluje sistem robotov, da bi odkrili RDF dokumente in HTML dokumente z vstavljenjo RDF vsebino. Swoogle razlaga te dokumente in njihove sestavne dele (npr. pogoji in trojicami) ter zapise in indekse smiselne v njihovih metapodatkih v svoji podatkovni bazi. Swoogle opravlja storitve za človeške uporabnike prek brskalniškega vmesnika in za programske agente prek RESTful²⁴ spletnih storitev. Uporablja se več tehnik za rang rezultatov poizvedb na podlagi PageRank²⁵ algoritma, ki je razvit na Google-u, toda prilagojen za semantiko in načine uporabe, ki so na voljo pri semantičnih dokumentih na spletu[13]. Iskalnik opravlja lezenje semantičnih dokumentov, tako kot je pri večini spletnih iskalnikov in je iskanje na voljo tudi kot spletna storitev. Iskalnik je prvotno napisan v Javi z uporabo jezika PHP za front-end in MySQL za bazo podatkov. Swoogle je sposoben iskati več kot 10 000 ontologij in indeksov, za več kot 1 300

²⁴ RESTful - je spletna storitev (imenovana tudi RESTful web API) je enostavna spletna storitev, ki se izvaja z uporabo HTTP in načelih REST

²⁵ PageRank - je algoritem analize povezav, ki ga uporablja Google za internet iskanje

000 spletnih dokumentov. Prav tako izračuna pomembnost dokumenta Semantičnega spleta. Tehnike, ki so uporabljene za indeksiranje so podobne google tipu uvrščanja strani, kot tudi kopanje dokumentov za medsebojne odnose, ki so podlaga za semantični splet.

Swoogle je bil razvit in se nahaja na Univerzi v Maryland-u, Baltimore County (UMBC), s financiranjem ZDA DARPA in National Science Foundation agencije. To je doktorsko delo Li Ding-a, pod mentorstvom profesorja Tim Finina[13]. Ker je to raziskovalni projekt, ki je narejen z nekomercialnim motivom, ni veliko reklame okoli Swoogle-a. Vendar pa je pristop k indeksiranju dokumentov Semantičnega spleta pristop, do katerega bo večina iskalnikov morala s časom priti.

Minilo bo še veliko časa dokler bo semantično iskanje postalo popolnoma uporabno. Veliko število spletnih razvijalcev ne pozna nove standarde in njihove strani zaradi tega niso zgrajene za uporabo semantičnih iskalnikov.

4.3 Mikroformati in RDFa

Semantični splet je okolje strukturiranih podatkov. Zagotavlja pa taka sredstva, ki pretvorijo vsako vsebino ali podatke v strukturirane podatke, da jih lahko uporabi druga programska oprema. To je dejansko tisto kar omogoča RDF. Semantični splet ne zahteva zapletene ontologije. Tudi brez uporabe OWL-ja in še bolj sofisticiranih ontologij, so lahko močna izmenjava in povezovanje podatkov omogočeni na obstoječem spletu z uporabo samega RDF-ja. Semantični splet ne obstaja le na spletnih straneh. RDF deluje znotraj aplikacij in podatkovnih baz, ne le na spletnih straneh. Pri Semantičnem spletu ne gre le za rudarjenje²⁶, t.j. iskanje in brskanje. Razvijalci aplikacij, ponudniki vsebin in končni uporabniki imajo korist od uporabe Semantičnega spleta (RDF) v okviru svoje storitve, ne glede na to, ali so izpostavili, da so metapodatki RDF za zunanje stranke. RDF je koristen brez kakršnega koli rudarjenja podatkov - lahko je vstavljen prav v vsebino avtorskih orodij in pregledno ustvarjen ob objavi podatkov. RDF naredi vsebino bolj obvladljivo in omogoča razvijalcem in ponudnikom vsebin, da pri tem ne gledajo na relacijske podatkovne modele.

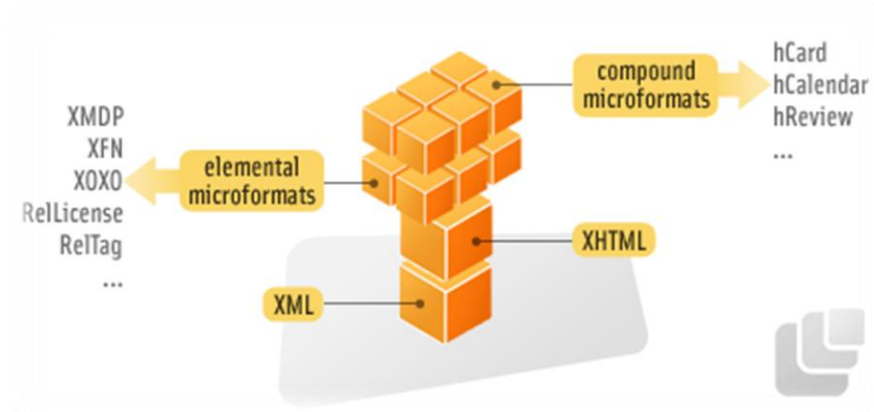
²⁶ Data mining

Prav tako daje končnim uporabnikom boljši način za zbiranje in upravljanje vsebine, ki jo najdejo.

Mnogi se strinjajo, da so Mikroformati le sredstvo za dosego cilja, in da bo tehnologija v središču Semantičnega spleta RDF. To se lahko zgodi, vendar nam ponujajo Mikroformati možnost, da se začnemo igrati s semantičnim spletom zdaj in tako se učimo in razumemo, kaj bo splet v prihodnosti. Ko RDF prevzame večjo vlogo zunaj akademskih krogov, bomo v dobrem položaju, da izkoristimo večino Semantičnega spleta.

4.3.1 Mikroformati

Povejmo nekaj o mikroformatih. Kaj so? Zakaj se uporabljajo?



Slika 7: Diagram Mikroformatov

Mikroformat je spletni pristop, ki temelji na semantičnem označevanju, ki poskuša s ponovno uporabo obstoječih HTML/XHTML oznak izraziti metapodatke in druge attribute na spletnih straneh in v drugih kontekstih, ki podpirajo (X)HTML, kot je npr. RSS. Ta pristop omogoča programski opremi samodejno procesiranje informacij, namenjene za končne uporabnike (kot so kontaktni podatki, geografske koordinate, koledar dogodkov in podobno). Mikroformati so izdelani najprej za ljudi potem za stroje, in so skupek preprostih, odprtih podatkovnih formatov, ki so zgrajeni na obstoječih in široko sprejetih standardih. Namesto da bi vrgli stran tisto kar deluje danes, imajo Mikroformati namen reševati preprostejše probleme, najprej s prilagajanjem na sedanje vedenje in vzorce uporabe (npr. XHTML, blogging)[14]. Hitro postajajo način za izmenjavo informacij med različnimi aplikacijami. Leto 2007 je treba imenovati kot leto Mikroformatov. Ker si vedno več ljudi poskuša deliti podobne podatke z zelo različno uporabo, so se začeli hitro zavedati, kako je pomembno imeti standarde za

izmenjavo informacij. Web 2.0 norčevanje je povzročilo številne Mikroformate, ali miniaturne standarde za veliko vrst informacij.

Specifikacije mikroformatov[14]:

- hCard, XFN – Ljudje in organizacije
- hCalendar – Koledarji in dogodki
- hReview, VoteLinks – Mnenja, ocene in pregledi
- rel-tag – Oznake, ključne besede in kategorije
- rel-license – Licence
- XOXO – Seznami in osnutki
- in še veliko več.

Mikroformati niso nov jezik, temveč so v tesni povezavi z obstoječimi spletnimi stranmi, pogosto brez potrebe po kakršni koli vidni spremembi. Večinoma so način za dodajanje semantike programiranju. Če ste bralec bloga in ste videli "Oznake" ali "Kategorije" pri določenih postih, ste verjetno naleteli na Mikroformate v akciji. Že leta 2010 omogočajo Mikroformati kodiranje in ekstrakcijo dogodkov, kontaktne informacije, socialne odnose in tako naprej..

XHTML in HTML standardi omogočajo vgradnjo in kodiranje semantike v attribute označevalnih oznak. Mikroformati izkoriščajo te standarde, kar kaže na prisotnost metapodatkov z uporabo naslednjih lastnosti:

- class
- rel
- rev

Primer:

V tem primeru so kontaktni podatki predstavljeni na naslednji način:

```
<div>
  <div>Ljiljana Perkovic</div>
  <div>Satbiro d.o.o.</div>
  <div>0038551841335</div>
  <a href="http://satbiro.hr/">http://satbiro.hr/</a>
```

```
</div>
```

Z hCard markupom mikroformata to postane:

```
<div class="vcard">
  <div class="fn"> Ljiljana Perkovic </div>
  <div class="org"> Satbiro d.o.o.</div>
  <div class="tel0038551841335"></div>
  <a class="url" href="http:// satbiro.hr /">http:// satbiro.hr </a>
</div>
```

Torej, uporabljeno ime (fn), organizacija (org), telefonska številka (tel) in spletni naslov (url) so bili identificirani s pomočjo posebnega razreda. Celotna stvar je zavita v class = "vCard", kar pomeni, da drugi razredi v obliki hCard (kratica za "vCard HTML") niso imenovani po naključju. Obstajajo seveda še drugi hCard razredi. Programska oprema, kot so priključki brskalnika, lahko izvelečejo tiste informacije o osebi in jih prenesejo v neko drugo aplikacijo kot je npr. adresar.

Uporaba Mikroformatov z HTML-jem omogoča dodatno oblikovanje in semantične podatke, ki jih aplikacije lahko uporabljajo. Na primer aplikacije, kot so spletni iskalniki, ki zbirajo podatke o online virih, ali namizne aplikacije, kot so e-poštni nabiralniki ali urejevalniki programske opreme, ki lahko prevajajo nekatere podrobnosti. Uporaba Mikroformatov lahko tudi olajša "mash up", tako da izvaža vse geografske lokacije na spletni strani v npr. Google Maps, in se jih tako predstavi v prostoru. Nekaj razširitev brskalnika, kot je Operater za Firefox in Oomph za Internet Explorer, zagotavljajo zmožnost zaznavanja Mikroformatov v HTML dokumentih. Ko so vključeni hCard ali hCalendar, take razširitve brskalnika omogočajo, da se jih izvede v formate, ki so združljivi s pripomočki za upravljanje stikov in koledarja, kot je npr. Microsoft Outlook. V maju leta 2009 je Google napovedal, da bo parsali Mikroformate hCard, hReview in hProduct, in jih uporabili za popularizacijo strani, ki so rezultat nekega iskanja. Razširili so še uporabo hCalendar za prireditve in hRecipe za kuharske recepte. Mikroformati niso edina rešitev za zagotavljanje "bolj inteligentnih podatkov" na spletu, obstajajo tudi alternativni pristopi, ki pa so še v fazi razvoja. Lahko uporabimo RDFa, da bi vgradili poljubne slovarje v HTML, in sicer vključevanje domensko specifičnih znanstvenih podatkov na svetovnem spletu, kot so npr. živalski ali kemijski podatki, kjer ne obstajajo mikroformati za take podatke. Poleg tega standardi, kot je W3C-jev

GRDDL²⁷, omogočajo Mikroformatom, da se pretvorijo v podatke združljive s Semantičnim spletom. Postalo je težko (čeprav je mogoče) uporabljati kombinacijo različnih mikroformatov. V nasprotju lahko RDF predstavi vse informacije, vključno s tistimi pridobljeni iz Mikroformatov na določeni strani. Tukaj imajo mikroformati korist od RDF-ja - splošna pravila orodij Semantičnega spleta omogočajo lažjo ponovno uporabo obstoječih orodij, kot je npr. povpraševalni jezik²⁸ in kombiniranje izkazov iz različnih izvorov kar enostavno sodi v samo bistvo Semantičnega spleta. GRDDL je "most" za pristop k Mikroformatom; opredeljen je splošni postopek, s katerim se lahko Mikroformate shranjene v datoteki XHTML preoblikuje v RDF. Druga tehnologija je RDFa, ki opredeljuje XHTML1.1 modul in daje možnost virtualne uporabe skoraj vseh RDF slovarjev, kot pripombo vsebine XHTML. Nekaj podobnega kot so Mikroformati z nekoliko bolj strogim in boljšim načinom vključevanja različnih slovarjev znotraj istega dokumenta.

Če smo se dotaknili podrobnostih o mikroformatih, povejmo tudi nekaj o RDFa in kakšna je njegova vloga.

4.3.2 RDFa

RDFa (ali Resource Description Framework – v – atributih) je priporočilo W3C-ja, ki dodaja niz razširitev atributov v XHTML za vključevanje bogatih metapodatkov v spletne dokumente. Ekstenzije za XHTML – RDFa omogočajo, da semantično opišemo naše podatke, ne da bi za to žrtvovali uporabnost naše strani. Hkrati pa našo stran naredimo še bolj prijazno za spletne iskalnike in druge »nežive« obiskovalce naših strani[15]. Lahko bi ga opisali kot CSS po pomenljivosti. Ko uporabimo RDFa lahko brskalniki, iskalniki in ostala programska oprema razumejo več o straneh, pri tem pa ponujajo več storitev ali boljše rezultate za uporabnika. Na primer, če brskalnik ve, da gre na eni strani za dogodek, kot je konferenca, lahko ponudi, da ga dodate na svoj koledar, pokaže njegovo lokacijo na zemljevidu, poišče hotele ali letalo in poljubno število drugih stvari. Bistvo RDFa zagotavlja množica atributov, ki se lahko uporablja za prenos metapodatkov v jezik XML. Ti metapodatki se lahko uporabijo za označevanje dogodkov na koledarju, kontaktnih podatkov, različnih dovoljenj, itd... S tem, da ima vgrajene metapodatke, igra RDFa pomembno vlogo pri velikem načrtu možnosti za semantični splet. Podatki spletnih strani niso berljivi le s človeškimi očmi ampak

²⁷ GRDDL - je označevalni format za Gleaning Resource Descriptions from Dialects jezike. To je priporočilo W3C, ki uporabnikom omogoča pridobitev RDF trojice iz dokumentov XML, vključno z XHTML

²⁸ Query language - so računalniški jeziki, ki se uporabljajo za izdelavo poizvedb v podatkovnih bazah in informacijskih sistemih

tudi z avtomatiziranimi procesi, ki lahko zbirajo podatke in s tem povezujejo metapodatke ter nato opravljajo naloge, ki so veliko bolj zapletene kot tiste, ki jih zmorejo tipične računalniške aplikacije zdaj. RDFa omogoča izrazom, ki so del več samostojno razvitih slovarjev, da bodo svobodno pomešani. Zasnovan je tako, da je mogoče razčleniti jezik brez posebnega znanja izrazov slovarja, ki ga uporabljate. Lastnosti metapodatkov, ki so na voljo v XHTML-ju omogočajo samo izražanje informacij o samem dokumentu, medtem ko RDFa omogoča, da dokument vsebuje metapodatke o drugih dokumentih in virih.

Spodaj je seznam RDFa atributov. V primerih, ki sledijo, bomo pokazali, kako jih uporabiti v XHTML dokumentu[15].

- **about** – URI ali CURIE (Compact URI) določa vir metapodatkov; v njegovi odsotnosti se za to privzame trenutni dokument
- **rel in rev** – določa razmerje ali obratno razmerje z drugim virom
- **href, src in resource** – določa vir partnerja
- **property** – določa lastnost za vsebino elementa
- **content** – izbirni atribut, ki prevladuje nad vsebino elementa pri uporabi lastnosti atributa
- **datatype** – izbirni atribut, ki določa podatkovni tip besedila, ki je določeno za uporabo lastnosti atributa
- **typeof** – izbirni atribut, ki določa RDF tip predmeta (vir metapodatkov).

Lepota RDFa je, da sledi petim načelom - "načela interoperabilnih metapodatkov"[15]:

1. Neodvisnost izdajatelja - vsaka stran lahko uporabi svoje standarde
2. Ponovna uporaba podatkov - podatki niso podvojeni. Ločeni XML in HTML oddelki niso potrebni za isto vsebino
3. Lastno zadrževanje - HTML in RDF so ločeni
4. Shema modularnosti - atributi se lahko ponovno uporabijo
5. Razširjenost - dodatna polja, se lahko dodajajo in XML transformacije lahko pridobijo semantiko podatkov iz datotek XHTML

RDFa uporablja nekaj preprostih XHTML atributov za kategorizacijo podatkov, do te mere, da ni na voljo za mikroformate. RDF trojka posreduje več informacij v obliki, ki je zlahka

predstavljena z iskalniki. Ta kategorizacija je vse bolj koristna, celo nujna, ko postaja splet vse bolj semantične narave.

Primeri RDFa[15]

V nadaljevanju sledi primer dodajanja Dublin Core (Dublin Core, DC, skupnost, ki je že opredelila slovar uporabnih konceptov za opis listin, vključno z obema ustvarjalcema, naslovom in datumom, kjer naslov dejansko pomeni naslov dela) metapodatkov v XHTML datoteko. Dublin Core podatkovni elementi so podatki, ki se običajno dodajajo v knjige ali članke (naslov, avtor, predmet itn.).

```
<div xmlns:dc="http://purl.org/dc/elements/1.1/"
  about="http://www.example.com/books/wikinomics">
  <span property="dc:title">Wikinomics</span>
  <span property="dc:creator">Don Tapscott</span>
  <span property="dc:date">2006-10-01</span>
</div>
```

Poleg tega RDFa omogoča odlomkom in besedam besedila, da so povezani s semantičnim markupom:

```
<p xmlns:dc="http://purl.org/dc/elements/1.1/"
  about="http://www.example.com/books/wikinomics">
  In his latest book
  <cite property="dc:title">Wikinomics</cite>,
  <span property="dc:creator">Don Tapscott</span>
  explains deep changes in technology,
  demographics and business.
  The book is due to be published in
  <span property="dc:date" content="2006-10-01">October 2006</span>.
</p>
```

XHTML+RDFa 1.0 primer

Spodaj je primer popolnega XHTML + RDFa 1.0 dokumenta. Uporablja Dublin Core in FOAF (ontologija za opis ljudi in njihovih odnosov z drugimi ljudmi in stvarmi):

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML+RDFa 1.0//EN"
```

```

"http://www.w3.org/MarkUp/DTD/xhtml-rdfa-1.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
      xmlns:foaf="http://xmlns.com/foaf/0.1/"
      xmlns:dc="http://purl.org/dc/elements/1.1/"
      version="XHTML+RDFa 1.0" xml:lang="en">
<head>
  <title>John's Home Page</title>
  <base href="http://example.org/john-d/" />
  <meta property="dc:creator" content="Jonathan Doe" />
  <link rel="foaf:primaryTopic" href="http://example.org/john-d/#me" />
</head>
<body about="http://example.org/john-d/#me">
  <h1>John's Home Page</h1>
  <p>My name is <span property="foaf:nick">John D</span> and I like
    <a href="http://www.neubauten.org/" rel="foaf:interest"
      xml:lang="de">Einstürzende Neubauten</a>.
  </p>
  <p>
    My <span rel="foaf:interest" resource="urn:ISBN:0752820907">favorite
    book is the inspiring <span about="urn:ISBN:0752820907"><cite
    property="dc:title">Weaving the Web</cite> by
    <span property="dc:creator">Tim Berners-Lee</span></span>
  </span>
  </p>
</body>
</html>

```

V zgornjem primeru vidimo, da dokument URI predstavlja dokument HTML, torej dokument URI skupaj z "#me" nizom `http://example.org/john-d/#me` predstavlja dejansko osebo, za razliko od dokumenta o njih. *foaf:primaryTopic* v glavi nam pove URI osebe in njenega dokumenta. *foaf:nick* lastnost (v prvem `span` elementu) vsebuje vzdevek za to osebo in *dc:creator* lastnost (v `meta` elementu) nam pove kdo je ustvaril dokument. Hiperpovezava do spletne strani Einstürzende Neubauten strani vsebuje `rel="foaf:interest"`, kar nam pove, da se John Doe zanima za to skupino. URI njihove spletne strani je vir strani. *foaf:interest* znotraj drugega `p` elementa se nanaša na knjigo s številko ISBN. Atribut `resource` določa vir na podoben način kot atribut `href`, ampak brez opredelitve hiperpovezave. Nadalje v odstavku, element `span` vsebuje atribut `about`, ki opredeljuje knjigo kot drugi vir za

določitev metapodatkov o tem. Naslov knjige in avtor so opredeljeni znotraj tiste oznake z uporabo lastnosti *dc:title* in *dc:creator*.

Tukaj so enake trojke, ko se zgoraj navedeni dokument samodejno pretvori v RDF/XML:

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description rdf:about="http://example.org/john-d/">
    <dc:creator xml:lang="en">Jonathan Doe</dc:creator>
    <foaf:primaryTopic>
      <rdf:Description rdf:about="http://example.org/john-d/#me">
        <foaf:nick xml:lang="en">John D</foaf:nick>
        <foaf:interest rdf:resource="http://www.neubauten.org/">
          <foaf:interest>
            <rdf:Description rdf:about="urn:ISBN:0752820907">
              <dc:creator xml:lang="en">Tim Berners-Lee</dc:creator>
              <dc:title xml:lang="en">Weaving the Web</dc:title>
            </rdf:Description>
          </foaf:interest>
        </rdf:Description>
      </foaf:primaryTopic>
    </rdf:Description>
  </rdf:RDF>
```

Prihodnost spletnih brskalnikov bo verjetno sodelovanje s semantično označenimi podatki na katere ste naleteli na spletu s posebnimi aplikacijami, ali na vašem operacijskem sistemu ali na spletu. To pomeni, da se bodo kontaktne informacije, ki jih vidite na spletni strani povezale z vašo priljubljeno aplikacijo o stikih, prireditve se bodo povezale z vašo priljubljeno aplikacijo koledarja, lokacije se bodo povezale z vašo priljubljeno aplikacijo kart, telefonske številke se bodo povezale z vašo priljubljeno aplikacijo VOIP, itd. Hitreje ko se bojo ljudje odzvali na te vrste standardov, prej bo splet resnično postal en interaktivni vir informacij.

4.3.3 Mikroformati vs. RDFa

Kaj je boljše uporabljati? RDFa ali Mikroformate? No pa si pogledjmo naprej nekatere razlike med temi dvema standardi:

1. Različna zgodovina

Microformats.org stran je zrasla iz dela skupnosti razvijalcev Technorati, medtem, ko je RDFa zrasel, kot koncept W3C Semantic Web skupine.

2. Različen markup podatkov

Mikroformati uporabljajo atribut "class" za zagotovitev značilnosti npr. class = "locality". Dokler W3C uporablja novejša attribute kot je "property" za zagotovitev značilnosti npr. property = "dc:location".

Kam nas pelje prihodnost?

Obstaja očitno nasprotje med tema dvema načinoma oblikovanja semantičnih markupov. Več ljudi trenutno uporablja Mikroformate (ki so prav tako priznani od več brskalnikov), ampak RDFa je razvil W3C in uporablja XML imenski prostor, ki zagotavlja prilagodljivo obliko za semantične vsebine. Ko se bodo XHTML 2 standardi začeli bolj pogosto uporabljati, bodo, upajmo, razvijalci začeli več uporabljati RDFa za markup Mikroformatov (to je zanimiva razprava). V nasprotnem primeru to lahko pomeni, da bomo programirali za različne brskalnike in za povezovanje v različnih formatih. Začetna prednost Mikroformatov je bila, da so delali brez težav v obstoječi HTML kodi, ko so se uporabljali znotraj strani, niso bile potrebne nobene posebne oznake, ki bi lahko prekomerno omejile eno HTML različico niti niso povzročile stran neveljavno. Slaba stran je, da je v prvi vrsti zahteva uporabo posebnih poimenovanj atributov razreda - rekonstrukcijo spletnih mest, da imajo hCard potrebno preimenovanje CSS razredov ali dodajanje več oznak div/span za dodajanje v posebej imenovane razrede. Prav tako je bilo marked-up dele naslova ali karkoli drugega treba ustrezno popraviti za pravilno delovanje.

Zdaj pa so zgradili bolj prilagodljiv RDFa, saj je bil ustanovljen kot bolj popolni XHTML - in ena ključnih značilnosti XHTML je "X" del – kar pomeni je "razširljiv", torej lahko dodate imenski prostor za lastne namene, ne da bi poškodovali dokument. RDFa je dodal "property", "role" in "about" attribute za informacije o označevanju primerne za stroje, prednost tega pa je, da jih lahko uporabite, ne da bi morali nujno spremeniti svoj CSS. Slaba stran pa je, da morate imeti tehnično dobro oblikovan dokument, ki bo deloval v okviru XHTML - bolj zahteven model kodiranja dokumenta od prejšnjih različicah HTML. (Seveda, lahko dodate RDFa kodo v svojo vsebino, brez da bi naredili celotno XHTML stran veljavno, ampak to je dokaj tvegano, saj se odpira vprašanje ali bodo iskalniki lahko pravilno razčlenili stran, če ni

pravilno oblikovana. Medtem, ko brskalniki in iskalniki ves čas (ne)uspešno interpretirajo neveljavno kodno stran, obstaja tveganje, da bo uničen markup RDFa lahko večji, če stran ni veljavna, saj je markup odvisen od strojne pravilne prepoznave XHTML markup elementov in njihove pravilne razlage.)

RDFa je veliko bolj razširjena in uporabna rešitev. Če se Semantični splet ne bi zgodil, bi bilo potrebno zahtevati markup močnejši od mikroformatov. Mikroformati imajo manj ovir pri izdelavi. Oni so konceptualno preprostejši in bolj osredotočeni na posamezne slovarje. Težje je za začetnike, če delajo z neomejeno rešitvijo in jih tak način vodi naprej. Imenski prostori z osnovo URI so nejasni (uporabljaš URL za nekaj, kar ni ravno spletna stran) in podvrženi napakam (na videz nepomembne razlike, kot so dodatni slash(črka »/«) ali www. naredijo različen imenski prostor). Če pogledaš kako je HTML tagsoup²⁹ izdelan v nasprotju z XHTML, je jasno, da bo večina uporabila preprosto rešitev, ki načeloma naredi določeno delo, čeprav se tako izognejo razširljivosti, eleganci, itd. Večina uporabnikov bo brez razmišljanja kopirala kodo in se naučila delovanja iz vira drugih strani, samo zaradi tega mislim, da mikroformati imajo velike možnost za preživetje v takih razmerah. Podatki kažejo, da se je uporaba RDFa povečala za 510% med marcem 2009 in oktobrom 2010, iz 0,6% spletnih strani na 3,6% spletnih strani. Zato se lahko v veliki meri zahvalimo prizadevanjem ljudi na portalu Yahoo! (z SearchMonkey), Google-u (z Rich Snippets) in Facebook-u (z Open graf), to je zahvala vsem ki priporočajo uporabo RDFa. Te informacije nam prebujajo optimizem, ker nam kažejo, da je Semantični splet že tu in v veliki meri. Uporaba mikroformatov ni bistveno napredovala v tem enakem obdobju, razen hatom mikroformata. Obstajata dve vrsti mikroformatov: eden za HTML (ki ne uvaja nikakršnih novih elementov ali atributov) in drugi za XHTML (ki dodaja nekaj novih atributov, ampak nobene nove oznake). Predlagan RDFa standard se lahko razume kot vrsta Mikroformatov za XHTML, saj temelji na formatu RDF XML. Resnica je, da obstajata dve skupini ljudi (RDFa, Mikroformati), ki delajo v istem prostoru tako, da ga popolnoma prekrivajo. Vendar za zdaj deluje ločevanje HTML vs. XHTML dobro za kategorizacijo teh dveh vgrajenih prizadevanj dokumentov.

Poleg pregledov razumeta Google in Yahoo najmanj štiri druge vrste vgrajenih dokumentov: ljudi / podjetja, izdelke, dogodke in video posnetke. Ker obstajata dva organa za standarde na

²⁹ Tafsoup - v spletnem razvoju, tagsoup se nanaša na oblikovan markup napisan za spletno stran, ki je zelo podoben kot HTML, vendar ne vključuje pravilno sintakso HTML in strukture dokumenta.

delovnem mestu, obstajajo za vsakega Mikroformatske in RDFa špekulacije. Spodnji pregled povzema te povezave in poda primere teh vgrajenih dokumentov[16].

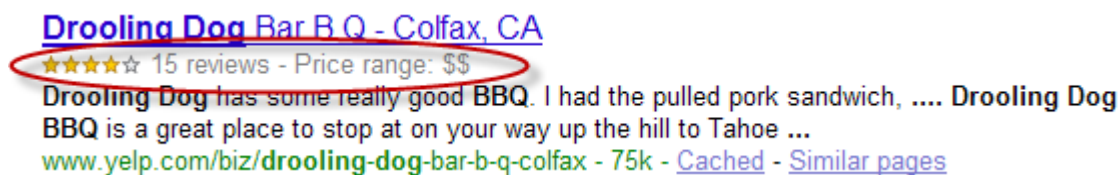
- Pregledi: hReview (microformat) v:Review (RDFa)
- Ljudje/Podjetja: hCard/vCard (microformat) v:Person (RDFa)
- Dogodki: hCalendar (microformat)
- Izdelek: hProduct (microformat) v:Product (RDFa)
- Video: Facebook Share (microformat), Yahoo! SearchMonkey (za video posnetke) (RDFa)

V zadnjih letih se je korist za strukturirane podatke zelo okrepila s predstavitvijo podatkov, da bi tako pritegnili potrošnike. Yahoo-jevo oblikovanje izboljššanega rezultata iskanja s svojo platformo SearchMonkey ter kasnejša uvedba Google-ovog Rich Snippets-a, dovoljuje spletnim skrbnikom, ki uporabljajo strukturirane podatke, da ustvarjajo več privlačnih seznamov. Teorija (ki jo podpira statistika v poročilu Yahoo-a) kaže, da lahko privlačnejši seznam iskanja povečajo razmerje prikazov in stopnjo iskanja.

Pokazali bomo še en primer, kako lahko uporabimo Mikroformate in RDFa skupaj, in sicer Google-ov Rich Snippets[17]:

Rich Snippets podaja uporabnikom na prvi pogled primeren povzetek podatkov o njihovih rezultatih iskanja. Trenutno podpiramo podatke o pregledih in ljudeh. Pri iskanju izdelka ali storitve uporabniki enostavno vidijo preglede in ocene, ter dobijo pri iskanju osebe pomoč razlikovati osebe z istim imenom. Gre za preprosto spremembo prikaza rezultatov iskanja, vendar so naši poskusi pokazali, da lahko uporabniki najdejo pomembne nove podatke. To pomeni, da če se jim zdi informacija iz strani koristna in ustrezna, jo bolj verjetno kliknejo. Zdaj pa smo na začetku procesa odpiranja tega uspešnega poskusa, tako da ima vse več spletnih strani možnost sodelovanja. Kot oblikovalec spletnih strani, lahko pomagate s kodiranjem vaše strani s strukturiranimi podatki v standardnem formatu. Če želite videti prikaz Rich Snippets-a, Google poišče markup formatov (Mikroformate in RDFa), ki jih lahko enostavno dodate na svoje spletne strani. V večini primerov je to tako hitro, kot je zapisovanje obstoječih podatkov na vaših straneh z nekaj dodatnimi oznakami. Na primer,

tukaj je nekaj pomembnih HTML vrstic z Yelp³⁰ pregledne strani za " Drooling Dog BarBQ", preden je dodan markup podatkov:



Slika 8: Prikaz delovanja Rich Snippets-a

```
<h1>Drooling Dog Bar B Q</h1>
...

<em>based on 15 reviews</em>
...
<strong>Price Range:</strong> $$
```

in zdaj z markup-om mikroformatov:

```
<div class="hreview-aggregate">
  <div class="item vcard">
    <h1 class="fn org">Drooling Dog Bar B Q</h1>
    ...
    
    <em>based on <span class="count">15</span> reviews</em>
    ...
    <strong>Price range:</strong> <span class="pricerange">$$</span>
  </div>
</div>
```

ali pa uporabite RDFa markup. Vsak format od navedenih deluje:

```
<div xmlns:v="http://rdf.data-vocabulary.org/#" typeof="v:Review-aggregate">
  <div rel="v:itemreviewed">
    <span typeof="v:Organization">
      <h1 property="v:name">Drooling Dog Bar B Q</h1>
      ...
      
      <em>based on <span property="v:count">15</span> reviews</em>
      ...
      <strong>Price Range:</strong> <span property="v:pricerange">$$</span>
    </span>
  </div>
</div>
```

³⁰ Yelp - <http://www.yelp.com/>

Z vključitvijo standardnih opomb na vaši strani, vaši podatki ne bodo samo strukturirani za Googlove rezultate iskanja, ampak tudi za vsako storitev ali orodje, ki podpira enak standard. Ko bodo strukturirani podatki vse bolj razširjeni na spletu, pričakujemo, da bomo našli veliko novih aplikacij, ki jih podpirajo in smo navdušeni nad možnostmi. Veliko prejšnjega dela na strukturiranih podatkih se je osredotočilo na razprave okoli kodiranja. V Google-u, kjer so zagovorniki za kodiranje mikroformatov, zagovarjajo različno RDF kodiranje in se zavzemajo za njihovo lastno kodiranje. Toda, po delu na Rich Snippets projektu smo spoznali, da lahko strukturirani podatki na spletu sprejmejo in bodo sprejeli več kodiranja, zato je dan poudarek na to s sprejemom obeh, kodiranje mikroformatov in RDFa kodiranje. Vsako od kodiranj ima svoje pluse in minuse, zato je ta debata lepa intelektualna vaja. Verjamemo, da je pomembno imeti skupen slovar: jezik objektnih tipov, lastnosti objekta in vrste lastnosti, ki omogočajo strukturirane podatke tako, da se razumejo z različnimi aplikacijami. Razpravljali smo kako obravnavati problem slovarja, in ugotovili, da smo potrebovali narediti naložbe. Google bo v sodelovanju z drugimi gostil slovar, ki ga lahko uporabljajo različne Googlove storitve in druge spletne strani.

5. Sklep

Web 2.0 tehnologije predlagajo lahko in enostavno uporabo tehnologij za gradnjo in izkoriščanje informacij, ter jo pretvarajo v znanje v obliki kolektivne človeške inteligence. Te so izkazale v praksi njihove uporabnosti in učinkovitosti, kot sredstvo ponujanja priložnosti za izmenjavo in obdelavo kolektivnega znanja mnogih pedagogov/avtorjev. Njihove pomanjkljivosti so pomanjkanje stroge logične podlage za sklepanje novih znanj, za dokazovanje kakovosti človeškega ustvarjenega znanja ter tehnologij za samodejno pridobivanje znanja, upravljanja in povezovanja. Semantični splet ponuja zapleteno razvito logiko tehnologij za obdelavo znanj v spletnem okolju, ampak ta tehnologija za samodejno pridobivanje znanja in vključevanje ni dovolj zrela, predvsem orodja, ki jih je težko uporabiti za domeno strokovnjakov ali uporabnikov in ne podpirajo sodelovanja upravljanja znanj. Naš glavni zaključek je, da bi povezovanje Web 2.0 aplikacij, Semantične spletne logike in znanja bilo treba načrtovati za novo in razburljivo okolje - socialni Semantični splet. V socialnem Semantičnem spletu se nekateri formalno predstavljeni deli človeškega znanja lahko kodirajo in utemeljijo glede na orodja Semantičnega spleta in jih je lahko tudi enostavno vzdrževati prek socialno usmerjenih tehnik Web 2.0. Semantične tehnologije bodo zagotovile učinkovito iskanje (prek semantičnih iskalnih pristopov) virov za učenje, informacije ali storitev, samodejno upravljanje znanja, natančnost in doslednost integracije. Medtem ko bodo Web 2.0 tehnologije zagotovile sodelovanje človeške udeležbe. Za Web 3.0 naj bi bila značilna profesionalizacija Web-a 2.0, ljudje in skupnosti pa naj bi postale še bolj vključene v izdelavo vsebin. Virtualni svet bo postal pravzaprav paralela realnemu, v njem pa bomo imeli več virtualnih jazov. Ena od ključnih zadev bo tudi, kako filtrirati vsebine, ki so zanimive za posameznike. E-trgovina se bo decentralizirala, ljudje pa bodo lahko kupovali preko blogov in spletnih strani s profili. Ključna beseda Web-a 2.0 je "iskanje", nova ključna beseda pa bo verjetno "odkritje". Splet bo o nas vedel zelo veliko, morda nas bo splet poznal celo bolj, kot poznamo sami sebe, saj se bodo različne podatkovne baze začele pogovarjati med seboj, kar pomeni, da se bodo začele informacije med seboj povezovati in tvoriti neko celoto. Kljub vsem optimističnim napovedim pa je treba biti pozoren tudi na negativne strani. Pojavilo se bo ogromno vprašanj in dilem, ki se bodo navezovala na zasebnost posameznika. Pomisliti je treba tudi na to, koliko vseh ljudi, ki sodelujejo v raznih spletnih skupnostih, resnično prispeva in ustvarja določene vsebine, koliko pa je samo pasivnih opazovalcev. Tudi oglaševanje se bo spremenilo, oglaševalcem se bodo odpirale vedno nove možnosti, kar zagotovo ne bo odgovarjalo vsem oglaševalcem zaradi vse večje medsebojne konkurence.

Priloge

Seznam slik

SLIKA 1: PRIKAZ SESTAVIN WEB 2.0.....	9
SLIKA 2: ELEMENTI SEMANTIČNEGA SPLETA.....	17
SLIKA 3: PRIKAZ DELOVANJA WEB 3.0	22
SLIKA 4: SEMANTIČNI ISKALNIK BING	26
SLIKA 5: PRIKAZ ISKANJA V BING ISKALNIKU.....	27
SLIKA 6: SEMANTIČNI ISKALNIK SWOOGLE.....	28
SLIKA 7: DIAGRAM MIKROFORMATOV	30
SLIKA 8: PRIKAZ DELOVANJA RICH SNIPPETS-A.....	41

6. Literatura

- [1] Wikipedia contributors. (2011) Wikipedia, The Free Encyclopedia. Dostopno na: <http://sl.wikipedia.org/wiki/Splet> (ogled 20.3.2011)
- [2] Wikipedia contributors. (2011) Wikipedia, The Free Encyclopedia. Dostopno na: http://hr.wikipedia.org/wiki/Web_2.0 (ogled 21.3.2011)
- [3] James Governor, Dion Hinchcliffe and Duane Nickull, Web 2.0 Architectures. O'Reilly Media, 2009
- [4] Wikipedia contributors. (2011) Wikipedia, The Free Encyclopedia. Dostopno na: http://sl.wikipedia.org/wiki/Splet_2.0 (ogled 25.3.2011)
- [5] IT PRO Fit for business (2010), Q&A: Conrad Wolfram on communicating with apps in Web 3.0, Dostopno na: <http://www.itpro.co.uk/621535/q-a-conrad-wolfram-on-communicating-with-apps-in-web-3-0> (ogled 10.4.2011)
- [6] Morpox Blog (2007), Web 3.0 – Semantični splet, Dostopno na: <http://blog.morpox.si/2007/uporabniske-izkusnje/web-30-%E2%80%93-semanticni-splet/> (ogled 10.4.2011)
- [7] Wikipedia contributors. (2011) Wikipedia, The Free Encyclopedia. Dostopno na: http://en.wikipedia.org/wiki/Semantic_Web (ogled 10.4.2011)
- [8] Wikipedia contributors. (2011) Wikipedia, The Free Encyclopedia. Dostopno na: http://it.wikipedia.org/wiki/Web_3.0 (ogled 11.4.2011)
- [9] Toby Segaran, Colin Evans, and Jamie Taylor, Programming the Semantic Web, O'Reilly Media, 2009
- [10] Article for free (2009), Web 3.0 Technology, Dostopno na: <http://www.articleforfree.com/Articles/Technology/Web-3-0-Technology-380190732.aspx> (ogled 11.4.2011)
- [11] Basant Narayan, Cloud Computing (2009), Web 3.0: The Intellignet Web, Dostopno na: <http://www.techno-pulse.com/2009/11/web-30-intelligent-web.html> (ogled 2.5.2011)
- [12] Wikipedia contributors. (2011) Wikipedia, The Free Encyclopedia. Dostopno na: http://en.wikipedia.org/wiki/Bing_%28search_engine%29 (ogled 5.5.2011)
- [13] Wikipedia contributors. (2011) Wikipedia, The Free Encyclopedia. Dostopno na: <http://en.wikipedia.org/wiki/Swoogle> (ogled 7.5.2011)
- [14] Microformats, Dostopno na: <http://microformats.org/about> (ogled 15.5.2011)
- [15]] Wikipedia contributors. (2011) Wikipedia, The Free Encyclopedia. Dostopno na:

<http://en.wikipedia.org/wiki/RDFa> (ogled 20.5.2011)

[16] Taskboy :: An Exploration Blog (2010), [Microformats, RDFa and their semantic uses](#), Dostopno na: <http://taskboy.com/blog/?bid=1217> (ogled 25.5.2011)

[17] Google Webmaster central blog (2009), Introducing Rich Snippets, Dostopno na: <http://googlewebmastercentral.blogspot.com/2009/05/introducing-rich-snippets.html> [ogled 26.5.2011]

Joey Bernal, Web 2.0 and Social Networking for the Enterprise, IBM Press, 2010

Mark Watson, Scripting Intelligence; Web 3.0 Information Gathering and Processing, Apress, 2009

Daniel Harris, Web 2.0 Evolution into The Intelligent Web 3.0, 2008