

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Igor Medved

Organizacija informacij na spletu: primer folksonomij

Diplomsko delo

Ljubljana, 2009

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Igor Medved

Mentorica: doc. dr. Tanja Oblak Črnič
Somentor: doc. dr. Gregor Petrič

Organizacija informacij na spletu: primer folksonomij

Diplomsko delo

Ljubljana, 2009

ORGANIZACIJA INFORMACIJ NA SPLETU: PRIMER FOLKSONOMIJ

V realnem svetu prevladujoč, ekspertni način organiziranja informacij v obliki hierarhij, se je na spletu izkazal za neustreznega. Svetovni splet je preobsežen, da bi ga lahko strokovnjaki sproti in natančno klasificirali. Kot ena izmed alternativ taksonomijam so se izkazale folksonomije, ki omogočajo uporabnikom (samo)urejanje spletnih virov s pomočjo prostih oznak [tagov] za ponovni kasnejši priklic. Za razliko od taksonomij gre pri folksonomijah za urejanje informacij od spodaj navzgor, brez ekspertov in hierarhije. To sledi participatorni logiki spleta 2.0, kjer uporabniki sami ustvarjajo vsebine, ki jih lahko s pomočjo družbenih oznak tudi sami klasificirajo. Širši pregled trenutnega razvoja folksonomij dopolnjujem z njihovo potencialno pomembno vlogo v semantičnem spletu. Družbene oznake trenutno delujejo zgolj na opisni ravni, z dodano semantično plastjo v obliki natančne definicije oznake, pa bi na ta način klasificirane informacije lahko razumeli tudi računalniki. S tem bi bilo omogočeno bolj učinkovito iskanje informacij, ne le na opisni, ampak tudi na pomenski ravni.

Ključne besede: družbeno označevanje, folksonomija, organizacija informacij, splet 2.0, semantični splet.

THE ORGANIZATION OF INFORMATION ON THE WEB: THE CASE OF FOLKSONOMIES

In the real world, prevalent hierarchical organization of information proved to be inadequate on the web. Web is overloaded with information and experts cannot classify them on time and with needed precision. Folksonomies have proved to be one alternative to taxonomies on the web as they enable users categorizing web resources by means of tags (freely chosen keywords) for later retrieval. In comparison to taxonomies, folksonomies enable classification of information from the bottom up, without the experts and the hierarchy. This follows the logic of participatory web 2.0, where the users generate content, which they can also classify by themselves. The broad overview of current development of folksonomies is complemented with the view of their potentially important role in the semantic web. Tagging currently occurs only on a descriptive level, but with the added semantic layer in the form of the exact definition behind the tag, information classified in this way could also be understood by the computer. That would enable a more effective retrieval of information not only on the syntax but also on the semantic level.

Key words: social tagging, folksonomy, organization of information, web 2.0, semantic web.

Kazalo

1 Uvod	7
2 Splet kot informacijski sistem	10
2.1 Informacijska družba.....	10
2.2 Splet kot informacijski sistem.....	11
2.3 Hierarhični vs. hipertekstovni klasifikacijski sistemi	14
2.4 Klasificiranje informacij v spletnem okolju.....	16
2.4.1 Spletni imeniki	16
2.4.2 Fasetna klasifikacija	18
2.4.3 Spletni iskalniki	19
2.4.4 Osebni zaznamki	20
2.4.5 Družbeni zaznamki - folksonomije.....	21
2.4.6 Semantični splet	24
3 Hipertekstovnost spleta	26
3.1 Kultura hiperteksta.....	26
3.2 Splet 2.0	27
3.3 Divja hipertekstovnost	30
3.4 Hipertekstovnost spleta 2.0	32
4 Tehnična raven folksonomij in družbenega označevanja	36
4.1 Pojmi in ključne besede, oznake in metapodatki	36
4.2 Proces označevanja	37
4.3 Funkcije oznak	39
4.4 Učinek dolgega repa.....	40
4.5 Tipologija folksonomij.....	41
4.6 Slabosti družbenega označevanja.....	45
4.7 Motivacijski dejavniki označevanja.....	46
4.8 Vizualizacije oznak	48
5 Semantična raven folksonomij kot informacijskega sistema	52
5.1 Metapodatki semantičnega spleta	52
5.2 Semantične folksonomije	53
5.3 Družbena raven folksonomij	59
5.3.1 Folksonomije – virtualne skupnosti?	60
5.3.2 Folksonomije in pametne množice	62

6. Sklep	65
Literatura.....	67

Kazalo slik

Slika 2.1: Taksonomija v tipični drevesni strukturi	12
Slika 2.2: Povezanost enot v spletnem hierarhičnem sistemu	14
Slika 2.3: Odprava hierarhije in seznama, ker povezave zadoščajo.....	14
Slika 2.4: Primerjava starega in novega sistema priklica informacij	15
Slika 2.5: Spletna stran DMOZ	17
Slika 2.6: Fasetna klasifikacija.....	18
Slika 2.7: Primer spletnega fasetnega klasifikacijskega sistema	19
Slika 2.8: Prikaz označevanja spletnega vira	23
Slika 2.8: Prikaz označevanja spletnega vira	23
Slika 3.1: Seznam priporočil na Amazon.com na podlagi nakupov ostalih uporabnikov	33
Slika 3.2: Primerjava starejše in novejše verzije vpisa pojma Web 2.0 (Wikipedia):	34
Slika 3.3: Realizacija hipertekstovnih prizadevanj v spletnih sistem	35
Slika 4.1: Kognitivni proces označevanja	38
Slika 4.2: Kognitivni proces označevanja in kategorizacije	39
Slika 4.3: Porazdelitev oznak v obliki dolgega repa	41
Slika 4.4: Široka in ozka folksonomija	42
Slika 4.5: Pomoči pri označevanju s priporočenimi in priljubljenimi oznakami	43
Slika 4.6: Kategorizacija sistemov družbenega označevanja	48
Slika 4.7: Oblak oznak	49
Slika 4.8: Vizualizacija najbolj pogosto uporabljenih besed Hillary Clinton	49
Slika 4.9: Vizualizacija najbolj uporabljenih besed Baracka Obame	50
Slika 4.10: Aplikacija Worldexplorer	50
Slika 5.1: Twineov bookmarklet	54
Slika 5.2: Shranjen izdelek na Twine.com	55
Slika 5.3: Semantični graf Twine za knjigo Jurski park	56
Slika 5.4: Zigtag	57

1 Uvod

Zelo specifičen problem sodobne družbe, ki je danes že v celoti odvisna od spletnih tehnologij, je informacijska zasičenost oziroma preobilje informacij. A ta problem je z nami pravzaprav že vse od takrat, ko se je človek zavedel pomena znanja in ga poskušal z učinkovitim načinom upravljanja informacij tudi nasloviti. V širšem zgodovinskem kontekstu je kraljevala zlasti Aristotlova taksonomična, centralno urejena hierarhična ureditev, ki so jo med drugimi vzeli za svojo tudi Carl Linné (v biologiji), John Dewey (v bibliotekarstvu) in Microsoft (s sistemom map v okolju Windows). Zdi se torej, da je bila organizacija informacij vedno tesno povezana z razvojem znanosti, filozofije ali novih tehnologij in napredek na enem izmed teh treh področij je pogosto porajal kakšno novo različico upravljanja z njimi.

Tudi v digitalnem okolju spleta je obvladovanje informacij že vse od začetka spleta postalo eno izmed njegovih najbolj integralnih, propulzivnih in dobičkonosnih področij. Spletna giganta Yahoo in Google sta svoj imperij zgradila predvsem kot posrednika (tujih) informacij; Yahoo je s spletnim imenikom Yahoo!Directory uporabnikom ponudil **klasično taksonomično ureditev** informacij v obliki kategorij, spletni iskalniki na čelu z Googlom pa so bolje izkoristili potencialne digitalnega okolja, predvsem hipertekstovnosti kot načina strukturiranja informacij, in nadomestili hierarhično, centralno ureditev. Glavne pomanjkljivosti taksonomičnega klasificiranja podatkov v spletnem okolju se nanašajo predvsem na njegovo urejanje. Tudi velika ekipa urednikov ne more vzdrževati kataloga spletnih strani, ker je splet preprosto preobsežen. Poleg tega je kategorizacija informacij izrazito subjektivno področje, kar velja tako za urednike, ki se morajo pri kategorizaciji odločiti za eno izmed možnih in ustreznih kategorij oziroma podkategorij, kakor tudi za uporabnike, ki imajo različne predstave o tem, kje iskati določeno informacijo. S tem sta odpravljena ključna razloga za uporabo taksonomij v spletnem okolju, ki sta v knjižničnem okolju sicer glavni vrlini; celovita obravnava področja in preprost dostop do informacij.

Spletni iskalniki hierarhijo, kategorije in vlogo urednikov, kot že rečeno, odpravljajo. Delujejo tako, da najprej iz seznama vseh zapisanih mest izberejo vsa tista spletna mesta, ki vsebujejo iskani izraz, v drugi fazi pa ta spletna mesta s pomočjo programskega algoritma razvrstijo po pomembnosti. Toda kljub danes nepogrešljivi vlogi iskalnikov rezultati iskanja pogosto niso zadovoljivi ali pa se taki zdijo le na prvi pogled. Tudi ob na videz uspešnem iskanju v resnici ne vemo, koliko dobrih in kakovostnih informacij spletni iskalnik ni prikazal med zadetki. Večina

objavljenih podatkov je namreč nestrukturiranih in čeprav je ravno odsotnost strukture eden izmed glavnih razlogov za eksponentno rast spleta, danes vse bolj otežuje iskanje in najdbo relevantnih oziroma najboljših zadetkov.

Na navidez paradoksalno rešitev preobilja informacij v spletnem okolju nakazuje Weinberger (2007, 192), ko pravi, da je ključ do uspeha v ustvarjanju še več informacij in v izgubi nadzora nad njimi. Ideja je preprosta, več kot bomo imeli informacij o informacijah, lažje jih bomo našli. Kot ena bolj ustreznih rešitev za osebno in hkrati družbeno indeksiranje ter shranjevanje digitalnih informacij so se izkazale **folksonomije, neekspertni, uporabniški sistemi klasificiranja informacij**, ki so se pojavili v začetku tega tisočletja. Folksonomije so v osnovi lastni, uporabniški katalogi spletnih strani in so vidni tudi drugim uporabnikom, saj so shranjeni na spletu. Niso urejeno hierarhično v obliki kategorij, ampak temeljijo na uporabi oznak, prostih ključnih besed, s katerimi lahko sami opišemo vsebino spletnega vira ali naš odnos do njega. Analogija z realnim svetom je preprosta: spletne folksonomije so kakor naše knjižne police, kjer lahko knjige razvrstimo po lastnih parametrih, bodisi po abecedi, debelini knjige, avtorju, založbi, barvi ali pa izrazito subjektivnem parametru, kakovosti.

Za razliko od taksonomij gre pri folksonomijah torej za urejanje informacij od spodaj navzgor, za obliko ročnega indeksiranja vsebin brez pomoči strokovnjaka. Njihovo mesto je v spletu 2.0 oziroma v participatornem spletu, kjer personalizirano urejanje informacij na nek način zvesto sledi njegovi ideji : **vsebino spleta ustvarjajo uporabniki, zato jo naj sami klasificirajo in tudi ovrednotijo**. Ob dovolj velikem številu uporabnikov lahko folksonomije, kot je denimo Delicious, s svojimi rezultati uspešno konkurirajo tudi spletnim iskalnikom.

Oblakova in Petrič (2005) svetovni splet v grobem razdelita na informacijsko in komunikacijsko raven, pri čemer se prva nanaša na prostor za skladiščenje znanja, dokumentov ali informacij, kakor tudi na učinkovite načine priklica teh dokumentov s strani uporabnikov. Komunikacijska raven pa obravnava vse tiste elemente, ki omogočajo (ne)posredno človeško komuniciranje prek spleta. V nalogi se bomo osredotočili predvsem na informacijsko raven spleta s poudarkom na folksonomijah kot novem načinu klasificiranja informacij. V prvem delu naloge tako najprej opredelimo informacijsko družbo in vlogo informacije v njej ter svetovni splet, ki je s tehnologijo hiperteksta želel nasloviti predvsem problem učinkovitega priklica informacij. Skozi primerjavo klasičnega, taksonomičnega klasifikacijskega sistema s hipertekstovnim prikažemo njegove glavne omejitve pri uporabi v digitalnem okolju spleta in kot eno izmed možnih alternativ priklica informacij (poleg najbolj priljubljene – spletnih iskalnikov) predstavimo folksonomije. Tu skušamo predvsem odgovoriti na vprašanje, zakaj so folksonomije

postale med uporabniki tako priljubljene. Eden izmed glavnih razlogov za to je njihova preprosta uporaba in odsotnost kakršnih koli pravil v obliki kategorij, hierarhij ali klasifikacijskih shem. To pa po drugi strani pomeni tudi slabost, tudi družbene oznake v folksonomijah namreč ne odražajo (točne) semantične vsebine označenih predmetov, pač pa zgolj uporabnikovo subjektivno idejo o tem, kaj informacija predstavlja. Prav na ta problem se zato osredotočamo v zadnjem delu naloge, ko opišemo novi, prihajajoči splet, ki bo nasledil splet 2.0, ki bo po Weschu (2007) namenjen učenju stroja [*teaching the machine*]. Prihaja namreč semantični splet, v katerem bodo podatki dobili pomen, tako da jih bodo razumeli tudi računalniki. Kljub priročnosti in trenutni priljubljenosti folksonomij postavljamo hipotezo, da lahko njihova moč in učinkovitost v celoti pride do izraza šele v naslednji generaciji spleta v obliki **semantičnih folksonomij**, kjer bo vsaka oznaka opremljena s točno definicijo. Priklic informacij bo končno lahko potekal tudi na semantični ravni. Če se tako povrnemo na uvod uvoda, verjetno se nam ponovno obeta velika sprememba v organizaciji informacij, ki bo tokrat povezana tako s filozofskim, tehnološkim kot znanstvenim preskokom.

2 Splet kot informacijski sistem

2.1 Informacijska družba

Vannevar Bush, ameriški vojaški znanstvenik in vizionar, je v 40. letih prejšnjega stoletja opazil, da skupna znanstvena dejavnost po svetu tako hitro ustvarja znanje, da je tekoče spremljanje vseh novosti izpostavil kot resen problem prihodnosti (Rheingold 2004, 209). Z vse večjim vlaganjem v znanost, porastom števila znanstvenikov ter z razvojem tehnike in računalništva se je kmalu po 2. svetovni vojni pojavila potreba po oblikovanju učinkovitih, računalniško podprtih informacijskih sistemih, ki bi lahko to novo ustvarjeno znanje zbrali, obdelali in pripravili za učinkovito nadaljnjo rabo. Tako smo 2. polovici prejšnjega stoletja vstopili v dobo informacij oziroma v obdobje **informacijske družbe**, v središču katere je bila nova informacijska tehnologija.

Naisbitt (glej Roszak 1994, 22) opiše prehod iz industrijske v informacijsko družbo z novim ekonomskim redom, v katerem podobno kot smo nekoč masovno proizvajali avtomobile sedaj masovno proizvajamo informacije in znanje. Toda Roszak (1994, 22-23) meni, da informacijska tehnologija ne nadomešča industrijskega sistema, ki je bil prav tako vedno odvisen od znanja in inovacij, temveč se nanj še vedno opira oziroma se z njim prekriva. Tudi Castells (1996, 204) zelo podobno ne razmišlja o delitvi na industrijsko in postindustrijsko ekonomijo, ampak razlikuje med dvema proizvodnjama, kmetijsko-industrijski in storitveni, ki pa obe temeljita na znanju in inovacijah. Glavna razlika se nanaša na proces **informatizacije** oziroma na revolucionarni razvoj informacijskih tehnologij in na njihovo vključevanje v sfero družbene in gospodarske aktivnosti, kar je omogočilo oblikovanje infrastrukture za vzpostavitev globalnega gospodarstva. Prav informacijska tehnologija in pojav računalnikov je za Castellsa (1996, 61) jedro trenutne transformacije družbe, ki jo opiše skozi **paradigmo informacijske tehnologije** in zajema naslednje ključne značilnosti:

1. informacija je njen temeljni material, kateri se prilagaja tehnologija in ne obratno, kot je bilo značilno za prejšnje tehnološke revolucije;
2. informacije so integralni del človekove dejavnosti;
3. kateri koli sistem, ki uporablja nove informacijske tehnologije, zaradi kompleksnosti interakcij deluje po principu mrežne logike, kar zagotavlja določeno strukturo, ki pa ni omejujoča in dovoljuje nadaljnje inovacije;
4. nova tehnološka paradigma temelji na fleksibilnosti, ki poleg procesov transformira oziroma modificira tudi same organizacije in institucije;
5. konvergenca različnih tehnologij v visoko integriranem sistemu narašča in sproti

odpravlja zastarele tehnologije, predvsem pa tiste, ki niso združljive z novimi.

Če je Gutenbergova iznajdba tiskarskega stroja v 15. stoletju omogočila, da je informacija postala pravica slehernika in ne zgolj privilegij peščice, tudi danes sodobna informacijska tehnologija spreminja družbo in naš odnos do informacije. Kot ključno v procesu novega pojmovanja informacije Roszak (1994, 11-20) izpostavi Shannonovo *Matematično teorijo komuniciranja*, po kateri informacija ni več povezana s semantično vsebino trditve, pač pa postane povsem **kvantitativno** sredstvo izmenjave komunikacij. Posledično je informacija ločena od svojega konvencionalnega pomena, njena definicija pa postane tako širša in manj specifična. Lahko je karkoli, kar je posredovano in (lahko zgolj metaforično) razumljeno kot sporočilo, tudi na primer živčni impulz. S tem so informacije izgubile svojo vrednost, saj so bile z marginaliziranjem pomena vse zreducirane na isto raven. Ta radikalna sprememba v dojemanju informacije je omogočila posvojitve računalnika v polje razumevanja informacij in hkrati kot **kvalitativni** element od nje oddaljila človeka (kot njenega oddajnika ali prejemnika). Roszak (1994, 87) zato zavrača vsakršen bolj ambiciozen poskus definicije informacije in jo sam opredeli zgolj kot diskreten klobčič dejstev, včasih uporaben, včasih trivialen, ki ni nikoli produkt misli.

V informacijski dobi osebnih računalnikov naj bi se, kot je že v 60. letih prejšnjega stoletja napovedal McLuhan (glej Roszak 1994, 163), zgodila demokratična renesansa – takojšnji dostop do informacij naj bi povzročil večjo (politično) vpletenost ljudi. McLuhan je imel v mislih predvsem televizijo, toda zgodilo se je ravno nasprotno, nihče ni namreč predvidel situacije, ko lahko obstaja preveč informacij. Problem ni bil več njihov primanjkljaj, pač pa preobilje oziroma poplavljanje javnosti s tolikšno količino informacij, da njihovo učinkovito procesiranje ni bilo več mogoče. Kot medij prihodnosti so televizijo nato zamenjali osebni računalniki in nove hipertekstovne omrežne tehnologije, ki bi na ravni informacijskega sistema omogočale enostaven in učinkovit priklic informacij, kakor tudi vrnilo upanje po večji politični participaciji posameznika v družbi.

2.2 Splet kot informacijski sistem

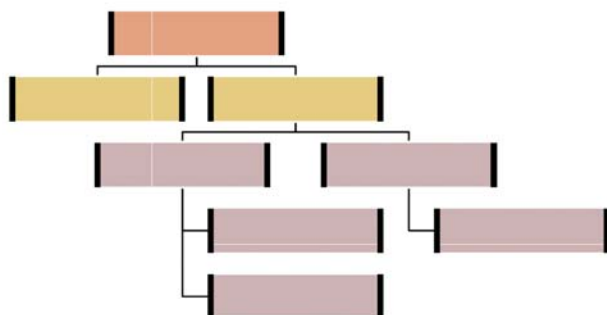
Rojstvo svetovnega spleta¹ (angl. World Wide Web) je tesno povezano prav z nezadovoljivo organizacijo informacij. Raziskovalec Tim Berners Lee je konec 80. let prejšnjega stoletja,

¹ Svetovni splet nam je uporabnikom viden kot zbirka digitalnih dokumentov v obliki spletnih strani, ki so med seboj lahko povezane in do katerih dostopamo s pomočjo brskalnika. (Petrič 2003).

nezadovoljen z učinkovitostjo informacijskega sistema in razdrobljenostjo informacij na inštitutu CERN, predstavil idejo informacijskega sistema, ki bi temeljil na hipertekstovnosti in bi omogočal obstoj povezav med besedilnimi deli. Zaradi že obstoječega decentraliziranega internetnega omrežja med seboj povezanih računalnikov je lahko Berners Leejeva ideja leta 1993 zaživela in svetovni splet se je z inštituta naglo razširil po svetu ter tako upravičil svoje globalno poimenovanje.

S hiperpovezavami hipertekstovni informacijski sistemi odpravljajo tradicionalne hierarhične modele organiziranja informacij oziroma **taksonomije**, ekspertne sisteme razvrščanja informacij s strogo centralizirano strukturo, katerih izvor sega že v čas **Aristotela**, ki je vsa živa bitja razvrstil po vrsti na rastline in živali, slednje pa nato v kategorije po gibanju po kopnem, zraku ali vodi. Njihova struktura je prikazana na naslednji sliki, na vrhu je glavna kategorija, ki se nato veji na podkategorije te pa naprej na naslednje podkategorije.

Slika 2.1: Taksonomija v tipični drevesni strukturi



Razvrstilni sistem v obliki drevesa je v 18. stoletju nato s klasifikacijo živih bitij izpopolnil biolog **Carl Linné**, najbolj uporabljena taksonomija pa je verjetno **Deweyjeva decimalna klasifikacija** iz leta 1876, ki se je z delnimi popravki ohranila vse do danes.² Ameriški bibliotekar Melvin Dewey, je takrat celotno človeško znanje razdelil v naslednjih deset glavnih kategorij (OLCL 2003):

000 Splošna skupina

100 Filozofija. Psihologija

200 Religija

² Deweyjeva klasifikacija je očitno pristranska in zastarela, vendar pa se je pravzaprav ne da več "popraviti". Vezana je namreč na resnični svet, na knjigo, ki je lahko le na eni polici in v eni kategoriji hkrati. Popravljanje Deweyjevih zmot in premeščanje knjig v nove kategorije na nove police bi za knjižnice pomenil prevelik logistični zalogaj, ki poleg tega najbrž nikoli ne doživel konca, saj se znanje in vedenje nenehno spreminja. (Weinberger 2007).

300 Družbene vede
400 Jezikoslovje
500 Naravoslovne vede. Matematika
600 Tehnologija (Uporabne vede)
700 Umetnost
800 Književnost
900 Geografija. Zgodovina

Taksonomični, hierahični sistemi razvrščanja informacij so bili logična izbira v realnem okolju, **digitalno okolje** pa je zahtevalo odpravo nekaterih njihovih omejitev. Organizacija znanja in idej, značilna za digitalno okolje, se namreč močno razlikuje od organizacije fizičnih predmetov:

- v realnem svetu je predmet lahko umeščen le v eno kategorijo in podkategorijo, saj razdelitev na denimo dva predalčka ni mogoča. V digitalnem svetu pa te omejitve ni, pojmi so lahko poljubno razvrščeni v desetine ali stotine kategorij;
- v realnem svetu veliko ljudi uporablja le eno razvrstitveno drevo, digitalno okolje pa omogoča, da si lahko vsaka oseba ustvari svoje drevo;
- v realnem svetu ima oseba, ki ima v lasti informacije, navadno v lasti tudi drevo, po katerem so informacije organizirane (muzeji, arhivi, knjižnice). V digitalnem svetu pa uporabniki organizirajo informacije, ki večinoma niso njihove. (Weinberger 2007, 100-106).

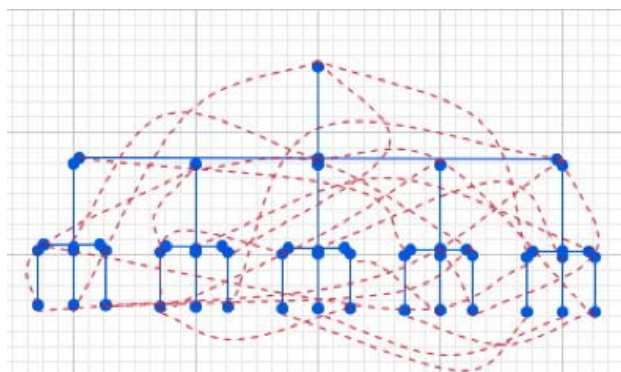
Za taksonomije kot prevladujoči sistem organizacije informacij v realnem svetu Quinterellijeva (2005) pravi, da morajo biti poleg usposobljenih strokovnjakov in izobraženih uporabnikov, ki vedo, kje najti želeno informacijo, za uspešno rabo izpolnjeni tudi naslednji pogoji:

- omejen korpus
- formalne, jasno določljive kategorije
- stabilne in omejene enote
- jasne meje med enotami

Slabosti taksonomičnega razvrščanja vsebine so še najbolj očitne v spletnem okolju; korpus informacij je neskončno velik, uporabniki so med seboj nepovezani in prihajajo iz različnih družbenih okolij ter pogosto niti ne vedo natančno, kaj iščejo. Poleg tega, da je uporabnik taksonomij primoran sprejeti ekspertno določen pomen predmeta in ga kot takega tudi najti, Shirky (2005) ugotavlja, da je svetovni splet zaradi svoje hipertekstovnosti taksonomično izjemno težko obvladljiv. Slednje grafično ponazori z naslednjima skicama. Prva ponazarja

hierarhični sistem spletne organizacije informacij, ki je prepleten s številnimi povezavami med posameznimi enotami v različnih kategorijah in podkategorijah.

Slika 2.2: Povezanost enot v spletnem hierarhičnem sistemu

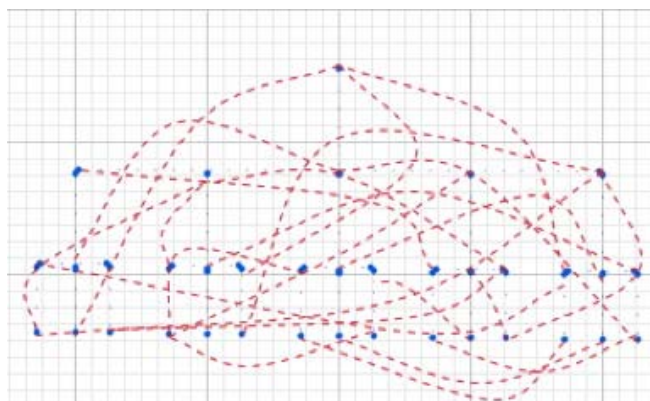


Vir: Shirky (2005).

Ob komentarju na skico, ki se nanaša na Yahoojev spletni imenik, Shirky (2005) pravi: “Na koncu so pri tem razvoju zgrešili pravi konec. Če obstaja dovolj povezav, hierarhija ni več potrebna. Ni police. Ni seznama. Povezave same so dovolj.”

Slednjo ugotovitev prikaže z drugo risbo, kjer je hierarhija odpravljena, ker ta zaradi hiperprepletenosti enot preprosto ni več potrebna:

Slika 2.3: Odprava hierarhije in seznama, ker povezave zadoščajo



Vir: Shirky (2005).

Ta struktura je značilna predvsem za spletne iskalnike, ki so prvi korak k odpravi hierarhij v razvoju sistematizacije priklica informacij v spletnem okolju.

2.3 Hierarhični vs. hipertekstovni klasifikacijski sistemi

Dreyfus (2001, 94) je do odprave hierarhičnih modelov organizacije informacij izrazilo

skeptičen in črnogled, saj meni, da bi se z digitalizacijo posledično morali odpovedati tudi zmožnosti, da bi našli večino informaciji, ki jih potrebujemo. Ker je internet anarhično strukturiran in brez neke centralne avtoritete, so ljudje odvisni od spletnih iskalnikov, da za njih najdejo informacijo, ki jo potrebujejo. Spletni iskalniki pa so računalniški programi, ki so sposobni obdelati informacije le na sintaktični ravni, ne pa tudi na semantični. Dreyfus (2001, 94) opozarja, da se lahko, v kolikor sprejmemo odvisnost od spletnih iskalnikov, poslovimo od velike večine arhiviranega znanja na spletu. Nekoliko nostalgичno Dreyfus (2001, 11) zato organizacijo in priklic informacij razmeji na kulturo starih knjižnic in na kulturo hiperteksta.

Slika 2.4: Primerjava starega in novega sistema priklica informacij

KULTURA STARIH KNJIŽNIC	KULTURA HIPERTEKSTA
Klasifikacija	Diverzifikacija
a. stabilna	a. fleksibilna
b. hierarhično organizirana	b. 1-nivojska
c. določena s specifičnim interesom	c. dovoljuje vse možne asociacije
Pazljiva selekcija	Dostop vsemu
a. kvaliteta enot	a. inkluzivnost
b. avtentičnost besedila	b. dostopnost besedil
c. odstrani star material	c. shrani vse
Stalne zbirke	Dinamične zbirke
a. ohranitev fiksnih besedil	a. intertekstualna evolucija
b. brskanje po zanimanju	b. igrivo brskanje

Vir: Dreyfus (2001, 11).

Prva struktura ima izrazito hierarhično strukturo in je značilna predvsem za knjižnice. Pri gradnji in uporabi tovrstnega sistema sodelujejo najprej **strokovnjaki**, ki v fazi klasifikacije pripravijo hierarhijo oziroma seznam kategorij, **knjižničarji**, ki opravijo selekcijo in knjigo ali predmet glede na njegovo vsebino umestijo v eno izmed vnaprej določenih kategorij ter v končni fazi **uporabniki**, ki v iskanju določenega naslova brskajo po teh kategorijah.

Za strukturo hiperteksta, kamor spadajo poleg spletnih iskalnikov tudi folksonomije, pa je značilna popolna sploščenost oziroma odsotnost kategorij in hierarhije, zato sta **vlogi eksperta in selektorja** odpravljeni oziroma **združeni v uporabniku**. Na tem mestu se lahko vprašamo, kakšna je usoda najbolj tradicionalnega vira informacij – javnih knjižnic v novem digitalnem redu. Spremenjen namreč ni le način klasificiranja informacij, pač pa tudi njihov priklic. Kljub temu, da knjižnice pogosto pojmujejo zgolj kot prostor za skladiščenje knjig, Roszak (1994,

182) meni, da bi lahko predstavljale vezni člen med računalniškimi sistemi in družbo, saj imajo knjižničarji neprimerljivo več znanja od povprečnega uporabnika, tudi na računalniškem področju. Vsaj tako pomemben kot kvantitativni vidik informacij, je tudi njihova **kakovostna raba**, česar se uporabniki sicer dobro zavedajo. Po raziskavi Pew Instituta (Horrigan 2006) o tem, zakaj ljudje iščejo znanstvene informacije po spletu, velika večina (71 %) izpostavlja priročnost medija, le 13 % pa jih meni, da so te informacije tudi bolj kredibilne od ostalih virov informacij.

Roszak (1994, 182) v tem kontekstu izpostavi tudi problematiko privatiziranja procesa iskanja informacij³, s čimer se ne dela škoda le javnim knjižnicam, ki bi zaradi zagotavljanja točnih in brezplačno dostopnih informacij vsem državljanom (tudi revnim in zapostavljenim) morale imeti posebno mesto v vsaki demokratični družbi, ampak tudi sami kakovosti informacij. V informacijski dobi zato ni ključna tehnologija, ampak predvsem **družbena organizacija te tehnologije**. Informacija mora ostati javna dobrina in mora biti brezplačna, ker ljudje želijo, da takšna ostane in kot je vedno bila v javni knjižnici. Svetovni splet kot informacijski sistem v tej luči nadaljuje tradicijo zagotavljanja brezplačnih informacij ljudem in je odlično dopolnilo javnim knjižnicam, zaradi izpostavljenih kakovostnih razlik pa nikakor ne tudi njihov nadomestek. Občutek, da sami prek spleta in spletnih iskalnikov lahko najdemo ustrezne in kakovostne informacije je namreč zelo varljiv; Swanson (v Dreyfus 2001, 23) opozarja, da bo v dovolj velikem sistemu skoraj vsak priklic informacij vrnil nekaj uporabnih dokumentov. Napačna pa je predpostavka, da je zgolj zato tak sistem priklica informacij uspešen. Ne vemo namreč, koliko *boljših* dokumentov je sistem zgrešil.

2.4 Klasificiranje informacij v spletnem okolju

Oglejmo si najbolj pogosto uporabljene sisteme razvrščanja informacij v spletnem okolju, pri čemer začenjamo z najbolj hierarhično obliko, spletnimi imeniki.

2.4.1 Spletni imeniki

V svetovnem spletu se povsem klasični, taksonomični način klasificiranja podatkov pojavlja v obliki **spletnih imenikov**, ki so spremljevalci svetovnega spleta skoraj od njegovega nastanka. Yahoo! Directory, prvi spletni imenik, je bil ustanovljen že leta 1994. Gre za katalog spletnih mest, ki ga ureja ekipa urednikov, ki obiskujejo in vrednotijo spletna mesta ter jih nato

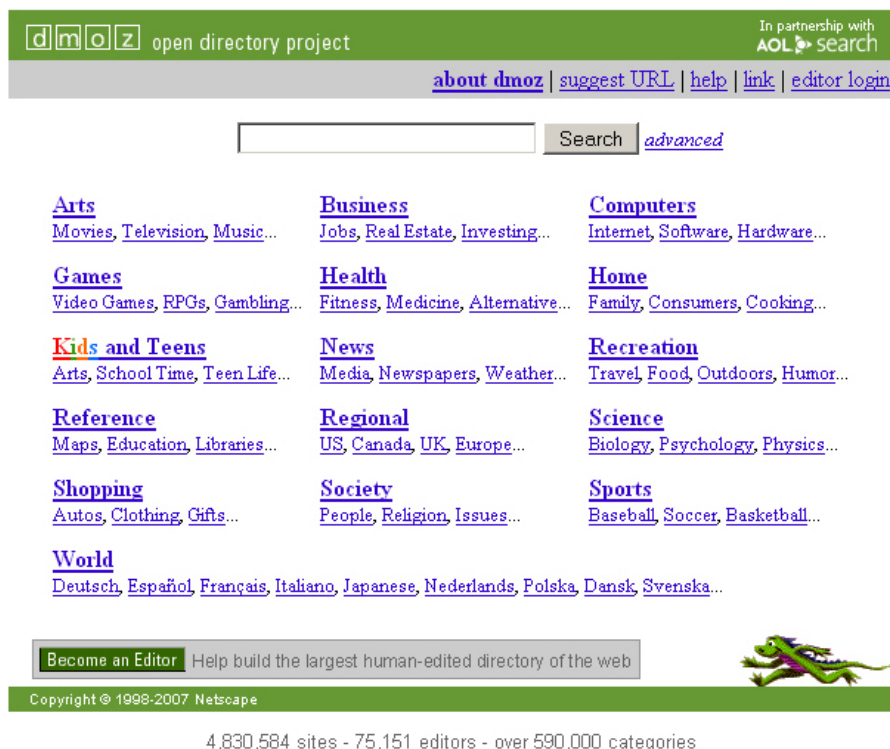
³ Roszak je imel pri tem v mislih predvsem podjetja in lastnike baz podatkov, ki so javnosti prodajale podatke in informacije, ki so bile tako ali tako last javnosti oziroma javna dobrina.

razvrščajo v predmetno označene kategorije in podkategorije. Skupaj ima imenik 14 glavnih kategorij.

Zaradi že omenjene drugačne narave hipertekstovnega okolja, kjer ni strokovnjakov, ki bi pomagali uporabniku najti, kar ta išče, imajo spletne taksonomije precej pomanjkljivosti. Med njimi je z vidika uporabniškega priklica informacij najbolj problematično semantično reduciranje spletnega mesta na en pomen oziroma njegovo umeščanje v le eno kategorijo. Uporabniki namreč iščejo spletna mesta po različnih poteh.⁴

Omenimo še spletni imenik **Open directory** (dmoz.org), ki je v lasti korporacije Netscape in velja za največji tovrstni sistem na spletu. Za nas je relevanten, ker ga podobno kot folksonomije, ki jih bomo opisali v nadaljevanju, lahko ureja vsak, ki to želi. Mesečno več tisoč aktivnih urednikov v sistem tako vpisuje nove spletna mesta, kjer jih je doslej v več kot 590.000 kategorijah zbranih že skoraj 5 milijonov (dmoz.org).

Slika 2.5: Spletna stran DMOZ.org



16 glavnih kategorij se veji v nadaljnje podkategorije.

⁴ Da bi uporabniki vendarle našli, kar želijo in to tudi tam, kjer Yahoo ni predvidel, da iščejo, se podkategorije pojavljajo v več kategorijah. S tem namenom je bila uvedena oznaka @, ki označuje podkategorije, ki izvirno ne spadajo v kategorijo, kjer se v trenutnem pogledu nahaja. Podkategorija *Books & Literature* je denimo v nadrejeni kategoriji *Entertainment* označena na naslednji način: *Books & Literature@*, ker je izvirno sicer umeščena v kategorijo *Literature*. Yahoo je nato z namenom, da ne bi nastala zmešnjava v obliki polihierarhij, uvedel še eno pravilo, ki določa, da se lahko vsako spletno mesto v imeniku pojavi le v treh različnih (pod)kategorijah.

Vir: Dmoz (2009).

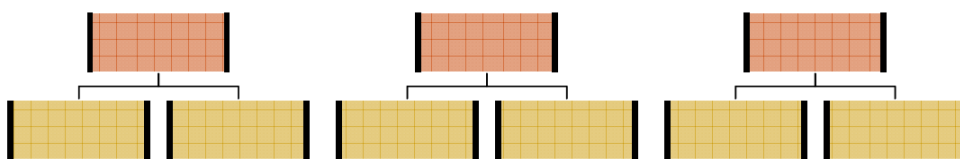
Tudi Open directory ima kot hierarhični sistem poleg že omenjene Yahoojeve težave še eno, ki se nanaša prav na njegovo odprtost. Zaznamujejo ga namreč številna odstopanja in izjeme od sicer jasno definirane uredniške politike, nemalo pa je tudi nesoglasij in posledično izključitev med uredniki. V kolikor strokovnjaki za klasificiranje niso strokovno usposobljeni in ne sledijo skupni taksonomiji, kar je v digitalnem okolju spleta težko zagotoviti, tovrstne težave z discipliniranjem ekspertov-amaterjev niso ravno presenetljive.

Z vidika taksonomičnega klasificiranja vsebin Shirky (2005) izpostavlja predvsem dve problematični lastnosti spletnih imenikov, in sicer **branje misli** (*mind reading*) in **vedeževanje** (*fortune telling*). Z razvstitvijo vsebine v določeno kategorijo ni nujno, da je izražen pomen vsebine, kot jo razume posamezni uporabnik, taksonomični sistemi pa temeljijo prav na predpostavki o tem, da znajo brati misli uporabnikom. Druga pomanjkljivost se nanaša na kriterij stabilnosti enot; z napredkom znanosti in zlasti z družbeno-političnimi spremembami se spreminja tudi kontekst informacij, zaradi česar so nerevidirane taksonomije večjih informacijskih sistemov znanstvenih inštitucij pogosto nenatančne in zastarele. Informacije tako pogosto obstanejo v času in v razmerah, v katerih so nastale, in niso pravi odsev sedanjosti.

2.4.2 Fasetna klasifikacija

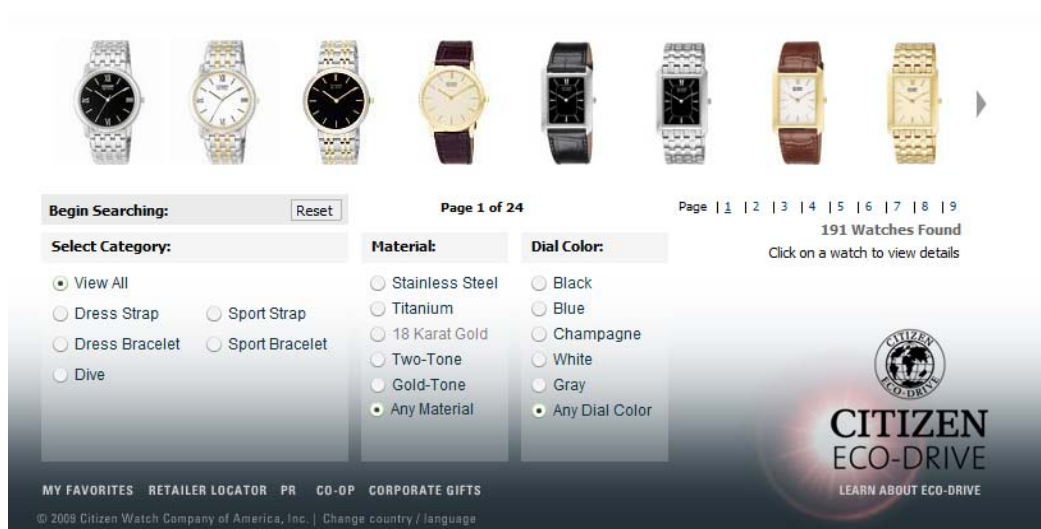
Zanimiva različica hierarhične organizacije informacij je tudi **fasetna klasifikacija**, katero je v 30. letih prejšnjega stoletja izumil indijski bibliotekar Ranganathan. Fasetna klasifikacije temelji na polihierarhični ureditvi. V njej so pojmi razvrščeni po nekaj osnovnih kategorijah oziroma fasetah, ki tako skupaj tvorijo različne hierarhije. Ročna ura ima denimo lahko opredeljene naslednje fasete: *cena* (do 100 €, nad 100 €...), *poreklo* (Švica, Japonska), *barva*, *spol* (moška ali ženska) in *tip* (digitalna ali analogna). Za razliko od klasične taksonomije je dostop do pojma mogoč po številnih poteh, saj ne zahteva umestitve predmeta v točno eno določeno kategorijo. Glavna uporabna vrednost fasetne klasifikacije je, da torej **obdrži drevesno strukturo, vendar ne določi vrstnega reda vej.**

Slika 2.6: Fasetna klasifikacija



V spletnem okolju se fasetna klasifikacija pojavlja predvsem v spletnih trgovinah pa tudi kot način prikaza digitaliziranih zbirk izobraževalnih ustanov, pri čemer so fasete oziroma klasifikatorji ekspertno določeni vnaprej in se lahko z minimalnimi stroški sproti dodajajo.

Slika 2.7: Primer spletnega fasetnega klasifikacijskega sistema



Glavne kategorije/fasete so pas, material in barva.

Vir: Citizen Watch (2009).

Fasetna razvrstitev odlično deluje, kadar je v njej veliko število predmetov, ko so meje med lastnostmi jasno opredeljive (na primer po spolu, barvi ali poreklu) in kadar je težko napovedati pot uporabnikovega brskanja po seznamu (Weinberger 2007, 78-84). Zlasti slednja ugotovitev je pomembna tudi z vidika naše naloge. Fasete kljub hierarhični ureditvi v nasprotju s taksonomijami izkoriščajo prednosti digitalnega okolja, saj uporabniku s polihierarhično distribucijo enot omogočajo večjo preglednost, skrbnikom sistema pa hitro dopolnjevanje ali spreminjanje kategorij.

2.4.3 Spletni iskalniki

Spletni velikan Google ni bil prvi, ki je odpravil hierarhijo in v obliki spletnega iskalnika ponudil

nehierarhično posredovanje informacij, se je pa v trenutku, ko se je pojavil na spletu, osredotočil na najpomembnejše, na sledenje uporabnikovemu brskanju po hiperpovezavah. Pomemben del Googlovega uspeha pri ponujanju najbolj relevantnih zadetkov iskanja je bil **algoritem PageRank** (PR), ki vsaki indeksirani strani dodeli številko, ki odraža njeno pomembnost. Pri tem imajo povezave s spletnih strani, ki imajo višji PR, večjo težo od tistih z nižjim PR-om in so zato tudi višje uvrščene v seznamu iskalnih zadetkov. Google je s PageRankom premeteno izkoristil **modrost množice** oziroma uporabnikov spletne skupnosti (Surowiecki 2005), katere koncept bomo razložili v nadaljevanju.

Spletni iskalniki imajo kljub svoji hitrosti in obsežnosti vrnjenih zadetkov tudi nekaj večjih pomanjkljivosti. Predvsem velja, da vrnjenih zadetkov ne znajo postaviti v kontekst, v katerem jih uporabnik išče. Ajax je na primer računalniški jezik, pa tudi nizozemski nogometni klub in kraj v Kanadi, za katerim mora uporabnik v seznamu zadetkov že dobro pobrskati, če ga želi najti. Iskalniki so zelo učinkoviti pri tarčnem iskanju, ko uporabnik ve, kaj želi najti in zna to tudi ubesediti v iskalni pojem, manj pa, ko ne ve točno, kaj išče. Iskalnike lahko razumemo tudi kot trenutne ponudnike informacij. Ko jih zapremo, moramo iskanje ponoviti in se pri tem spomniti istega iskalnega pojma, če želimo ponovno najti isti zadetek.

2.4.4 Osebni zaznamki

Poleg spletnih direktorijev oziroma imenikov poznamo še en (avto)taksonomičen način shranjevanja spletnih virov. To so zaznamki (angl. *bookmarks* v brskalniku Mozilla Firefox) ali priljubljene povezave (angl. *favorites* v brskalniku Internet Explorer)⁵, ki so torej v brskalnik shranjeni URL⁶ naslovi spletne strani, do katerih je mogoče s klikom takoj dostopiti. Zaznamki so zelo priročen način shranjevanja priljubljenih naslovov spletnih strani z namenom, da se jih ni treba zapomniti ali za dostop do njih vedno znova uporabljati iskalnikov. Shranjuje jih torej sam uporabnik, lahko tudi v kategorije oziroma mape, ki jih prav tako sam določi.

Kljub priročnosti in hitri dostopnosti imajo zaznamki nekaj pomanjkljivosti; shranjeni so lokalno, v brskalniku, zato se do njih ne da dostopati iz drugega računalnika ali jih deliti z drugimi uporabniki. V brskalniku so urejeni v hierarhični seznam na podoben način, kot je urejena mapa Moji dokumenti v programskem okolju Windows. Lahko se jih združuje v posamezna gnezda

⁵ Prvi brskalnik, ki je omogočal uporabo zaznamkov, je bil brskalnik Mosaic iz leta 1993, kjer se je seznam zaznamkov imenoval *hotlists*.

⁶ URL (uniform resource locator) je naslov, s katerim je naslovljena spletna stran oziroma predmet v svetovnem spletu.

oziroma mape in tudi poljubno poimenuje, za uporabo posameznega zaznamka pa ga je treba najprej poiskati s seznama vseh zaznamkov. Če ima uporabnik veliko število shranjenih zaznamkov in map, mora za preglednost in uporabnost teh porabiti torej precej časa.

2.4.5 Družbeni zaznamki - folksonomije

Programer Joshua Schachter, ki je imel težave ravno z obsežnostjo in s tem nepreglednostjo svojega seznama osebnih zaznamkov, se je organizacije spletnih informacij lotil tako, da je razvil lasten sistem shranjevanja spletnih virov. Zanimive spletne strani, na katere je naletel pri brskanju, je začel označevati s svojimi oznakami, tako imenovanimi "tagi", preprostim metapodatkovnim opisom vsebine shranjenega spletnega vira, ki jih je shranjeval na svojo spletno stran. Konec leta 2003 je sistem odprl tudi za druge uporabnike in nastal je spletni servis **del.icio.us**⁷; kolektivni, kolaborativni spletni projekt označevanja spletnih strani in virov (Green 2005).

Zaznamki, prej shranjeni lokalno v brskalniku, so z objavo na spletu tako postali **družbeni zaznamki**, dostopni vsem. Ker se shranjujejo z uporabo oznak, ki jih uporabnik določi sam, po želji in kot se mu zdi najbolj smiselno, za razliko od osebnih zaznamkov torej **niso urejeni hierarhično**, pač pa po **prosto izbranih ključnih besedah**. S tem je omogočeno lažje iskanje vsem, ki uporabljajo družbene zaznamke, ne le tistemu, ki je določen zaznamek prvi shranil. Pogoji za to je sicer smiselna uporaba ključnih besed za označevanje spletnih virov in do določene mere skupno kulturno ozadje in referenčni sistem uporabnikov.

Delicious je postal izjemno priljubljen in kmalu so mu sledili tudi drugi podobni projekti, ki so temeljili na družbenem označevanju spletnih virov, kar je kot svojevrsten družbeni fenomen Wal (2007) kasneje poimenoval z imenom **folksonomija**. Preden razložimo princip delovanja tega sistema, se ustavimo pri slovenski terminologiji s tega področja. V slovenščini doslej še ni bilo veliko poljudnih in prav tako tudi ne akademskih objav na to tematiko, zato sem se za ustrezno poimenovanje oprl predvsem na letak z opisom zanimivega evropskega projekta MELT, ki so ga leta 2007 izdali pri Fakulteti za računalništvo in informatiko⁸. Za glavne termine bomo v nadaljevanju tako uporabili naslednja slovenska poimenovanja:

⁷ Spletni servis prej znan pod domeno Del.icio.us se je julija 2008 preimenoval v Delicious.com.

⁸ Namen projekta MELT (A Metadata Ecology for Learning and Teaching) je omogočiti uporabnikom učnih vsebin v šolah dostop do bolj uporabnih tipov metapodatkov, ki bi omogočili najti tiste vire, ki ustrezajo njihovim potrebam, jeziku, kulturi in njihovim priljubljenim načinom poučevanja in učenja. Učne vsebine bi bile opremljene z **obogatnimi metapodatki**, kar bi omogočilo učiteljem in učencem enostavno združeno iskanje po tej vsebini. Obogatitev metapodatkov bi temeljila prav na uporabi folksonomij in družbenega označevanja.

tag	oznaka
tagging	označevanje
social tagging	družbeno označevanje
bookmark	zaznamek
bookmarking	postavljanje zaznamkov
social boobookmarking	družbeno postavljanje zaznamkov
folksonomy	folksonomija

Avtor izraza *folksonomy* Wal (2007) folksonomijo razume kot **družbeno klasificiranje od spodaj navzgor** in pravi, da je folksonomija rezultat osebnega prostega označevanja informacij in predmetov (česarkoli z URL-jem) za kasnejši osebni priklic, pri čemer so oznake dostopne tudi drugim. Beseda folksonomija je sicer skovanka angleške besede *folk* (ljudje) in grške *nomia* (upravljanje) ter nakazuje pomemben premik v načinu uporabe spletnih informacij. Z možnostjo bogatenja informacij z lastnimi metapodatki (v tem primeru oznakami) ljudje niso več pasivni uporabniki spletnih virov, pač pa aktivni ustvarjalci njihovega pomena. Na preplet družbenega in osebnega se v svoji definiciji opira tudi Mathes (2005), ki folksonomijo opredeli kot sistem, v katerem uporabniki dokumentov in medijev za lastno uporabo ustvarjajo metapodatke, ki so na voljo tudi drugim uporabnikom v skupnosti.

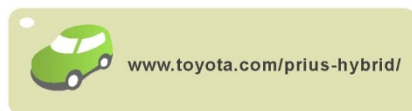
Oznake ustvarjajo med seboj nepovezani uporabniki, običajno brez posegov administratorja ali programskega algoritma v njihov opis. Proces dodeljevanja ključnih besed oz. oznak spletnim vsebinam imenujemo **družbeno označevanje** (angl. social tagging).⁹ Po definiciji *tag* pomeni oznako v obliki koščka papirja, blaga ali kovine, ki se drži predmeta in vsebuje informacije o njem (Cambridge Dictionary 2008). Iz razlage pojma torej ni določeno, kakšne so informacije, ki jih oznaka vsebuje, koliko jih je in kdo vse je sodeloval pri njenem nastanku. Prav tako oznaka ne določa razmerja do predmeta; avtor oznake ni nujno lastnik ali avtor označenega predmeta. Te lastnosti se zelo dobro ujemajo tudi z oznako, ustvarjeno v spletnem okolju folksonomij. Guyjeva in Tonkinova (2006) ji tako pravita **prost izbor ključnih besed**.

⁹ Popularnost družbenega označevanja strmo narašča. Po raziskavi ameriškega inštituta Pew je 28 % obiskovalcev spleta že uporabilo oznako, 7 % pa jo je nazadnje uporabilo prejšnji dan (Rainie: 2007).

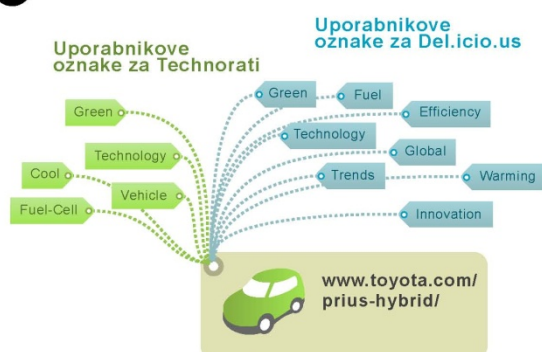
Slika 2.8: Prikaz označevanja spletnega vira

Družbeno označevanje

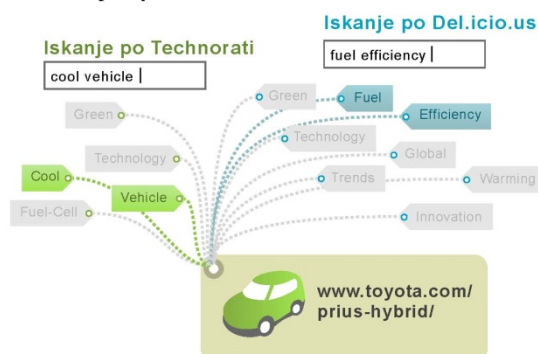
- 1 Spletna vsebina (npr. spletna stran, video, fotografija, blog)



- 2 Uporabniki naletijo na spletno vsebino in jo po svoje označijo



- 3 Drugi uporabniki najdejo označeno vsebino z iskanjem po oznakah



Prerejeno po Elliance (2006).

Shirky (v Weinroth 2005) s pojavom folksonomij ugotavlja, da se z njimi umikamo od binarne, ekskluzivne kategorizacije, v svet verjetnosti, kjer je razumevanje predmeta prepuščeno ljudem, kar poimenuje **organska organizacija informacij**. Glede na prejšnje sisteme ekspertnega, uredniškega razvrščanja informacij, gre za radikalni prelom, saj je klasifikacija v folksonomijah torej prepuščena posameznemu uporabniku in njegovemu lastnemu razumevanju označene vsebine. Locke (2005) ugotavlja, da v klasifikacijskih sistemih med seboj za prevlado tekmujejo tri vrednote, in sicer **primerljivost**, **vidnost** in **nadzor**, pri čemer sta v folksonomijah na račun vidnosti (sprotnega takojšnjega ustvarjanja in spreminjanja vsebine) zapostavljeni primerljivost

(možnost prenosa taksonomij iz enega konteksta v drugega) in nadzor (omejitev uporabnikovih interpretacij). Folksonomije ne zahtevajo skoraj nikakršnega tehničnega znanja za uporabo aplikacije in označevanja ne omejujejo z besediščem ali jezikom, s čimer bolj odražajo uporabnikov odnos do označene vsebine, še posebej kadar je ta tudi avtor označene spletne vsebine. Bolj problematična je seveda uporabna vrednost za druge uporabnike, česar se bomo dotaknili v poglavju tehničnega pregleda folksonomij.

2.4.6 Semantični splet

Semantični splet, imenovan tudi podatkovni splet¹⁰, je naslednji večji tehnološki korak v razvoju spleta, četudi je obris njegove vizije celo starejši od spleta, ki ga trenutno uporabljamo. Še preden je nastal svetovni splet, je Berners-Lee namreč napisal program Enquire, ki je omogočal povezovanje podatkov in tudi omejeno opisovanje njihovih medsebojnih relacij (Weinberger 2007). Za razliko od navadne spletne povezave, ki le povezuje točko A s točko B, je bila povezava v Enquireju obogatena z metapodatki in je na primer lahko opredelila, da A vsebuje B ali da je A ustvaril B. V razvoju Berners-Leejevega svetovnega spleta se je ideja tovrstnih pametnih povezav kasneje umaknila praktičnosti ustvarjanja preprostih A do B hiperpovezav. S tem se je sprožil nesluten razmah uporabe spleta, ki bi bil lahko s časovno zamudnim doslednim metapodatkiranjem sicer ogrožen.

Danes, tri desetletja kasneje, prihaja počasi v ospredje Enquirejev naslednik, a v rahlo drugačni obliki. Semantični splet predstavlja idejo, kjer je ogromna količina podatkov na svetovnem spletu zapisana v računalniku razumljivi obliki (Lavbič in drugi 2006). Ne gre torej za nov splet, ampak le za podaljšek obstoječega oziroma za dodajanje nove plasti, katerega bistvo je v strukturiranju informacij na način, da bodo inteligentni agenti (stroji) lahko razumeli njihov pomen in tako med drugim denimo nadomestili klasično, nenatančno iskanje po ključnih besedah.

Trenutni splet zato lahko razumemo kot decentralizirano platformo distribuiranih prezentacij, medtem ko bo semantični splet decentralizirana platforma distribuiranega znanja (Tauberer, 2008). Resnična uporabnost svetovnega semantičnega pa bo dosežena, ko bo uporabnik prek

¹⁰ Semantični splet je kot nekakšen naslednik družbenega spleta 2.0 hitro dobil tudi oznako splet 3.0. Pri tem ne gre nujno za sinonimno pojmovanje, saj bo semantičnost prav mogoče osrednja lastnost spleta 3.0, ne pa nujno tudi edina. O'Reillyjev (2007) splet 3.0, kateremu poimenovanju je sam sicer izrazito nenaklonjen, se tako bolj kot na dedukcijo spleta na semantičnost, nanaša na prelom z obstoječo paradigmo zamejitve spleta na zaslon in tipkovnico. Prihodnost vidi v instrumentalizaciji naših aktivnosti in uporabi storitev, ki so vezane na resnično življenje, na primer v uporabi pametnih mobilnih telefonov ali kreditnih kartic v kombinaciji z naprednimi aplikacijami.

računalnikovega povezovanja in avtomatskega sklepanja iz trenutno razdrobljenih podatkov v celoti izkoristil vsebino spleta. V praksi bi uporaba semantičnega spleta potekala nekako takole: proizvajalec na spletni strani predstavi in opiše nov izdelek, spletne trgovine objavijo informacijo o ceni in zalogi, tretja spletna stran pa o izdelku objavlja mnenja in priporočila obstoječih kupcev. Vse te distribuirane informacije inteligentni agent analizira in poda najboljši odgovor na zahtevano iskanje, na primer: »Poišči izdelek kategorije X v cenovnem razponu od A-B, ki je odlično ocenjen in je na zalogi«.

Semantični splet obeta radikalno spremembo v načinu priklica informacij, saj bo z računalniškim razumevanjem vsebine omogočal iskanje tudi na semantični ravni in ne zgolj sintaktični.

3 Hipertekstovnost spleta

3.1 Kultura hiperteksta

V naslednjih podpoglavjih bomo izhajali iz teze, da je glavna spremljevalka spleta 2.0, ki temelji na uporabniškem soustvarjanju spletnih vsebin, na novo realizirana hipertekstovnost, ki kot temeljna ideja svetovnega spleta (Oblak in Petrič 2005) omogoča obstoj povezav med besedilnimi deli. Že leta 1945 je idejo hiperteksta orisal računalniški pionir Vannevar Bush z načrtom elektronske naprave **memex**, ki bi omogočala enostavno in fleksibilno shranjevanje informacij in dokumentov na asociativni ravni z možnostjo kasnejšega dodajanja novih povezav, zapiskov in komentarjev ljudem različnih profilov in poklicev (Petrič 2003, 29). Čeprav je ostal le pri ideji, je Bush ob iskanju nadgradnje obstoječih nenaravnih in neuporabnih sistemov indeksiranja informacij odlično zadel bistvo hiperteksta, kot ga poznamo danes:

Človeška misel deluje na osnovi asociacij. V nekem trenutku ima v prijemu določen element, nakar v trenutku skoči na drugega, ki je sugeriran z asociacijo misli, v skladu z neko kompleksno mrežo poti, ki jih nudijo možganske celice [...] Človek ne more upati, da bi lahko ta proces v popolnosti repliciral na umetnem nivoju, vsekakor pa se bo iz njega učil. Ključna ideja, ki jo je vredno povzeti iz te analogije, je selekcija. Selekcija na osnovi asociacije, in ne na osnovi indeksiranja. (Bush v Petrič 2003, 29).

Prvi, ki je novo obliko elektronskega teksta in načina objavljanja poimenoval s predpono hiper, je bil leta 1962 Theodor H. Nelson, ki je hipertekst razumel kot “nezaporedno pisanje – tekst, ki se veji in omogoča bralcu izbire, ter je najboljše bran na interaktivnem zaslonu ... [t]o so nizi delov teksta, povezani s povezavami, ki nudijo bralcu različne poti”. (Nelson v Landow 2006).

K razumevanju koncepta hipertekstovnosti so nato pomembno prispevali poststrukturalisti, ki so se odrekli avtorju in na njegovo mesto postavili bralca oziroma sam akt branja. Julia Kristeva (Walker 2005) je tako zapostavila vlogo avtorja z idejo intertekstualnosti, po kateri noben tekst ne more obstajati sam po sebi, temveč je vedno del omrežja eksplicitnih in implicitnih namigov in citatov drugih besedil. Z vidika proučevanja hipertekstovnosti elektronskega teksta pa je s pojmom odprtega, idealnega teksta, najbolj aktualen Roland Barthes, ki je v osnovi „... galaksija znakov, ne struktura pomenov; nima začetkov; je reverzibilen; do njega dostopamo s številnih vhodnih točk, nobena od teh ni avtoritativno proglašena za glavno.“ (Barthes v Petrič

2003, 36).

Barthesov koncept odprtega teksta zato, kot ugotavlja Bolter (v Landow 2006), posebej hipertekstovnost, kot jo predvsem po zaslugi svetovnega spleta poznamo danes. In četudi je slednja njegov sestavni del, potencial hiperteksta do razvoja novih spletnih tehnologij v veliki meri ni bil neuresničen oziroma je bil podrejen komercialni funkciji, ki, kot ugotavlja Petrič (2005), „po eni strani omejuje potencialno svobodo delovanja bralca svetovnega spleta kot informacijskega sistema, po drugi pa spodbuja svobodo avtorja spletnega mesta.“

Ko govorimo o poststrukturalistični odpravi avtorja, lahko ugotovimo, da se (spletnega avtorja) na nek način odpravlja tudi v spletu 2.0, kjer se odbrane informacije postavlja v kontekst, ki ustreza bralcu, potrošniku informacij. A še bolj kot o odpravi avtorja lahko govorimo o **sodelujočem avtorstvu**, ki pomeni premik od dojemanja spletnih uporabnikov kot potrošnikov k aktivnim ustvarjalcem spleta. Šele s tem strukturnim zasukom se pravzaprav uresničuje resnični demokratični potencial spleta, ki je bil doslej reduciran bolj ali manj le na splošno dostopnost do informacij, ne pa tudi na možnost njihovega soustvarjanja.

3.2 Splet 2.0

Z današnje časovne perspektive je prelomna točka v razvoju spleta verjetno pok dot-com mehurčka spomladi leta 2000. Tvegano, prekomerno investiranje v spletna podjetja, ki so temeljila na trhlih poslovnih modelih, se je končalo s sesutjem tehnološke borze in zdesetkanjem spletne industrije. Tovrstni pretresi so tipični spremljevalci tehnoloških revolucij, ki navadno pomagajo ločiti zrno od plevela, in omogočijo preboj najboljših idej.

In res, na izpraznjenem prizorišču so novo priložnost izkoristila preživela podjetja in novi igralci, ki so medij in njegove uporabnike očitno razumeli bolje. Pojavili so se številni spletni servisi, ki so omogočili uporabnikom soustvarjati spletne vsebine, kar je temeljito spremenilo samo pokrajino spleta. Vprašanje, kaj se dogaja, je razmeroma dolgo čakalo na svoj v teoriji podprt odgovor. Plod viharjenja idej z novomedijske konference leta 2004 je bilo O'Reillyjevo verziranje svetovnega spleta z oznako 2.0 (*web 2.0*), ki je tako končno postal oprijemljiv koncept.

Preden se osredotočimo na vsebinske vidike spremenjenega spleta, se ustavimo najprej pri številki 2.0 in opredelimo njen pomen. V svetu računalništva pomeni oznaka naslednje stopnje

določeno programsko nadgradnjo ali izboljšavo. Pirc (2007) pravi, da je 2.0 sama ideja, da je moč komuniciranja pri uporabnikih in jo lahko apliciramo na vsa področja družbenega življenja, med drugim tudi na posel, oglaševanje ali politiko in torej ne zgolj na spremenjeno pokrajino digitalnega okolja. Splet 2.0 ali participatorni splet temelji na **mobilizaciji spletne skupnosti** oziroma na **uporabniškem soustvarjanju spletnih vsebin**. Konkretno se v spletu 2.0 aktivnost uporabnikov izraža v ustvarjanju in soustvarjanju raznovrstnih spletnih vsebin in družbenega programja, kot so med drugim blogi, wiki¹¹ in folksonomije oziroma družbeni sistemi za označevanje spletnih vsebin.

Za predhodnika spleta 2.0 velja **komercialni splet ali splet 1.0**, ki ga lahko najpreprosteje opišemo kot zbirko spletnih strani. V njem je vsebino ustvarjala le peščica spletnih avtorjev, uporabniki spleta pa so bili ustrezno terminološko degradirani v obiskovalce, katerih podrejenost se je kazala zlasti v načinu centraliziranega in hierarhičnega konzumiranja spletnih informacij. Obiskovalca se je prizadevalo zadržati čim dlje na spletnem mestu in napraviti vtis nanj, bolj kot z vsebino in ponudbo dodatnih, spletu prilagojenih storitev, pa so se podjetja ukvarjala z obliko svoje internetizirane vizitke oz. brošure. To je bil korporativiziran splet, reduciran zgolj na vlogo dodatnega orodja v promocijskem spletu, s čimer se je močno odmaknil od demokratičnega koncepta spleta, ki ga je njegov izumitelj Berners Lee (Berners Lee in drugi 2001) v izhodišču zastavil kot prostor medsebojnega sodelovanja enakovrednih uporabnikov.

Mnogi znani avtorji in pionirji spleta 2.0 so do samega termina in njegove verzifikacije postopoma zavzeli kritično distanco in ga ožigosali kot „marketinški slogan“ (Shaw 2005), „še en mehurček“ (Winer 2005) in „zamegljevanje“ (MacManus 2005), ki izrablja napredne tehnologije za ponovno ustvarjanje borzne histerije. K takšnemu razmišljanju je pripomoglo tudi dejstvo, da splet 2.0 kljub jasni paradigmi koncepta mobilizacije in demonopolizacije komuniciranja, ki se je kmalu začel pojavljati in uveljavljati tudi na drugih družbenih področjih (opazen je zlasti v novejših marketinških prijemih, ki v proces ustvarjanja izdelkov, storitev ali oglasov vabijo neposredno uporabnika¹²), ni doživel splošno sprejete definicije. S kančkom cinizma bi lahko celo rekli, da so se teoretiki našli sindroma spleta 2.0 in po participatornem principu kar vsi po vrsti ustvarjali nove definicije.

To velja tudi za avtorja termina "web 2.0" Tima O'Reillya, ki je svojo originalno definicijo vmes

¹¹ Wiki (havajsko *hiter*) je oblika hipertekstovnega spletnega mesta, ki ga z dodajanjem, spreminjanjem in brisanjem vsebin lahko ureja kdorkoli. Najbolj znan wiki je spletna enciklopedija Wikipedia.

¹² V Sloveniji smo v zadnjem obdobju spremljali najmanj dve medijsko intenzivno podprti interaktivni kampanji, za Itak (Mobitel) in Cockto, pri katerih so uporabniki v spletnem okolju ustvarjali stripe.

večkrat spreminjal, a se naposled vrnil k prvotni in splet 2.0 opredelil kot „poslovno revolucijo v računalniški industriji, ki je posledica premika k uporabi interneta kot platforme in poskusa razumevanja pravil za uspeh na tej platformi. Glavno izmed teh pravil se glasi: zgradi aplikacije, ki izkoriščajo omrežne učinke na način, da postajajo boljše, več ljudi kot jih uporablja. (To je, kar sem drugje poimenoval koriščenje kolektivne inteligence).“ (O’Reilly 2006).

Kljub navidezni neizčiščenosti lahko ocenimo, da je ta definicija z vidika razumevanja spleta 2.0 dobra in nam bo v veliko pomoč tudi pri razumevanju družbeno označevalnih sistemov, opisanih v naslednjih poglavjih. Za boljšo predstavbo, kaj splet 2.0 natančneje predstavlja, pa na tem mestu najprej povzemimo nekatere glavne razsežnosti, kot jih v spletu 2.0 prepozna O’Reilly (2005):

- **Splet kot platforma:** selitev aplikacij, na primer odjemalca elektronske pošte (Gmail.com), enciklopedije (Wikipedia.org) ali fotoalbuma (Flickr.com) z namizja na splet in s spleta na druge naprave. Za razliko od licenčnih računalniških programov dohodek podjetij ne izhaja iz licenčnine za uporabo storitve, ta je v glavnem brezplačna, temveč predvsem iz prodaje oglaševalskega prostora in donacij uporabnikov. Kelly (2005) do leta 2015 napoveduje zaton razvoja namiznih operacijskih sistemov. Pomemben bo le še splet in dostop do njega, ne glede na napravo, s katero bomo to storili.
- **Uporabniško posredovana vsebina:** splet 2.0 temelji na zbiranju podatkov, ki jih v sistem prispevajo uporabniki sami. Informacija je ključna, čemur je prilagojena tudi arhitektura in oblika spletne strani, ki je v funkciji vsebine. Programersko znanje ni več pogoj za ustvarjanje spletnih vsebin, saj so aplikacije prilagojene povprečnemu uporabniku. Demokratizacija trga informacij, v katerem je vsak lahko producent in hkrati urednik vsebin, pomeni resno grožnjo tradicionalnim množičnim medijem, ki so v spletnih uporabnikih – avtorjih dobili resnega konkurenta. Po začetnem medlem odzivu tradicionalni mediji sedaj skušajo najti svoje mesto tudi na spletu.
- **Beta verzije:** klasični softverski cikel izdajanja novih različic programov vsakih nekaj let je zastarel koncept. Uspešne spletne aplikacije 2.0 se sproti razvijajo in izboljšujejo ter tako nenehno ostajajo v beta (testni) verziji. K razvijanju so povabljeni uporabniki sami, bodisi v odprtokodnih ali komercialnih projektih. Več kot je sodelujočih, večji je pretok znanja in boljše so aplikacije.

- **Koriščenje kolektivne inteligence:** osrednja lastnost spletnih velikanov, kot je na primer Amazon.com, ki je preživel kolaps spleta 1.0 in je eno vodilnih podjetij tudi v spletu 2.0, je verjetno prav koriščenje kolektivne inteligence. Amazon.com je iz uporabniške izkušnje naredil celo znanost. Uporabnik lahko skoraj na vsaki strani ocenjuje, sodeluje in tako s koriščenjem skupnih informacij medsebojno izboljšuje poslovne odločitve. Na tovrstnem principu delovanja temelji večina spletnih servisov spleta 2.0, kar je v kontekstu njegove komercializacije lahko tudi nov povod za kritiko v obliki reartikulacije pojma *koriščenje* v *izkoriščanje* kolektivne inteligence.

- **Bogata uporabniška izkušnja:** med vzroki, zakaj svetovni splet že v 90. letih prejšnjega stoletja ni postal orodje sodelujočega avtorstva, Berners Lee (developerWorks Interviews 2006) izpostavlja za povprečnega uporabnika prezapleten spletni jezik HTML, ki je večini onemogočal dostop do urejanja svetovnega spleta in s tem zaviral osnovno človeško željo po kreativnosti in izražanju. Pisni in ne zgolj bralni dostop do svetovnega spleta so omogočili šele WYSIWYG (WhatYouSeeIsWhatYouGet – kar vidiš, to dobiš) vmesniki, na katerih temeljijo med drugim tudi blogi, ki omogočajo uporabnikom enostavno zapisovanje vsebin. Tehnologija oz. skupek tehnologij, ki je omogočila bogato uporabniško izkušnjo in tako uporabniku prijaznejši splet, je bil zlasti Ajax (*Asynchronous JavaScript and XML*). Javascript, programski jezik, ki poganja ajax, določa, kako bodo vsebine prikazane ter pri tem komunicira s strežnikom, XML pa opisuje vsebino. Ajax omogoča izmenjavo le določenih informacij med brskalnikom in strežnikom, zaradi česar je prenos podatkov optimalen. Namesto vsakokratnega osveževanja spletne strani, slednje poteka le v delu strani, kjer uporabnik spreminja vsebino, s čimer je odpravljeno čakanje na strežnik, da naloži celotno stran.

Kljub na prvi pogled morda manj razumljivemu konceptu, je v ozadju novega spleta torej zlasti poenostavitev njegove uporabe. Z odpravo nekaterih ovir, ki so v preteklosti preprečevale večjo vključenost uporabnikov v soustvarjanje spletnih vsebin, virtualno okolje spleta danes zaradi svoje povezovalne narave vse bolj spominja na naš realni svet. Ključno pri tem je spoznanje, da splet 2.0 poganja iz vajeti spuščena hipertekstovnost, ki pomaga uresničevati nekoč nedosegljive demokratične ideale.

3.3 Divja hipertekstovnost

Hipertekstovnost je bila del številnih podatkovnih sistemov že dolgo pred vzpostavitvijo svetovnega spleta. A pri tem ne moremo govoriti o isti entiteti, Walkerjeva (2005) namreč ugotavlja, da je hipertekstovnost s spremembo svojega habitata spremenila tudi svojo naravo; nekoč v zaprtih informacijskih sistemih nadzorovana in disciplinirana, je v svetovnem spletu, zlasti pa v novem spletu 2.0, postala **nepredvidljiva** in **divja**. Za ovrednotenje te spremenjene narave hipertekstovnosti je koristno najprej opredeliti njene izhodiščne, teoretično idealno zastavljene razsežnosti, ki v komercialnem spletu zaradi različnih okoliščin doslej niso bile povsem uresničene. Po Petriču in Oblakovi (2005) veljajo za hipertekst v odnosu do tradicionalnih množičnih medijev naslednje lastnosti:

- **decentraliziranost** (odsotnost središča oz. osrednjega besedilnega dela)
- **nelinearnost** (besedilni del nima ne začetka ne konca, kar omogoča možnost izbire na vsakem koraku)
- **aktivna vloga bralca** (bralec sam izbira pot skozi besedilo ali pa z dodajanjem svojih besedilnih elementov tudi sam postane tvorec hiperteksta)
- **večglasna dinamičnost** (opredeljuje sodelujoče avtorstvo in ne dovoljuje enega glasu, glasu tirana, ki bi preglasil ostale. Hipertekst je trenuten, saj se z dodajanjem, brisanjem in povezovanjem besedilnih delov nenehno spreminja.).

Če na tem mestu zanemarimo družbeno-ekonomske neenakosti pri dostopu do spleta in se osredotočimo zgolj na uporabnike z omogočenim dostopom, so njihovo večjo participacijo pri ustvarjanju spleta omejevale predvsem tehnične zahteve. Med njimi že omenjeno nepoznavanje jezika HTML, pa tudi še danes prevladujoča vloga spletnih iskalnikov in večjih spletnih portalov. Iskalni algoritmi in uredništva portalov namesto uporabnika namreč sami opravijo prvo selekcijo informacij. Tovrstna **disciplinirana hipertekstovnost**, kamor lahko uvrstimo tudi še vedno vseprisotne vsiljive spletne oglase, ki kot ugotavlja Petrič (2003, 52), ne temeljijo na principu sodelovanja med teksti, ampak je njihov namen izključno pridobivanje potrošnikov, je z vidika zgoraj naštetih idealnih lastnosti zelo problematična. Ko so hiperpovezave v primeru portalov **nadzorovane** (s strani spletnih urednikov), v primeru spletnih oglasov **vsiljene** (s strani oglaševalcev) in v primeru spletnih iskalnikov **algoritmično določene**¹³, o pravi

¹³ Google spletne strani razvršča s pomočjo algoritmične tehnologije PageRank in z analizo ujemanja hiperteksta. PageRank objektivno izmeri pomembnost spletnih strani, tako da reši enačbo z več kot 500 milijoni

decentraliziranosti in nelinearnosti teksta ni moč govoriti. V takem okolju ni prostora niti za sodelujoče avtorstvo niti za aktivnega bralca.

Do tu smo na kratko očrtali omejitve hipertekstovnosti spleta 1.0 in v kontekstu te naloge se sedaj lahko vprašamo, ali se je s spremembo tehnologije spleta, načina ustvarjanja in posredovanja informacij, spremenila tudi tehnologija hiperteksta; lahko s prihodom participatornega spleta govorimo tudi o **participatornem hipertekstu**? Walkerjeva (2005) s spletom 2.0 in z novo generacijo spletnih aplikacij opredeli namreč tudi rojstvo **divjega hiperteksta**, ki ni omejen s strani sistema ali lastništva oziroma avtorja. Kot najbolj očitne primere divjih hipertekstov razume velike označevalne projekte, ki ustvarjajo hipertekst in pomene znotraj njega brez očitnih avtorjev ali urednikov, ki bi nadzorovali povezave. Na videz torej nenadzorovani in kaotični sistemi družbenega programja, kot so folksonomije in wikiji, tako sicer dopuščajo uporabniško ustvarjene napake, nevednost in navidezno nepreglednost, vendar preprosta in fleksibilna struktura prinaša nekoč nedosegljive potenciale. S tem, ko je uporabnikom dopuščena popolna svoboda pri ustvarjanju povezav in dodajanju komentarjev, nastajajo sistemi, ki ne uresničujejo le na videz visoko zastavljenih idealov hiperteksta, temveč zelo spominjajo tudi na kolaborativno nadgradnjo Bushevega memexa, vizije intimne tehnologije in podaljška človekovega spomina, „ki bi uporabniku omogočal shranjevanje knjig, zapiskov in sporočil na način, da bi bili ti dostopni z visoko stopnjo hitrosti in fleksibilnosti“ (Bush v Petrič 2003).

Informacije se dopolnjujejo, popravljajo in na novo ustvarjajo s hitrostjo misli, neodvisno od prejšnjih centrov moči v obliki spletnih webmasterjev in informacijskih korporacij. Manifestacijo hiperteksta v okolju spleta 2.0 si zato velja ogledati поближе.

3.4 Hipertekstovnost spleta 2.0

Pri preučevanju realizacije hipertekstovnosti v spletu 2.0 sta prvo obsežnejšo raziskavo naredila Millard in Ross (2006), ki sta pod drobnogled vzela spletne sisteme **Flickr** (spletno mesto za shranjevanje osebnih fotografij), **WordPress** (zbirka blogov) in **MediaWiki** (Wikipedia), pa tudi družbeno-programске aplikacije kot so **Annotea** (označevalni projekt spletnih vsebin konzorcija W3C), **Twiki** (wiki za organizacije, ki podjetjem med drugim omogoča vzpostavitev internih blogov in intraneta) in raziskovalna nespletna hipertekstovna sistema **Xspect** ter **OHS/WWW**.

spremenljivkami in 2 milijardama izrazov. Nato izvede analizo ujemanja hiperbesedila in ugotovi, katere strani so ustrezne za določeno iskanje, ki se izvaja. (Vir: Google 2009).

Sisteme sta presojala skozi 5 kategorij in preučevala podporo glavnim hipertekstovnim kriterijem, ki so se z razvojem hipertekstovnih sistemov oblikovali skozi čas:

1) Iskanje

- iskanje po vsebini: klasično iskanje v iskalni vrstici po pojmu ali besedni zvezi
- iskanje po kontekstu: iskanje s pomočjo metapodatkov ali oznak, posredovanih od uporabnikov
- iskanje po strukturi: iskanje določenega strukturnega vzorca v sistemu

2) Struktura in vsebina

- podpiranje multipovezav: povezanost hiperpovezave z večimi viri
- združevanje enot v eno kategorijo, ki omogoča uporabnikom večjo preglednost in nadzor nad vnešenimi vsebinami
- dodatne navigacijske strukture, ki uporabnikom pomagajo pri naprednem iskanju
- sledenje uporabnikovi zgodovini: sledi, ki jih za seboj puščajo uporabniki spletnih aplikacij, lahko pomembno pomagajo tudi drugim uporabnikom, še posebej v sistemih elektronskega trgovanja. Spletna trgovina Amazon je vzpostavila zelo dovršen priporočilni sistem, ki uporabnikom na podlagi izkušenj drugih uporabnikov svetuje pri nakupnih odločitvah. V ozadju teh sistemov gre za tako imenovano **sodelovalno filtriranje**, ki posamezniku na podlagi mnenj in dejanj drugih uporabnikov v skupnosti omogoča bolj učinkovito prepoznavo za njih relevantnih vsebin iz sicer potencialno nepregledne množice različnih izbir (Herlocker in drugi 2004).

Slika 3.1: Seznam priporočil na Amazon.com na podlagi nakupov ostalih uporabnikov



Vir: Amazon (2008)

3) Dinamičnost / prilagodljivost

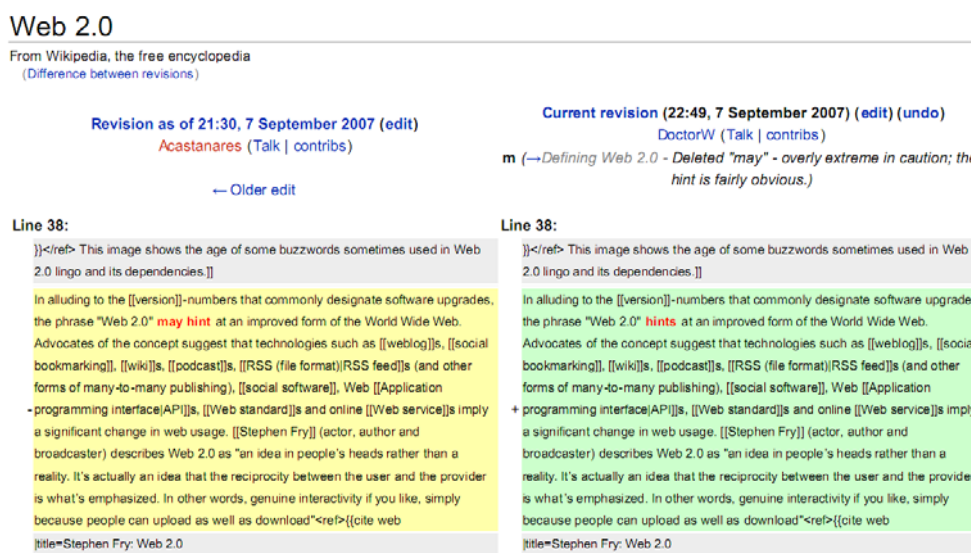
- dinamične vsebine, brez katerih si spleta 2.0 ni mogoče zamisliti. Med naprednimi programskimi jeziki, ki to dinamičnost omogočajo, je znan zlasti ajax.

- b) dinamične strukture, ki opredeljuje hipertekstovno strukturo sistema, na primer v obliki povezav ali zbirke povezav.
- c) izračun omrežnih statistik: pomeni možnost analize hipertekstovnega sistema, na primer izračun popularnosti določene oznake ali posameznega vira
- d) personalizacija: prilaganje aplikacije potrebam posameznega uporabnika, bodisi po vsebini (na primer združevanje oznak v zbirke) ali obliki (na primer prilagajanje videza bloga)

4) Shranjevanje različic

- a) shranjevanje različic enot: pomeni hrambo prejšnjih verzij zapisov, kar je za odstranjevanje napačnih ali zlonamernih vpisov zelo pomembno zlasti na Wikipedii.

Slika 3.2: Primerjava starejše in novejše verzije vpisa pojma Web 2.0 (Wikipedia):



Razlika med zapisoma je označena z rdečo barvo; may hint (starejši zapis) in hint (novejši).
Vir: Wikipedia. (2007).

- b) shranjevanje različic omrežja: s prihodom nove verzije aplikacije se ponavadi starejša verzija omrežja umakne iz spleta, hrambo zato omogočajo redki sistemi

5) Odprto urejanje

- a) zasebni komentarji: dovoljenje za objavljanje komentarjev ima le avtor teksta ali določena skupina uporabnikov
- b) javni komentarji: dovoljenje za objavo komentarjev ima katerikoli uporabnik sistema
- c) globalno sodelovanje: pomeni možnost neodvisnega ustvarjanja kateregakoli uporabnika

- d) omejeno sodelovanje: pooblastilo za sodelovanje uporabnik dodeli le prijateljem
- e) dodatna razširljivost sistema: možnost dodajanja novih funkcij sistemu in povezovanje sistema tudi z drugimi aplikacijami

Slika 3.3: Realizacija hipertekstovnih prizadevanj v spletnih sistemih

	Flickr	MediaWiki	Twiki	WordPress	Annotea	Xspect	OHS/WWW
Iskanje po vsebini	x	x	x	x	x	+	x
Iskanje po kontekstu	x	+	+	x	x	x	x
Iskanje po strukturi	+				x	x	+
Sledenje uporabnikovi zgodovini	x			x	x	x	
Združevanje enot v eno kategorijo	x	x	x	x	x	x	x
Dinamičnost vsebine	x	x	x	x	x	x	+
Dinamičnost strukture	+			x	x	x	+
Podpiranje multipovezav						x	x
Dodatne navigacijske strukture	x	x	x	+	+	x	x
Izračuni omrežnih statistik	x	x	x	x		x	x
Shranjevanje različic enot		x	x				x
Shranjevanje različic omrežij			x				
Zasebni komentarji	x	x	x	x	x	x	+
Javni komentarji	x	x	x	x	x	x	+
Sodelujoče avtorstvo	x	x	x	x	+	+	+
Omejeno avtorstvo	x	x	x	x	+	+	+
Personalizacija	x	+	x	x	x	x	x
Dodatna razširljivost sistema	x	x	x	x	+	x	x

X predstavlja popolno podporo, + pa delno podporo.

Vir: Millard in Ross (2006).

Kljub temu, da je bil v raziskavo vključen zelo omejen vzorec spletnih aplikacij spleta 2.0, lahko ugotovimo, da je večina prizadevanj hipertekstovne skupnosti v njem uresničenih, toda v celoti le kot zbirka različnih aplikacij, ki delujejo na skupni platformi spleta, ne pa v eni sami aplikaciji. Kot izpostavljata Millard in Ross (2006), so nekatera starejša, originalna hipertekstovna prizadevanja namenoma zapostavljena, predvsem v korist večje fleksibilnosti, prilagodljivosti in odzivnosti sistema oziroma zato, ker hipertekstovni kriterij preprosto ni relevanten ali primeren za določeno aplikacijo. Splet 2.0 tako sicer povsem ne izpolnjuje z visokimi kriteriji določene izhodiščne vizije hiperteksta, vendar je uporabniku prijazna alternativna različica, ki pragmatično in suvereno rešuje večino nerešenih problemov, s katerimi se je v praksi doslej soočal hipertekst.

4 Tehnična raven folksonomij in družbenega označevanja

4.1 Pojmi in ključne besede, oznake in metapodatki

Zaradi dokaj očitne terminološke zmede na področju družbenega označevanja opredelimo najprej razliko med pojmom in ključno besedo v ne nujno spletnem okolju. Reitz (2004), avtor bibliotekarskega in informacijskega terminološkega slovarja, **ključno besedo** [angl. keyword] opredeli kot pomembno besedo ali frazo v naslovu, predmetnem naslovu (deskriptorjih), v opisu vsebine, povzetku ali v zapisu besedila v spletnem katalogu ali bibliografski zbirki podatkov, ki se lahko uporabi kot iskalni pojem za priklic vseh zapisov, ki jo vsebujejo. Ključne besede zato lahko razumemo kot vrsto pojmov, ki so sami po sebi del vsebine vira, ki ga opisujejo, delujejo torej na opisni in ne semantični ravni.

Pojem [angl. term] pa je opredeljen kot beseda, fraza ali simbol, ki se v slovarjih, katalogih, indeksih ali podatkovnih bazah uporablja za opredelitev vsebine ali lastnosti dela. Pojem za razliko od ključnih besed viru lahko pripisuje tudi njegov pomen in hkrati ni nujno del njegove vsebine. (Reitz 2004).

Podobno analogijo lahko uporabimo tudi za opredelitev razlike med indeksiranjem s pomočjo meta ključnih besed in z družbenim označevanjem. **Metapodatki** (angl. metadata) so strukturirani podatki ali informacije, s katerimi opisujemo, pojasnjujemo in lociramo informacijske vire in ki nam omogočajo enostavnejše iskanje, dostop, uporabo in upravljanje s temi informacijskimi viri (Hodge 2001). Že stoletja se metapodatkovne sheme (z nizom ključnih besed in pojmov) sistematično uporabljajo v bibliotekarstvu in založništvu za opis in lociranje bibliografskega gradiva¹⁴, v svetovnem spletu pa so bile doslej aktualne predvsem na področju iskalnega marketinga.

Obiskovalci svetovnega spleta iščejo informacije večinoma s pomočjo spletnih iskalnikov, ki spletne strani indeksirajo in razvrstijo v svoj sistem po različnih kriterijih. Med pomembnejšimi so vsebina in strukturiranost spletne strani, njena obiskanost ter število vhodnih in izhodnih

¹⁴ Danes glavni označevalni jezik za opis bibliografskega gradiva je XML (eXtensible Markup Language), ki ga je konzorcij W3C (World Wide Web Consortium) predal v splošno uporabo leta 1998. Zaradi svoje enostavnosti, fleksibilnosti in omogočanja množične izmenjave podatkov v svetovnem spletu, je kmalu presegel svoj prvotni namen pokrivanja elektronskega založništva (Budimir 1999). XML v svetovnem spletu z dopolnjevanjem najbolj razširjenega jezika HTML odpravlja nekatere njegove pomanjkljivosti, zlasti počasnost, ki s širokopasovnimi internetnimi povezavami sicer ni več največji problem, ter neučinkovito iskanje. Jezik HTML namreč ni opisovalni jezik (metajezik), določa le, kako naj brskalnik uredi elemente spletne strani, medtem ko XML opisuje informacije in s tem določa pomen vsebine dokumenta.

hiperpovezav do drugih spletnih virov. Na vsebino in strukturo spletne strani ter na njen opis lahko za razliko od nekaterih drugih dejavnikov vpliva avtor sam, in sicer s pomočjo **meta ključnih besed** (angl. meta keywords), ki jih vstavi v izvorno kodo spletne strani¹⁵. Avtorji spletne strani društva Kalidej (kalidej.net) so na primer uporabili naslednje meta ključne besede:

```
<META content="tržno komuniciranje, trženje, marketing, kreativnost, ustvarjalnost, ideje, integrirano, odnosi z javnostmi, goldendrum, sof, magdalena, spletno, članki, prispevki, znanje, " name=keywords>
```

Meta keywords so torej bolj podobni Reitzevim ključnim besedam kot pojmom, saj naj bi (idealno) izražali nek objektivni povzetek vsebine spletnega mesta.

Oznake, s katerimi uporabniki v folksonomijah označujejo spletne vsebine, pa predstavljajo sicer neke vrste reinencijo metapodatiranja toda s pomembno razliko; uporabniki folksonomij oznake ustvarjajo večinoma za lastno uporabo, zato so te lahko izrazito subjektivne. Ker uporabnik sam določa pomen označeni vsebini, te ne ustrezajo le Reitzevi opredelitvi ključne besed temveč tudi pojma.

Terminološko s to ugotovitvijo nismo ničesar poenostavili, saj je, da bo zmeda popolna, uveljavljena sopomenka za tag beseda keyword oz. ključna beseda in bo večkrat uporabljena tudi v nadaljevanju naloge. Pomembno pa je razumevanje, da ključna beseda v spletnem okolju ne pomeni nujno isto kot v bibliografskem opisu vsebine. Glavna razlika izhaja iz odnosa do označene vsebine, ključne besede so v bibliografskih vsebinah kar se da objektivno določene, v folksonomijah pa so lahko tudi arbitrarne in povsem subjektivne, kar lahko najlažje ponazorimo s funkcijami oznak.

4.2 Proces označevanja

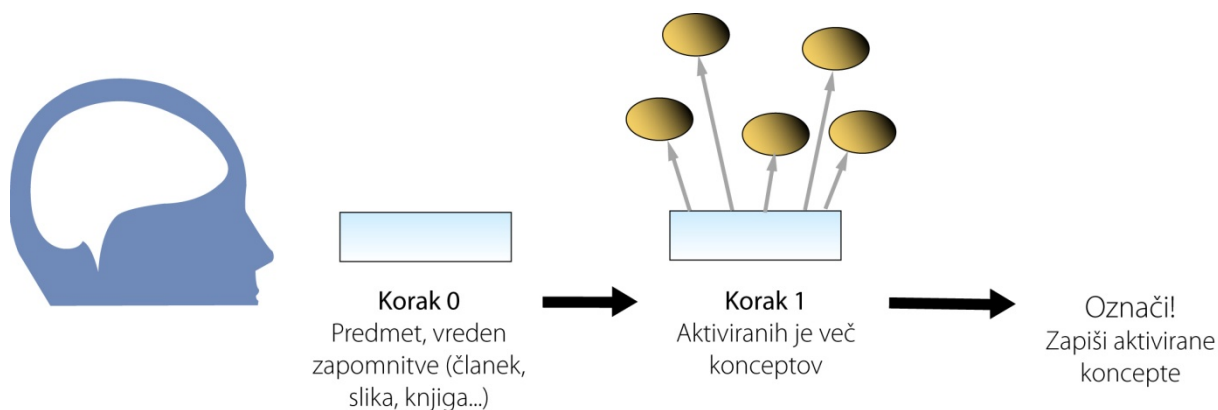
Proces označevanja lahko zaradi povratne informacije o kategoriziranju drugih uporabnikov razumemo kot kolektivni proces kategoriziranja informacij, ki poteka na zelo neformalen in ohlapen način. Gre za proces, ki za posameznika na individualni ravni pomeni predvsem **podeljevanja smisla ter pomena informacijam** (Golder in Huberman po Weicku 2007).

¹⁵ Zaradi zlorab določenega dela spletne industrije, ki je z irelevantnimi metapodatki preplavila spletne iskalnike, da bi s tem povečala obiskanost svojih spletnih strani, danes večina iskalnikov, med njimi tudi največji Google, ključne besede v izvorni kodi pri indeksiranju skoraj popolnoma ignorira in uveljavlja svoje metapodatkovne standarde.

Ta proces lahko opišemo tudi s terminom **generalizacija**, s katerim Roszak (1994, 89) loči informacije od idej. Generalizacijo razume kot osnovno dejanje inteligence, s katerim lahko opišemo in razberemo logične vzorce iz kopice podatkov in informacij. Pri tem Roszak (1994, 88) strogo loči informacije od idej. Informacije namreč niso nikoli produkt misli, prav tako informacije niso ideje, ki se jih lahko ustvarja in vrednoti le s pomočjo drugih idej. Zgolj ideje lahko vsebujejo informacije, jih definirajo in v končni stopnji ponovno ustvarjajo. Družbene oznake zato lahko opredelimo kot zelo splošne generalizacije, saj je njihov nastanek tesno povezan z mislijo, sploh kadar informacije ne le opisujemo, ampak tudi vrednotimo.

Za lažje razumevanje si oglejmo Sinhino (2005) primerjavo procesov označevanja in kategorizacije na kognitivnem nivoju. Sinha **označevanje** razume kot preprost proces, sestavljen iz dveh korakov, pri katerem gre za vrsto prostega asociiranja in označevanje predmetov z miselnimi koncepti, ki jih uporabnik poveže s predmetom (glej sliko 4.1).

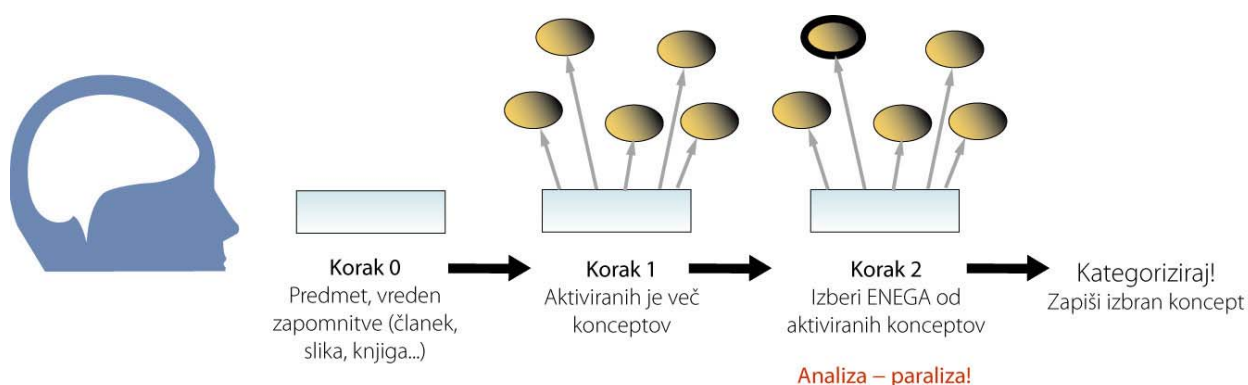
Slika 4.1: Kognitivni proces označevanja



Vir: Sinha 2005.

Kategorizacija pa proces podaljša za dodaten korak, saj od uporabnika zahteva odločitev za eno od asociacij (glej sliko 4.2). Izbira pogosto ni lahka, saj mora biti informacija shranjena tam, kjer jo bo uporabnik nekega dne najprej iskal. Zlasti v digitalnem okolju preobilja informacij mora biti odločitev preiščljena, ker so naknadni popravki kategorije težko izvedljivi in časovno zamudni.

Slika 4.2: Kognitivni proces označevanja in kategorizacije



Vir: Sinha (2005).

Označevanje zahteva od uporabnika, da torej razmisli o vsebini predmeta, ki ga označuje, ali da do njega zavzame tudi določeno stališče. To sicer zahteva določen miselni napor, a pomembna dodatna kognitivna korist, ki izhaja iz dejstva, da se moramo aktivno ukvarjati z vsebino, je, da si spletne vire tako lažje zapomnimo in da vemo, kje iskati, ko jih naslednjič potrebujemo. Narava označevanja je pogojena tudi s samo naravo informacije, zato je za celostno razumevanje označevanja, kot opozarja Shirky (v Weinroth 2005), ključen zlasti njegov kontekst: „... users tag differently on del.icio.us than on Flickr, so treating tags as purely atomic elements strips them of much of their value. A 'tag molecule', so to speak, includes not just the tag and the URI it pointed to, but the users, timestamp, service it was derived from, and other tags by that user pointing to the same URI. Getting that bit of encoding right seems to me to be the essential issue.“

4.3 Funkcije oznak

Golder in Huberman (2007) sta oznake v odnosu do vsebine razvrstila v 7 različnih kategorij:

Kategorija	Primer oznake
Na kaj (ali na koga) se predmet nanaša	“folksonomija” “splet 2.0”
Vrsta oz. zvrst	“diploma”
Avtorstvo ali lastništvo	“igor medved” “fdv”
Podrobnejša kategorija predmeta	“diplomska naloga o označevanju”
Lastnosti in značilnosti	“zanimivo” “dolgočasno”
Uporabnikovo razmerje do predmeta	“moja diploma”
Uporabnikov namen s predmetom	“preberi” “natisni”

Oznake, ki spadajo v prve tri kategorije, imajo širšo uporabno vrednost in bi jih brez težav lahko zamenjali tudi z metaoznaki. Te splošno razumljene oznake so zato pogosto uporabljene ključne besede številnih, med seboj nepovezanih uporabnikov folksonomij. Ob uporabi takšne oznake kot iskalnega pojma običajno torej dobimo veliko število zadetkov.

Za ostale štiri kategorije oznak pa velja, da imajo zelo nizko uporabno vrednost, saj so pogosto uporabljene samo enkrat. To so izrazito subjektivne oznake, ki pogosto opisujejo čustveno razmerje oz. odnos uporabnika do označenega predmeta in s tem ožijo uporabnost in razumevanje oznake zgolj na samega avtorja oznake. Skupnost od takšne izrazito subjektivne oznake torej nima pretirane koristi. Prav zaradi teh redkih oznak se oznake v sistemu porazdeljujejo po **zakonu moči**. Ta fenomen, ki ga opredeljuje teorija učinka dolgega repa, si zato oglejmo poglobljeno.

4.4 Učinek dolgega repa

Z naraščanjem števila uporabnikov v folksonomičnem sistemu eksponentno narašča tudi število različnih oznak ter pri tem sledi teoriji zakona moči (angl. power law). Med najbolj znanimi teorijami zakona moči sta vedno aktualno Paretovo načelo, ki pojasnjuje razporeditev moči v družbi in pravi, da 20 % ljudi razpolaga z 80 % bogastva ter Zipfovo odkritje o pogostosti pojavljanja besed v jeziku, ki temelji na spoznanju, da je v vsakem jeziku **nekaj** besed, ki imajo visoko frekvenco pojavljanja (*in, je*), **precej** besed, ki se pojavljajo pogosto (*knjiga, avto*) in **veliko** besed, ki se pojavljajo redkeje (*dnina, slamoreznica*).

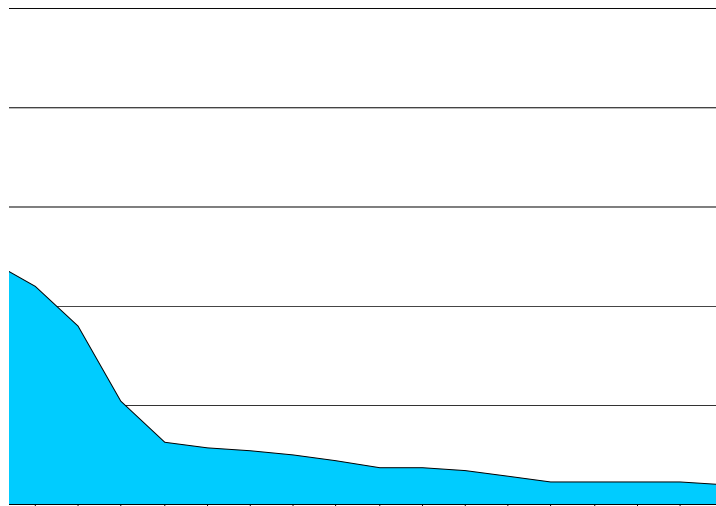
Kadar govorimo o folksonomijah, pa se za opis razporeditve oznak uporablja predvsem izraz **učinek dolgega repa** zanimive ekonomske teorije Chrisa Andersona (2004), ki opisuje povpraševanje po izdelkih zabavne industrije, kot so knjige, zgoščenke in filmi. Anderson ugotavlja, da Paretovo načelo 20-80 res drži; velik del prihodka ustvari dvajset odstotkov najbolj uspešnih izdelkov, tako imenovanih blockbusterjev, potem pa krivulja povpraševanja hitro pade, a se ne dotakne dna. Ta dolgi rep predstavlja ostalih 80 odstotkov izdelkov, ki so sicer manj tržno uspešni in priljubljeni, vendar pa tvorijo lepo nišo, ki so jo nekatera spletna podjetja, na primer Amazon in Netflix, znala spretno izkoristiti.¹⁶

Tej razporeditvi sledi tudi razporeditev oznak v folksonomijah, kjer zaradi dovoljenega

¹⁶ Po raziskavi Brynjolfssona (Brynjolfsson in drugi 2006) je med knjigami, ki jih prodaja Amazon, 30-40 odstotkov takih, ki jih v navadnih prodajalnah, predvsem zaradi njihovih prostorskih omejitev, ni mogoče dobiti.

poljubnega označevanja virov najdemo malo oznak, ki se pojavijo velikokrat in veliko oznak, ki se pojavijo v sistemu samo enkrat ali le nekajkrat. Primer takšne porazdelitve je tudi naslednji graf, ki prikazuje, kako so uporabniki označevali Andersenov članek na spletnem servisu Delicious.

Slika 4.3: Porazdelitev oznak v obliki dolgega repa



Na x osi število različnih oznak, na y osi pa število posamezne oznake.
Vir: Delicious (2008).

4.5 Tipologija folksonomij

Med posameznimi družbeno označevalnimi sistemi obstajajo pomembne razlike, ki jih Marlow in drugi (2006) razčlenijo v 7 tipoloških dimenzij:

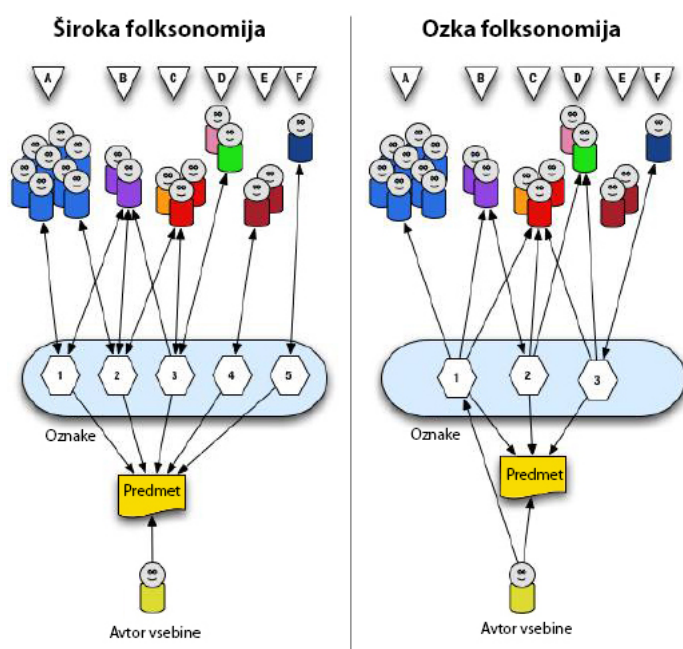
1. Pravice do označevanja
2. Pomoč pri označevanju
3. Zbiranje
4. Vrsta predmeta
5. Izvor predmeta
6. Povezanost predmetov
7. Družbena povezanost

1. Pravice do označevanja

Verjetno najpomembnejša tipološka dimenzija se nanaša na obseg pravic pri ustvarjanju, urejanju in brisanju oznak, ki sledi Walovi (2005) delitvi folksonomij na široke (označevanje vseh) in ozke (označevanje svojega). Za sistem **široke folksonomije** je značilno, da lahko

kdorkoli označi katerikoli vir. Uporabniki istemu predmetu dodajo različne oznake, s pomočjo katerih jih lahko kasneje znova prikličejo. Običajno je razporeditev oznak v takem sistemu v obliki dolgega repa. Za **ozko folksonomijo** pa je značilno, da predmet, na primer fotografijo na Flickru, lahko označi samo avtor vira in tisti uporabniki, ki jih za to pooblasti (na sliki 4.7.1 je slednje prikazano z dvosmernimi puščicami pri uporabnikih B in F). Število oznak za posamezni predmet je zato relativno majhno; ostali uporabniki lahko le iščejo po obstoječih oznakah, ne morejo pa dodajati svojih. Tako ozka folksonomija ima zaradi omejenih pravic označevanja predvsem drugačno, bolj ekspresivno logiko komunikacije, saj je v ospredje poleg označenih virov postavljen tudi njihov avtor/lastnik, ki je v širokih folksonomijah v smislu interakcije relativno nezanimiv del sistema. Ta razlika je opazna zlasti v spletni arhitekturi sistemov, Flickr denimo omogoča možnost neposrednega komentiranja označenih virov njihovih avtorjev, Delicious pa ne.

Slika 4.4: Široka in ozka folksonomija



Vir: Wal (2005).

Večina (širokih) folksonomij temelji na označevanju tujih vsebin za lasten prihodnji priklic. Zanimiva izjema je sistem Technorati, kamor se stekajo oznake avtorjev zapisov iz blogov in imajo podobno funkcijo kot meta oznake, torej z relevantnim opisom pritegniti na blog oziroma spletno stran čim več obiskovalcev.

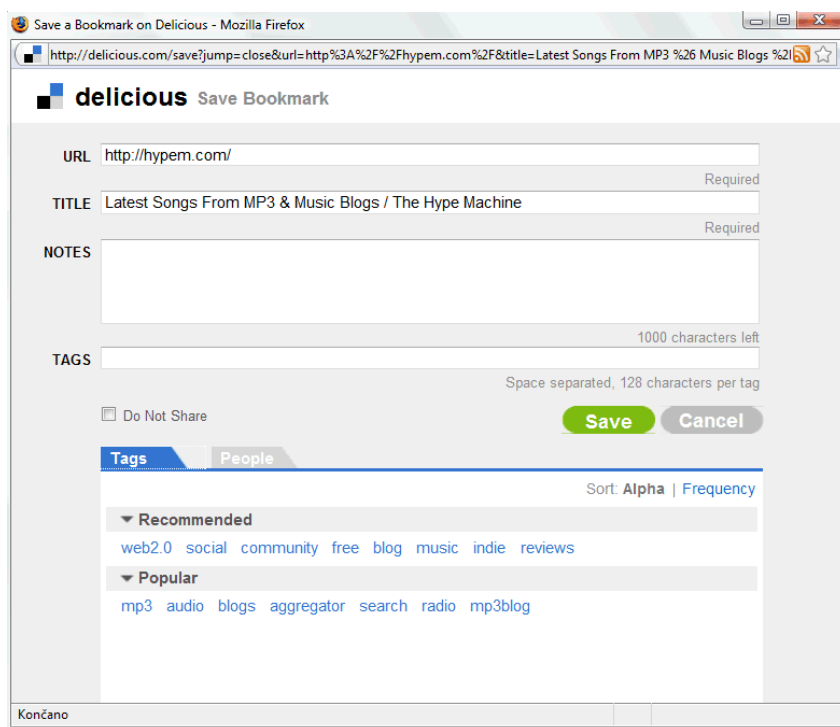
2. Pomoč pri označevanju

Marlow in drugi (2006) ločijo tri kategorije pomoči, ki so jih uporabniki lahko deležni pri

označevanju predmetov:

- slepo označevanje (uporabnik ne vidi oznak drugih uporabnikov)
- vidno označevanje (oznake drugih uporabnikov so vidne)
- označevanje s predlogi (sistem predlaga uporabniku možne oznake).

Slika 4.5: Pomoči pri označevanju s priporočenimi in priljubljenimi oznakami



Vir: Delicious (2009).

Vidno označevanje in označevanje vsebin s pomočjo predlogov ima lahko velik vpliv na porazdelitev oznak v sistemu. Bolj kot so oznake med uporabniki poenotene, bolj homogena je tudi folksonomija. Z manjšo raznolikostjo in večjo predvidljivostjo oznak pa, kot ugotavljata Guyjeva in Tonkinova (2006), folksonomije izgubljajo svojo pravo vrednost: „Obstaja resna nevarnost, da z urejanjem oznak sprejemamo uničujočo rešitev in izgubljammo dragocene metapodatke ... Tudi če predpostavljamo, da bi bil takšen dogovor mogoč, ali si res želimo kolaborativno ustvarjeno inačico zborne angleščine?“

3. Zbiranje

Označevalni sistemi se med seboj razlikujejo glede na način, kako se oznake za določen predmet zbirajo. Marlow in drugi (2006) ločijo dva modela:

- model vreče (“bag model”), značilen za Delicious, dovoljuje večkratno dodajanje oznake

predmetu, kar omogoča statističen pregled nad tem, ali med uporabniki obstaja konsenz o vsebini predmeta, ki ga označujejo;

- fiksni model ("set model") dovoljuje le enkratno dodajanje oznake predmetu (You tube). Oznake se naknadno torej ne da spremeniti.

4. Vrsta predmeta

Podobna raznovrstnost kot velja za oznake, velja tudi za same predmete označevanja. Kot je razvidno iz naslednje tabele, uporabniki lahko označujejo skoraj vse, kar je virtualno možno:

Spletna stran	Predmet označevanja
Delicious.com	spletne strani
flickr.com	fotografije in video
youtube.com	video
technorati.com	blogi
last.fm	glasba
yahoopodcasts.com	podcasti
friends4days.com	uporabniki
gmail.com	e-pošta
upcoming.yahoo.com	dogodki
citeulike.org	znanstveni viri
ning.com	družbena omrežja

5. Izvor predmeta

Spletne vsebine lahko v označevalni sistem prispevajo uporabniki (fotografije v Flickr), sistem sam (pesmi v Last.fm) ali pa je slednji odprt za označevanje tako novih kot že obstoječih predmetov (družbeni zaznamki v Delicious). Omejitve so pogojene z arhitekturo spletne strani (Flickr denimo ne omogoča objave zvočnega posnetka) ali z določenimi pravili (Cite U Like na primer dovoli samo shranjevanje družbenih zaznamkov z znanstvenega področja).

6. Povezanost predmetov

Predmeti v sistemu so lahko med seboj povezani, grupirani ali nepovezani. Pri tem igrajo oznake

uporabnikov ključno vlogo pri ustvarjanju povezav med sorodnimi predmeti. Ugotavljanje načina povezanosti in intenzivnosti povezav med enotami se avtorji lotevajo z analizo družbenih omrežij.

7. Družbena povezanost

Podobno kot viri so tudi uporabniki lahko v sistemu medsebojno povezani, nepovezani ali pa del skupine. Uporabniška povezanost je od sistema do sistema poimenovana različno, na primer *Friend* (Daily Motion), *Related Users* (Connotea) ali *Network* (Delicious). Vez med uporabniki je pogosto zgolj enosmerna in poteka praktično brez medsebojne komunikacije, pri čemer je edina skupna točka dveh uporabnikov njun skupno označeni predmet.

4.6 Slabosti družbenega označevanja

Oznaka je lahko med drugim poljubna črka, beseda, okrajšava, skupina besed in tudi lastno ime ali številka. Prosto razvrščanje informacij pomeni za uporabnika tako prednost kot tudi slabost. Guyjeva in Tonkinova (2005) opozarjata na nekaj večjih problemov v procesu označevanja, ki jih zaradi lažje preglednosti lahko strnemo v tri sklope:

1. Odsotnost skupnih pravil o zapisovanju oznak

Ker ne obstaja neka skupna konvencija o zapisovanju oznak, se pojavljajo različni zapisi ključnih besed, ki bi sicer lahko bili združeni v eno samo oznako. Tovrstne “napake” so opazne zlasti pri rabi:

- male in velike začetnice (Apple in apple)
- ednine in množine (apples in apple)
- posebnih jezikovnih znakov, ki so specifični za določen jezik (na primer zapisovanje slovenskih šumnikov)

2. Enkratne oznake

To so oznake, ki se v sistemu pojavijo le enkrat, razdelimo pa jih lahko na tri večje skupine:

- personalizirane oznake, brez pomena za druge uporabnike (“Tinin blog”)
- uporaba neobičajnih besednih zvez (geooznake¹⁷, datumi)

¹⁷ Možnost geografskega označevanja fotografij, kar pomeni, da lahko za vsako fotografijo natančno označimo, kje smo jo posneli.

- psevdo-hierarhično poimenovanje vsebin z oznakami z uporabo simbolov (oznake, uporabljene z namenom, da si uporabnik ustvari hierarhični seznam svojih oznak, bodisi po abecedi ali kakšnem drugem sistemu)
- slovnične napake, bodisi zaradi površnosti ali nepoznavanja slovničnih pravil oziroma jezika, v katerem uporabnik označi vsebino z oznako

Kot ugotavljata v raziskavi Guyjeva in Tonkinova (2005), je le 10–15 % oznak na Flickru in Delicious enkratnih. V primerjavi s predhodnimi raziskavami je trend tovrstnih oznak padajoč, kar si razlagata predvsem z večjo popularnostjo obeh spletnih strani in povečanim deležem identičnih “napačnih” oznak. Z možnostjo označevanja s pomočjo že ponujenih oznak pa lahko predvidevamo, da je ta trend postal le še bolj strm.

Vse bolj pereč je tudi problem **označevalnega onesnaževanja**(tag spamming) oziroma označevanje spletnih virov z namenom zavajanja drugih uporabnikov. Zaradi subjektivne narave označevanja, ki omogoča zelo svobodne interpretacije spletnih virov, odkrivanje zlorab predstavlja velik izziv. Koutrika in drugi (2007) kot najboljšo rešitev predlagajo moderatorja, ki bi sproti odstranjeval namerno zavajajoče označene vire, saj računalniški algoritmi pri fokusiranih napadih niso dovolj zanesljivi.

3. Razlike v pomenu

Beseda “pot” je enakozvočnica in ima lahko več pomenov. V slovenščini po njej lahko hodimo ali si ga obrišemo s čela. V angleščini (lonec) pa lahko v njem tudi kuhamo. Z oznako “pot” imamo lahko v mislih torej katerega koli izmed vsaj treh možnih pomenov.

Še precej večjo težavo predstavljajo sopomenke. Zaradi nekonsistentne rabe izrazov za isto stvar lahko zgrešimo velik del relevantnih zadetkov (vsebine povezane s televizijo so lahko denimo označene samo s “tv” ali samo s “televizija”, tako da je iskanje samo z eno ključno besedo nepopolno).

Golder in Huberman (2007) opozarjata tudi na relativnost posameznikove ravni specifičnega pri miselnih konceptih. Za nekoga je lahko pojem *javascript* preveč specifičen termin, za koga drugega pa pojem *programiranje* preohlapen, čeprav je vsebina smiselno označena z obema oznakama.

4.7 Motivacijski dejavniki označevanja

Locke (2005) folksonomijam zaradi zaradi individualne narave označevanja vsebin pravi tudi intimna omrežja, ker je komuniciranje med uporabniki sekundarnega pomena ali pa celo ni omogočeno in je sam proces ustvarjanja oznak neodvisno dejanje posameznika. Pri označevanju virov zaradi tega prihaja do razlik, nekateri označujejo zgolj zase, drugi pa poskušajo upoštevati tudi širšo družbeno korist. Z vprašanjem, zakaj ljudje označujejo in kako se komponenti družbenega in individualnega prepleteta skozi omrežne učinke folksonomij, se ukvarjajo številni avtorji.

Marlow in drugi (2006) opredelijo dva tipa motivacije označevanja, **organizacijskega** in **družbenega**. Pri prvem se uporabniki z izborom za folksonomije odločijo za alternativno od običajno taksonomične različice upravljanja informacij in pri tem razvijejo svoj osebni standard za označevanje virov – glavni motivacijski dejavnik je torej prihodnji priklic vira. Pri **družbenem tipu motivacije** pa je prisotna še dodana komunikacijska vrednost, saj izbira oznak uporabniku lahko služi tudi kot način izražanja in kot orodje za opredelitev specifičnih lastnosti označevanih vsebin. Slednje se po Marlowu (Marlow in drugi 2006) izraža skozi vzajemno deljenje informacij med uporabniki, z zbujanjem pozornosti s pomočjo neobičajnih oznak, z igro in tekmovanjem¹⁸, samonanašanjem (na primer z oznakami “prebral”, “videl v živo”) in z izražanjem mnenja o označenem viru s pomočjo oznake (na primer “zanimivo”, “neumno”).

Kjub navidezni preprosti uporabi folksonomij, njeni uporabniki lahko nastopajo v kar štirih vlogah (Voss 2007):

1. **Avtor vsebine** (oseba, ki ustvari vir)
2. **Zbiralec** (oseba, ki doda vir v označevalni sistem, a ga ne označi)
3. **Označevalec** (oseba, ki označuje vire)
4. **Raziskovalec** (oseba, ki išče vire s pomočjo oznak)

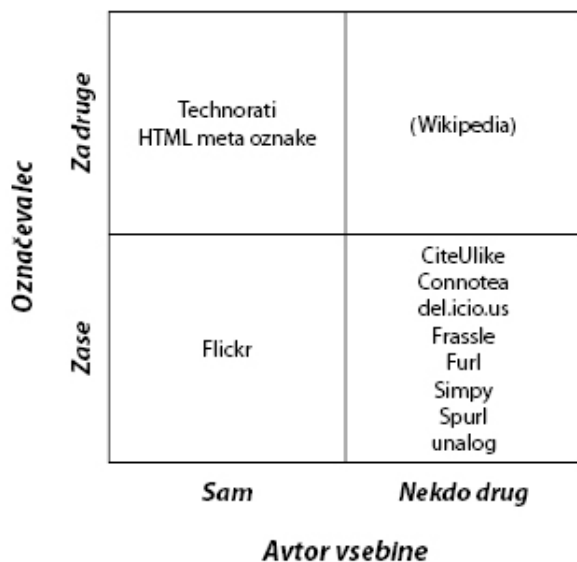
Uporabnik ponavadi nastopa hkrati v več vlogah; avtor bloga na Technorati je tako lahko avtor vira, zbiralec in označevalec, avtor družbenega zaznamka na Delicious zbiralec in označevalec, če kopira družbeni zaznamek od drugega uporabnika, pa raziskovalec in označevalec. Vezi, ki povezujejo uporabnike, pa so iste oznake in posledično lahko tudi isti spletni viri.

Uporabnike pri označevanju vodijo različni motivi, od povsem sebičnih (označevanje lastnih virov za lasten prihodnji priklic) do zelo altruističnih (označevanje drugih virov za priklic drugih

¹⁸ ESP Game je interaktivna igra v kateri dva tekmovalca hkrati označujeta naključno izbrane slike. Cilj igre je, da oba tekmovalca označita vir z enako oznako (Vir: <http://www.espgame.org/>. Dostopno 3. maj, 2008).

uporabnikov). Kot ugotavljajo Hammond in drugi (2005), na prakso označevanja vsebine vpliva zlasti narava aplikacije, kar je vidno tudi po naslednjem diagramu:

Slika 4.6: Kategorizacija sistemov družbenega označevanja



Vir: Hammond in drugi (2005).

Po tej kategorizaciji je najbolj sebičen sistem označevanja Flickr v levem spodnjem kvadrantu, najmanj pa Wikipedia v desnem zgornjem, ki sicer ni družbeno označevalni sistem, saj skupnost uporabnikov spletne enciklopedije namesto označevanja vsebin s prostimi oznakami te kategorizira s pomočjo skupne taksonomije.

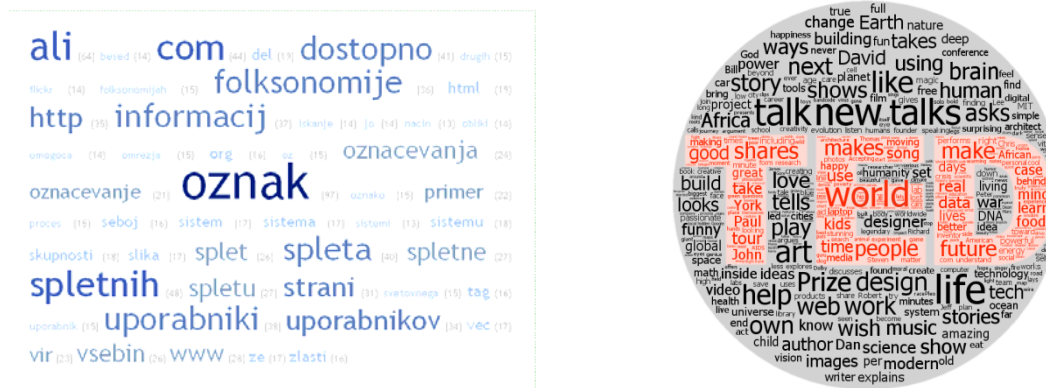
4.8 Vizualizacije oznak

Sočasno z označevalnimi sistemi se je pojavila tudi priljubljena grafična ponazoritev oznak, imenovana **oblak oznak** (angl. tag cloud). V njem se nahajajo najbolj pogosto uporabljene oznake, ki so običajno razporejene po abecednem vrstnem redu. Oblak oznak lahko razumemo kot orodje, s katerim ugotavljamo priljubljenost oznak, kar je grafično ponazorjeno z velikostjo njihovega izpisa. S klikom na določeno oznako v oblaku, se izpišejo vsi viri v sistemu, označeni s to oznako.

Hassan-Montero in Herrero-Solana (2006) opažata pomembno razlikovalno lastnost med priklicem informacij s pomočjo iskalne poizvedbe ali s pomočjo oblaka oznak. Pri prvem načinu mora uporabnik že ubesediti informacijo, ki jo išče, pri drugi pa mu brskanje po vmesniku

dovoljuje šele prepoznati informacijo, ki jo potrebuje. Vizualno brskanje je na ta način podobno branju hiperteksta, saj uporabniku omogoča iskanje z brskanjem.

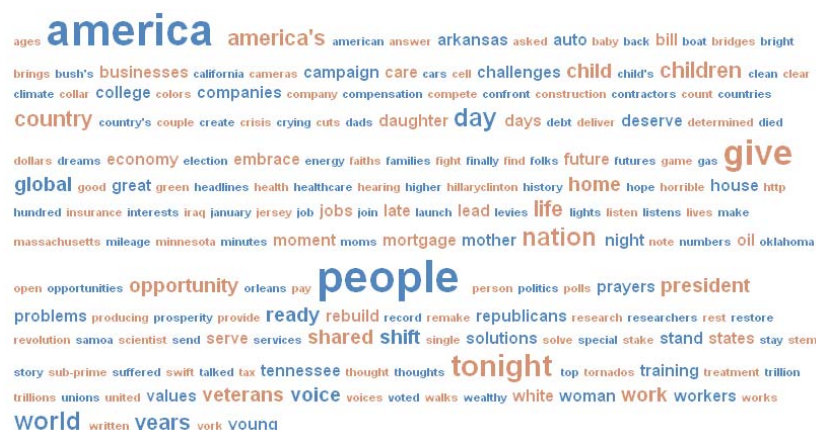
Slika 4.7: Oblak oznak



Klasičen oblak oznak na levi in njegova bolj barvita variacija na desni strani. Vir: Clark (2009).

Oblak oznak je lahko priročno orodje za ponazoritev vrste informacij, ki niso nujno povezane z oznakami ali označevalnimi sistemi, zato se bo njihov razvoj v prihodnje verjetno nadaljeval in uporaba širila tudi na druga področja. Lepa ponazoritev za to je naslednji primer aplikacije strukture oblaka oznak za ponazoritev miselnega toka ameriških demokratiških predsedniških kandidatov, Hillary Clinton in Baracka Obame, skozi analizo pogostosti uporabe besed v njihovih govorih. Oba kandidata sta imela govor na isti dan, 5. februarja 2008, na tako imenovani Veliki torek, ko so primarne ameriške volitve za demokratiško nominacijo potekale v 11 ameriških zveznih državah. Clintonova je največkrat uporabila besedi *America* in *people*, Obama pa besedo, na kateri je temeljila njegova politična kampanja, *change*.

Slika 4.8: Vizualizacija najbolj pogosto uporabljenih besed Hillary Clinton



Vir: Many Eyes (2009).

Slika 4.9: Vizualizacija najbolj uporabljenih besed Baracka Obame



Vir: Many Eyes (2009).

Zanimiva oblika prikaza oznak so tudi **zemljevidi oznak** (angl. tag maps), v katerih je pozicija besed določena z geografsko pozicijo označenega predmeta, velikost oznake pa ponazarja pomembnost pojma na specifični lokaciji. Yahoojeva aplikacija World explorer omogoča sočasen prikaz zemljevida in fotografij izbrane lokacije, ki so jih uporabniki tam posneli in naložili na spletni servis Flickr. Fotografije so pri tem označevali s tako imenovanimi geooznakami (angl. Geotags), ki vsebujejo informacije o geografski poziciji predmeta, torej natančno časovno širino in dolžino lokacije.

Slika 4.10 Aplikacija Worldexplorer



Zemljevid oznak lokacij San Francisca, desno izbor slik iz Flickr za oznako Golden Gate Bridge.
Vir: WorldExplorer (5. maj 2008).

Zemljevid oznak je tudi primer specifične nove pojavnosti oblike aplikacij spleta 2.0, tako imenovanih **lepljenk** (angl. mash up), ki temeljijo na principu transkluzije, vključevanju delov različnih dokumentov v drug dokument v hipertekst, s sklicevanjem na vire, s čimer se radikalno spreminja izkušnja branja. Lepljenka črpa podatke iz različnih virov (v našem primeru Yahoo Maps in Flickr), ki jih potem združi v eno aplikacijo (World Explorer) in tako ponudi popolnoma novo storitev.

5 Semantična raven folksonomij kot informacijskega sistema

5.1 Metapodatki semantičnega spleta

Semantični splet temelji na metapodatkih, ki opisujejo in povezujejo podatke oziroma znanje. Njihova glavna pomanjkljivost je različnost formatov, v katerih so zapisani, in s tem posledično neizmenljivosti. Večina sistemov jih je namreč zaklenila v zaščitene programe, s čimer ne sledijo logiki spleta, ki temelji prav na odprtosti njihove izmenjave. Rezultat prizadevanj po nekem skupnem formatu za njihovo zapisovanje in s tem zahtevani interoperabilnosti je standard RDF¹⁹, ki mu Tauberer (2008) pravi metoda za razgradnjo informacije na posamezne koške. V obliki trojčkov (osebek, povedek, predmet) omogoča ne le opis lastnosti (spletnih) virov, ampak tudi razmerij med njimi. To ga denimo pomembno loči od enostavnega zapisa v obliki formata XML, ki vsebuje podatke le o konceptu, ki ga opisuje, saj RDF s povezovanjem razpršenih informacij različnih virov ustvarja tudi znanje in pomen o opisanih konceptih.²⁰

Oglejmo si opis RDF na naslednjem primeru:

Osebek	Povedek	Predmet
Tomaž	ima	lokal

Osebek in predmet sta povezana s povedkom, ki opisuje njuno razmerje, v tem primeru lastništvo. Vse tri entitete so zapisane v obliki URI-ja²¹, ki poleg imena opredeli tudi točno lokacijo vira. To je nujno, saj sta mišljena točno določen Tomaž in točno določen lokal, zato ju ne moremo zamešati s katerim drugim spletnim virom, ki bi bil sicer lahko enako poimenovan.

Z nadaljnjim nizom stavkov v obliki trojčkov lahko opišemo celotno domeno, denimo biologijo, rastlinstvo ali v našem primeru ljubljansko gostinstvo. Domeno tovrstnih opisov posameznega področja imenujemo **ontologija**, ki je tipična hierarhična podatkovna struktura, ki vsebuje vse relevantne entitete, njihove relacije in pravila znotraj domene (Brejc in Jeričko 2005). Glavni prednosti ontologij sta predvsem enostavnost in fleksibilnost njihovega ustvarjanja, brez težav jih lahko namreč sproti dopolnjujemo in povezujemo z drugimi ontologijami. Pri povezovanju podatkov z novimi podatki se ontologije dodatno bogatijo in ustvarjajo novo znanje, čemur

¹⁹ Resource Definition Framework (RDF) oziroma opisni okvir vira.

²⁰ Obstajajo še naprednejši formati, med katerimi je trenutno najbolj razširjen jezik OWL.

²¹ Uniform Resource Identifier (URI) oziroma enotni označevalnik vira (na primer v obliki klasične spletne povezave <http://www...>).

Berners-Lee in drugi (2001) pravijo **učinek verige vrednosti** [angl. chain value effect].

S semantičnim spletom in formati kot sta RDF in OWL prehajamo torej iz spleta dokumentov v splet podatkov. Spivack (2007) pri tem celo reciklira termin *hiper*, ko pravi, da za podatke to pomeni vsaj tako veliko stvar, kot je za tekst pomenila iznajdba hiperteksta. V semantičnem spletu podatki niso več le podatki, ampak z vso dodatno, v ozadju skrito semantično prtljago, postajajo *hiperpodatki*. Da bomo lažje razumeli, kaj se se skriva za tem pojmom, opredelimo in definirajmo najprej hipertekst.

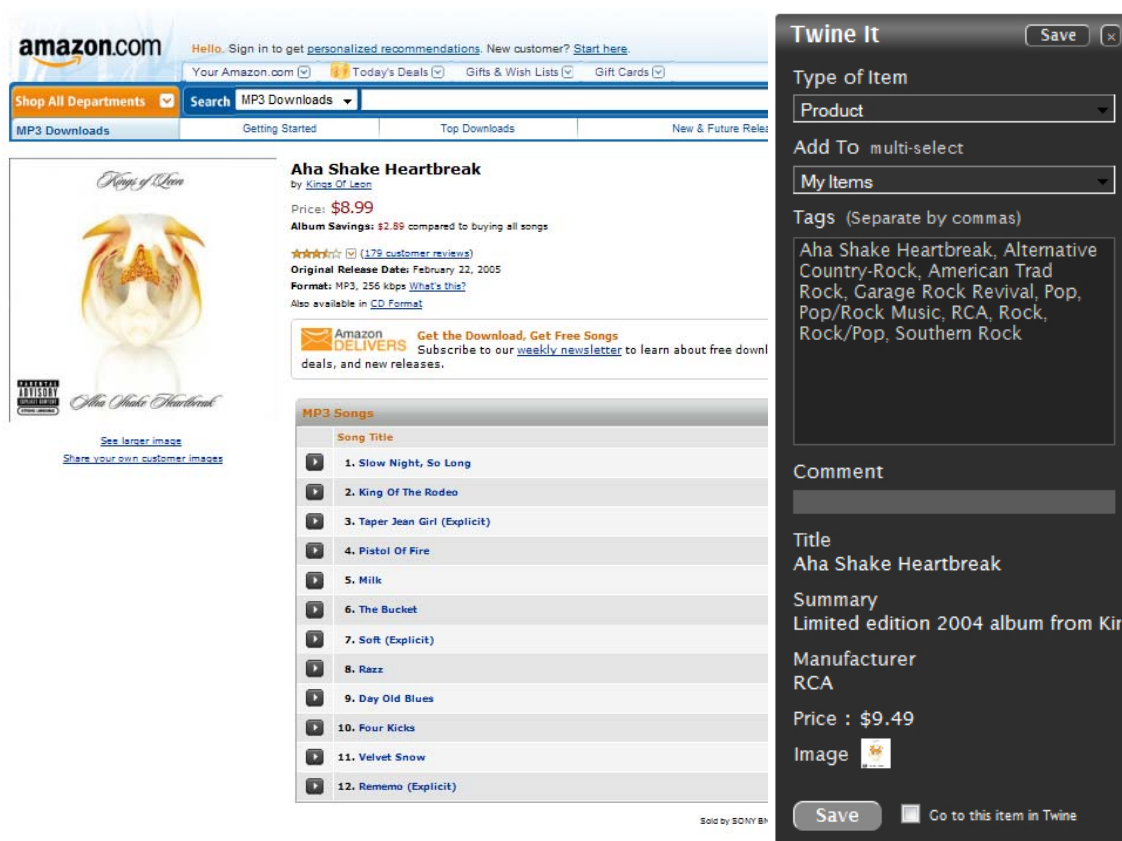
5.2 Semantične folksonomije

Na prvi pogled gre za dokaj paradoksalen koncept. Bistvena lastnost folksonomij je odsotnost nekih semantičnih pravil, ontologije pa zahtevajo definiranost in točnost pomenov, zato se samo po sebi zastavlja vprašanje, kako v uporabno celoto smiselno združiti oba koncepta oziroma ali je to sploh mogoče. Se naj folksonomije prilagajajo ontologijam in na ta način izgube tiste kvalitete, zaradi katerih so v spletnem okolju uspele preseči uporabnost in popularnost taksonomij, ali pa naj se ontologije prilagodijo folksonomijam in tvegajo okuženje s slabimi, nesemantičnimi metapodatki?

Kakšno vlogo bodo folksonomije igrale v prihajajočem semantičnem spletu tako še ni prav jasno, toda številni avtorji vidijo koristi v njihovem povezovanju. Prizadevanja denimo Spivacka (2005) in Gruberja (2007) gredo v smeri, da se ideji folksonomij in ontologij povežeta v tako imenovane na novo izgrajene **označevalne ontologije** (angl. tag ontology), kar bi lahko zaradi dodane semantične plasti razširilo uporabo folksonomij, približalo semantični splet povprečnemu uporabniku in kar je morda še najpomembneje, bistveno pospešilo njegov razvoj. Označevalna ontologija v praksi pomeni torej neko novo vrsto porajajoče se ljudske oziroma družbene ontologije, ki ji Spivack (2005) duhovito pravi tudi folkologija, saj briše meje med obema konceptoma.

Eden izmed novjših in bolj odmevnih tovrstnih semantičnih projektov je Twine.org, ki podobno kot Delicious omogoča osebno organizacijo spletnih vsebin, ki jih lahko delimo tudi z drugimi uporabniki. Twine trenutno omogoča shranjevanje petih različnih vrst vsebin: spletnih strani, izdelkov, knjig, slik in videov. Naslednja slika prikazuje način shranjevanja izdelka – zgoščenke (s strani Amazon.com), s pomočjo naprednega vmesnika (bookmarkleta).

Slika 5.1: Twineov bookmarklet



Vir: Twine (2009).

Če uporabimo Marlowovo tipologijo, označevanje s Twineom poteka s pomočjo že podanih predlogov, toda ne na podlagi oznak drugih uporabnikov, ampak glede na predloge, ki jih iz besedila izlušči računalnik. Če so oznake zgrešene, jih lahko uporabnik popravi ali doda svoje. Avtomatsko se iz spletne strani izluščijo tudi naslednje informacije:

- naslov izdelka
- povzetek
- proizvajalec oz. založnik
- cena
- slika izdelka.

Na strani twine.com pa se označeni (in tako shranjeni) izdelek izpiše v obliki, prikazani na naslednji sliki.

Slika 5.2: Shranjen izdelek na Twine.com

My Items

Items I Added (258) | Shared with You (5) | Shared by You (0) | Collected Items (2) | All Items

Share | Collect | Add Detail

Aha Shake Heartbreak
Product added by Igor on 11/10/2008
Private

Kings of Leon
Aha Shake Heartbreak

Description
Limited edition 2004 album from Kings Of Leon featuring the first single 'The Bucket'. Hand Me Down. H NME described it as, 'one of the best debut albums of the last ten years', Kings Of Leon have their wor strangely titled 'A-Ha Shake Heartbreak?', is more than a match for its illustrious predecessor. Fast pa Shake Heartbreak?, is a shoe-in for one of the best albums of 2004. Nomadic Southern evangelist Leon three sons and a nephew—in Tennessee quartet Kings of Leon, whose second album spins enough cus 20-something longhairs also fuel up on a filthy shotgun bass and relentless guitar riffs, complementing t 35-minute grab bag of ...

Original URL <http://www.amazon.com/Aha-Shake-Heartbreak-Kings-Leon/dp/B00078XKD4/ref=p>

Publisher RCA

Release Date Feb 22, 2005 (3 years ago)

Price \$11.98

Image http://ecx.images-amazon.com/images/I/41BpYVyuImL_SL160_.jpg

Detail Page <http://www.amazon.com/exec/obidos/ASIN/B00078XKD4/wwwradarnetwo-20>

Product Type Music

Places (1)
Tennessee [x]
Add places

People (3)
Caleb Followill [x]
Dolly Parton [x]
Leon Followill [x]
Add people

Organizations (0)

Other tags (14)
A-Ha Shake Heartbreak [x]
Aha Shake Heartbreak [x]
Alternative Country-Rock [x]
American Trad Rock [x]
Garage Rock Revival [x]
Me Down [x]
New Musical Express [x]
Pop [x]
Pop/Rock Music [x]
RCA [x]
Rock [x]
Rock/Pop [x]
Southern Rock [x]
The Bucket [x]
Add other tags

Recommended Items
No recommended items at the moment.

Vir: Twine (2009).

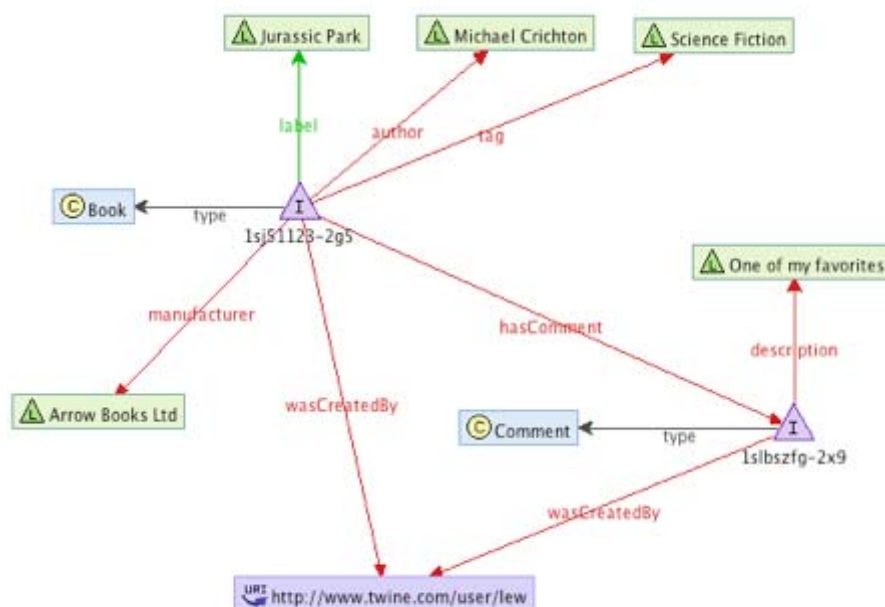
Kot je razvidno iz sklopa oznak na desni strani, računalnik zajetim informacijam doda še nekaj drugih metapodatkov. Inteligentnost Twina se kaže predvsem v tem, da je s pomočjo formata RDF in ontologij v ozadju (v tem primeru denimo ontologije MusicOntology) sposoben razumeti in na podlagi tega med seboj ločiti metapodatke po posameznih kategorijah. Tako razume, da je Tennessee kraj in da je Dolly Parton²² oseba.

Zanimiva je tudi vizualna predstavitev shranjenih predmetov v obliki semantičnega grafa, ki prikazuje povezave do označenega predmeta. Spodnji primer prikazuje semantični graf za knjigo

²² Neposrečna prisotnost Dolly Parton med semantičnimi oznakami kaže na nekatere pomanjkljivosti sistema, saj se na tem mestu pojavi le zato, ker je kot sokrajanka članov ocenjevane skupine mimogrede omenjena v Amazonovi recenziji te zgoščenke.

Jurski park in povezave do avtorja, založnika, vrste predmeta, zvrsti, oznak in tudi do drugih, uporabniško generiranih informacij; do uporabnika, ki je shranil ta vir in do komentarjev na ta vir.

Slika 5.3: Semantični graf Twine za knjigo Jurski park



Vir: Twine (2009).

Uporabniško ustvarjene informacije sproti dodatno bogatijo graf in predvsem z njihovo pomočjo se Twine uči o posameznikovih interesih, na podlagi katerih lahko uporabniku priporoča povezovanje s podobnimi uporabniki, skupinami ali predmeti: „Poganja ga semantično razumevanje, kar pomeni, da Twine spoznava tebe. Avtomatsko se uči o tvojih interesih in ustvarja povezave in priporočila, krojena tvojim potrebam. Najboljši del: več kot uporabljaš Twine, bolj uporaben postaja.“ (Twine 2008).

Poleg Twinea velja na tem mestu omeniti še Zigtag.com, Faviki.com in Entity Decriber (Connotea.org). Vsi omenjeni omogočajo shranjevanje virov z oznakami, ki imajo točno definiran pomen, a za razliko od Twina, kjer se semantični pomen oznak določi avtomatsko, mora pri teh servisih uporabnik v procesu označevanja sam izbrati pravo definicijo oznake. Spodnja slika prikazuje zigzagov bookmarklet, uporabljen pri shranjevanju spletne strani ameriškega podjetja Apple Inc. (Apple.com).

Slika 5.4: Zigtag



Vir: Zigtag (2009).

Vmesnik pri dodajanju oznake *Apple* ponudi vrsto definicij, med katerimi mora uporabnik izbrati pravo, kar je lahko sicer zamudno, vendar omogoča označevanje virov s skoraj taksonomično natančnostjo.

Analogija med paroma taksonomija–folksonomija in ontologija–folktologija je več kot očitna; podobno kot taksonomije so tudi ontologije zaradi natančnega opisovanja in definiranja konceptov izredno počasi rastoči sistemi, svetovni splet pa kot vrelišče idej in znanja preveč neučakan medij, da bi sedel križem rok in čakal na popolno rešitev. Označevalne ontologije so zato podobno kot folksonomije odličen primer realizacije spleta kot medija kolaborativne akcije, ki slej ko prej sam generira sprva sicer morda ne najbolj optimalne rešitve, a jih potem po principu evolucije sproti izboljšuje. Kot razmišlja Mika (2007), se v arhitekturi semantičnega spleta kritični del namreč nanaša prav na družbeno oziroma na uporabnika, ki prek ustvarjanja in vzdrževanja ontologij pravzaprav sam gradi semantični splet, četudi bo ta v končni fazi namenjen računalnikom. V tem kontekstu opozarja na asociativno komponento porajajočih se

ontologij, ki je kot eden ključnih kognitivnih procesov v folksonomijah seveda opazen zlasti pri označevanju, kjer mora uporabnik v nekaj sekundah najti ustrezne besede za opis predmeta, medtem ko je v strogo klasičnih ontologijah opazna predvsem njegova odsotnost.

Na tem mestu omenimo še eno sorodno lastnost, skupno obema konceptoma, in sicer pogajanje za pomen vira. Podobno kot v folksonomijah, kjer uporabniki to zavedno ali nezavedno počno z izbiro oznak (Mathes 2004), lahko opazimo to tudi v označevalnih ontologijah pri izbiri definicije oznak. Uporabnik ima tako možnost, da pri izbiri oznake natančno opredeli, kaj ta oznaka za njega in spletni vir pomeni. Ker se definicije za potrebe označevalnih ontologij črpajo iz spletnih virov, med drugim tudi iz sicer akademsko prav tako nepriznane Wikipedie s pomočjo sistema DBpedia (www.dbpedia.org), katerega se denimo poslužuje semantična folksonomija Faviki (www.faviki.org), to pomeni, da nikoli niso dokončne, ampak vedno podvržene presoji spletne skupnosti. Za tako odprte ontologije z vedno k pogajanjem podvrženim definicijam, seveda ne moremo reči, da bodo bolj točne od denimo klasičnih ontologij, bodo pa prav gotovo bližje potrebam širše spletne skupnosti, kar se je zgodilo tudi z vse prej kot natančnimi folksonomijami.

Omenimo za konec še folksonomije, ki niso bile ustvarjene z namenom uporabe v semantičnem spletu. Angeletou in drugi (2007) so že dokazali na primerih Flickrja in Deliciousa, da se tudi že izgrajene folksonomije da, v sicer omejenem obsegu, avtomatsko bogatiti s koriščenjem profesionalnih ontologij. Podobno je na primeru Delicious ugotovil Mika (2007), ki je s pomočjo oznak, nanašajočih se le na semantični splet, uspel iz klasične folksonomije izluščiti dokaj natančno in relevantno ontologijo, ki je pokrivala področje semantičnega spleta.

Vsem težavam, ki pri integraciji v semantični splet še čakajo nesemantične folksonomije, se v tem poglavju omenjeni semantični označevalni sistemi s sprotnim semantičnim izgrajevanjem oznak torej spretno izognejo. Z natančno, sicer uporabniško generirano definicijo, ki v ozadju spremlja oznako, prinašajo celo nekatere v družbeno označevalnih sistemih doslej pogrešane taksonomične kvalitete. Po eni strani ohranjajo organskost, fleksibilnost in hitrost folksonomičnega organiziranja informacij, po drugi strani pa ne dopuščajo več oznak brez pomena. Kljub temu omenimo, da tovrstne ontologije po natančnosti vendarle niso primerljive s pravo ontologijo, kjer je definicija vseobsežna in popolnoma formalizirana. Na področju ontologij lahko tako poleg natančnih in specifičnih ontologij, kjer bodo opisi in razmerja med koncepti nedvoumno zapisani (na primer v medicini), pričakujemo tudi sobivanje manj profesionalnih ontologij, ki bodo predvsem plod različnih spletnih kolaborativnih projektov. Tu

imamo v mislih torej tudi koriščenje podatkov obstoječih (klasičnih) folksonomij, na primer Flickr ali Delicious, pri čemer bodo dodatno prišle do izraza vse njihove poznane slabosti in pomanjkljivosti. Proste oznake, uporabljene v folksonomijah, namreč nimajo eksplicitnega pomena in so lahko pogosto ne le pomensko, ampak tudi slovnično nepravilne, pri čemer so največji problemi že omenjena večjezičnost, polisemični, sinonimni ali slovnično nepravilni izrazi, različna leksikološka oblika zapisov (na primer ednina ali množina), personalizirane oznake ter različne stopnje njihove specifičnosti. Euzenat (2001) je sicer v drugem kontekstu, ki pa ga lahko prenesemo tudi na naš primer, že ponudil tri strategije soočenja s tovrstnimi arbitrarnimi metapodatki pri njihovem povezovanju v semantične ontologije:

- izogibanje: stroški površnih metapodatkov so previsoki, da bi jih dopuščali;
- toleriranje: kljub nepopolnosti metapodatkov so koristi zajema večje od škode;
- čiščenje: sprotno izboljševanje sistema s pomočjo uporabnikov in računalnikov.

Uporaba posamezne strategije bo verjetno odvisna od tipa ontologije. Strokovne ontologije, ki bodo zahtevale eksplicitnost opisov, bodo najbrž temeljile na strategiji izogibanja uporabniško generiranih metapodatkov, uporabniške, ki bodo dinamične in sproščene, pa na toleriranju njihove nepopolnosti. Prav tako bo nek dinamičen pristop zahtevalo povezovanje strokovnih in uporabniških ontologij, kjer se kot najboljša rešitev sama po sebi ponuja tretja možnost, sprotno odpravljanje nepravilnosti in pomanjkljivosti.

5.3 Družbena raven folksonomij

S spletom 2.0 se je pojavila nova generacija družbenih omrežij, kjer v središču zanimanja niso nujno le člani omrežij, pač pa tudi skupna zanimanja ali hobiji, ki jih povezujejo. Za povprečnega uporabnika folksonomij tako denimo ni bistvenega pomena komuniciranje z drugimi uporabniki, temveč predvsem označevanje spletnih virov in raziskovanje označenih virov drugih uporabnikov. Po eni strani je opazna torej izrazito personalizirana in individualna raba sistema oznak ter uporaba folksonomij kot neke vrste osebnega spletnega skladišča, po drugi strani pa je zaradi splošne vidnosti oznak v omrežju in prednosti, ki jih ta lastnost prinaša, zelo pomembna tudi njihova širša družbena dimenzija. Razmerje med tema dvema motivoma, ki vodita uporabnike pri označevanju, zato velja osvetliti skozi nekatere obstoječe teorije omrežij oziroma skupnosti.

Zaradi svoje kompleksnosti je polje folksonomij idealno za proučevanje odnosov in relacij med

njenimi elementi, s čimer se podrobneje ukvarja uporabna znanost analize omrežij. Družbeno omrežje lahko definiramo kot družbena strukturo, sestavljeno iz več akterjev, ki so med seboj povezani. V našem kontekstu jih lahko opredelimo kot tripartitni sistem uporabnikov, virov v obliki spletnih vsebin in oznak. Do danes opravljene raziskave in preučevanja lastnosti omrežij in relacij med akterji v njih, so razkrile nekatere njihove zanimive lastnosti. Schmitz in drugi (2007) v njih prepoznajo obstoj **(ohlapno povezanih) skupnosti**, Mika (2006) pa jih z vidika preučevanja realizacije semantičnega spleta, opredeljuje kot **semantična družbena omrežja**. Schmitz in drugi (2007) so na primerih Delicious in Bibsonomy tudi ugotovili, da grafa družbenih omrežij teh folksonomij izkazujeta **lastnosti malega sveta**. To lastnost je prvi opisal Milgram (1967), ki je z eksperimentom ugotovil, da je povprečna dolžina najkrajše poti med dvema posameznikoma v ZDA, ki se sicer ne poznata, približno 6 medsebojnih poznanstev. Ta pot je v folksonomijah še precej krajša; v povprečju lahko v omrežju Delicious vsakega uporabnika, oznako ali vir lahko dosežemo z le 3,5 kliki.

5.3.1 Folksonomije – virtualne skupnosti?

Iskanje skupnih točk folksonomij in virtualnih skupnosti je z vidika njihovih starejših definicij precej nevhvaležna naloga, zlasti ob upoštevanju današnje, močno spremenjene narave spleta. Folksonomije denimo še vzdržijo zelo ohlapno opredelitev skupnosti Hammana (1997 v Trček 2003, 24), ki jih razume kot skupino ljudi, ki sodelujejo v družbeni interakciji in jih povezujejo nekatere skupne vezi med njimi in ostalimi člani skupine ter delijo vsaj nekaj časa skupni prostor, le delno pa to velja za starejšo Rheingoldovo definicijo virtualne skupnosti, ki jo opredeli „kot družbeni agregat, ki vznikne na internetu takrat, ko zadostno število ljudi nadaljuje s svojimi razpravami dovolj dolgo in s pravšnjim človeškim občutkom, z namenom, da v kibernetnem prostoru oblikujejo mreže osebnih odnosov.“ (1993 v Oblak 2000, 1063).

V večini folksonomij o poteku razprav in oblikovanju mrež osebnih odnosov (za zdaj) ne moremo govoriti. V nasprotju s tradicionalnimi skupnostmi so virtualne skupnosti neformalne entitete, v katerih norme ne dominirajo tako kot v tradicionalnih skupnostih, kjer se člani med seboj vidijo. Johnson (2001) na podlagi te ugotovitve virtualno skupnost definira kot “skupino ločeno s prostorom in časom, ki uporablja omrežno tehnologijo v takšni ali drugačni obliki za sodelovanje in komuniciranje.”

Pri opredeljevanju folksonomij z vidika virtualnih skupnosti se je zato smiselno najprej osredotočiti na skupno značilnost vseh zgornjih definicij, torej na interakcijo med člani omrežij,

česar v folksonomijah na prvi pogled ni opaziti. Kljub temu, da se folksonomije med seboj v marsičem razlikujejo in da obstajajo sistemi, kjer je poleg dejavnosti označevanja omogočeno tudi medsebojno spoznavanje uporabnikov (Flickr na primer omogoča funkcijo komentarjev in posebnih razpravljalnih desk), to namreč ni glavni fokus storitve ali ključni del uporabnikove izkušnje pri označevanju. To dodatno, a ne osrednjo lastnost sistema, bomo v nadaljevanju zato namerno zanemarili. Interes za drugega uporabnika se v folksonomijah namreč večinoma izraža v zanimanju za njegove označene vire in ne za njega osebno. Toda ker v folksonomijah individualno obnašanje poteka znotraj javnega, posameznikovo ravnanje lahko vpliva ali pa je pod vplivom drugih uporabnikov kljub odsotnosti neposrednega komuniciranja. Interaktivnost, kot osrednja strukturna odlika virtualnega medija, ki omogoča nastanek skupnosti (Trček 2003), poteka v folksonomijah na subtilnejši ravni kot v prevladujočih tipih virtualnih skupnostih spletnih forumov in klepetalnic, vendar je vsekakor prisotna, in sicer kot integralni del samega procesa označevanja. Omogoča jo takojšnja povratna informacija na posameznikovo izbiro oznake, ki jo lahko razumemo kot družbeni vidik folksonomij:

V trenutku, ko določiš oznako viru, zagledaš svojo skupino virov z isto oznako. V kolikor nisi zadovoljen, lahko oznako spremeniš ali dodaš novo. ... Toda prava moč izhaja z razširitvijo pogleda na vire vseh uporabnikov, ki se ujemajo s tvojo oznako. Spet rezultat morda ni v skladu s tvojimi pričakovanji. V tem primeru se lahko prilagodiš normi skupine, v poskusu vpliva na normo skupine obdržiš svojo oznako ali pa narediš oboje. (Udell v Mathes 2004).

Mathes (2004) temu pravi asimetrično komuniciranje med uporabniki skozi uporabo metapodatkov, pri čemer interakcijo med uporabniki z individualnim izbiranjem oznak razume tudi kot namerno ali nenamerno pogajanje za pomen vira.

Še bolj zanimivo od realizacije interaktivnosti pa je opazovati, kako je v folksonomijah realizirana druga razsežnost Johnsonove definicije virtualnih skupnosti, in sicer **sodelovanje**. Za tradicionalno skupnost velja, da njeni člani zaradi različno porazdeljenega znanja oziroma strokovnosti sčasoma v njej pridobijo določen sloves, naj si bo ta slab ali dober. Russell (2005) meni, da član skupnosti nikoli nima le enega slovesa, ampak si jih glede na različna področja delovanja oblikuje več; nekdo je lahko recimo ugleden matematik, a zelo povprečen kuhar. Podobno lastnost opazimo tudi v folksonomičnih omrežjih, virom nekaterih uporabnikov zaupamo bolj kot drugim, saj jih prepoznamo kot eksperte na področju, ki ga označujejo. Te uporabnike lahko označimo in shranimo v osebni seznam ali pa se na njihove vire naročimo z

RSS podajalnikom. Vidnost oznak v folksonomijah omogoča torej dinamičen, fleksibilen in decentraliziran pretok informacij in znanja med uporabniki, ki vire sicer označujejo zase, a hkrati z vidnostjo svojih oznak pomagajo tudi ostalim.²³

Ta egocentrično-altruističen preplet uporabe je po Fayju (2006) tudi sicer ena temeljnih značilnosti spleta 2.0. Izraža se v veliki **personalizaciji** uporabe spleta, od soustvarjanja spletnih vsebin do samourednikovanja s pomočjo RSS tehnologije²⁴ in nenazadnje samostojnega klasificiranja spletnih vsebin v folksonomijah. Po drugi strani pa splet 2.0 in specifično folksonomije temeljijo na **povezovanju, soustvarjanju in pretoku informacij** med uporabniki. Prek asociativne narave označevanja vsebin se prilagajajo načinu delovanja naših možganov in dovoljujejo svoboden pretok misli ter idej oziroma kot to misel razvije Hanno (2005), tako premeteno tudi izkoriščajo posameznikova sebična dejanja in jih usmerjajo k skupnemu dobremu.

5.3.2 Folksonomije in pametne množice

Zaradi specifične dinamike delovanja folksonomij nekateri avtorji iščejo razlago za njih v modernejših socioloških konceptih, da bi lahko ustrezneje določili strukturo folksonomij. Poskus v tej smeri je tudi Sinhina (2006) razlaga folksonomij skozi Surowieckijev koncept **pametnih množic**. Pri opredeljevanju te na novo konstruirane entitete Surowiecki (2005) najprej odpravi negativno konotacijo množic, saj jo razume tudi kot skupino ljudi, ki se v osnovi ne zavedajo, da so del skupine, na primer vlagatelji na borzi. Množic torej ne tvorijo le nogometni navijači na stadionu ali protestniki na ulicah. Pametna množica pa je po Surowieckiju (2005) tista, ki zna konstantno reševati probleme bolje kot njen posamezni član, kar je predvsem posledica večje kolektivne agregacije znanja in manjše pristranskosti pri sprejemanju odločitev.

Pametne množice idejno sledijo tudi Rheingoldovim **nevidnim množicam** oziroma **pametnim združbam** (angl. smart mobs), v katerih so uporabniki sodobnih tehnologij sposobni delovati povezano v korist skupnega cilja, čeprav se še nikoli prej niso srečali in se med seboj ne poznajo. Ključno Rheingoldovo (2003) idejo agregacije kolektivne moči na podlagi horizontalnega

²³ S to dimenzijo se folksonomije približajo še enemu teoretičnemu konceptu, in sicer **skupnosti prakse** (community of practice), ki jo Wenger (v Kloos 2006) opredeli kot skupino ljudi, ki si delijo skrb, nabor problemov ali strast do določene teme in ki redno poglobljajo svoje znanje in strokovnost na tem področju. Ker pa se skupnosti prakse lahko izoblikujejo le v kontekstu učenja in na podlagi medsebojnega dogovora med člani skupnosti (Kloos 2006), večina javno dostopnih folksonomij teh pogojev ne izpolnjuje.

²⁴ RSS (*Really Simple Syndication*) je format oziroma metoda za organizacijo novic in drugih vsebin, dostopnih na spletu, ki so pripravljene za distribucijo uporabnikom. RSS je uporabnikom posredovan v obliki kratkega opisa vsebine spletne strani in povezave na celotno stran. (Hrastnik v Ban 2006).

povezovanja in sodelovanja posameznikov Surowiecki (2005) torej prenese na področje znanja; tudi povsem povprečna množica se lahko na podlagi skupno zbranih informacij odloča bolje kot najpametnejši posameznik v njej. Če gre pri nevidnih množicah za nov način organiziranja kolektivnega delovanja²⁵, lahko folksonomije razumemo kot nov način kolektivnega organiziranja informacij. Skupno obema konceptoma je pri tem umikanje vertikalnega načina komunikacije (na primer množičnih medijsko posredovanih informacij) ali vertikalnega (taksonomičnega) organiziranja informacij v korist horizontalnemu, kjer so vsi udeleženci komunikacijskega procesa izhodiščno med seboj enakovredni.²⁶

Sinha (2006) tako ugotavlja, da skupnost v omrežju folksonomij deluje po štirih principih delovanja pametnih množic, kot jih za predpogoj prisotnosti pametne množice navede Surowiecki (2005):

- raznolikost mnenj**, ki se med drugim kaže v porazdelitvi oznak v obliki dolgega repa;
- decentraliziranost**: taksonomija je izrazito centralističen sistem upravljanja informacij, folksonomija pa s svobodnim kalsificiranjem njeno čisto nasprotje;
- medsebojna neodvisnost članov množice**: vse, kar povezuje uporabnike klasičnih folksonomij, je zavedanje o prisotnosti drugih;
- združevanje mnenj**: slednje je v označevalnih sistemih vidno v obliki oblaka oznak ali združevanju oznak v kategorije. Sicer to nista najbolj optimalni metodi za agregacijo mnenj, kar ugotavlja tudi Sinha (2006).

Uporabniki spleta so z družbenim označevanjem osmislili informacije in brez velikih stroškov v kratkem času družno pravzaprav na novo katalogizirali splet. Toda kljub temu se zdi, da je družbeno označevanje nekako obstalo v senci drugega velikega kolektivnega projekta znanja, spletne enciklopedije Wikipedia. Ingram (2008) se na primer odkrito sprašuje, kdo sploh še označuje, saj ugotavlja, da do večine od svojih 10.000 spletnih zaznamkov, shranjenih v sistemu Delicious, potem, ko jih shrani, nikoli več ne dostopi. Folksonomije, kot najbolj prepoznaven stranski produkt družbenega označevanja, so prišle do stopnje, ko se zdi, da njihov nadaljnji

²⁵ Primer decentralizirano organizirane množice anonimnih posameznikov na spletu so uporabniki bit-torrent tehnologije, ki so trn v peti zlasti medijski industriji, saj si prek hitro prilegajočih se spletnih omrežij brezplačno in neovirano izmenjujejo medijske vsebine.

²⁶ Zanimiv remiks obeh konceptov v nespletnem okolju je kolektivna družbenokritična akcija samoimenovane *folksonomične združbe* (folksonomic mob), ki je hodila v knjigarne in pod pretvezo novoustanovljenega Ministrstva za premeščanje (Ministry of Reshelving) premikala Orwellovo knjigo 1984. Knjigo so premeščali iz tradicionalnih kategorij 'fikcija' in 'literatura' v nove; na primer v 'trenutna dogajanja', 'politika', 'zgodovina' in 'resnični zločin'. Za seboj je združba puščala obvestila, da je bila knjiga 1984 premeščena in da je Ministrstvo predano pravilni klasifikaciji fikcijskih in stvarnih besedil. (McGonigal 2005).

razvoj ni več možen. Lahko temu rečemo tudi kriza identitete? Družbeno označevanje je nastalo iz potrebe posameznika, da lažje shranjuje in dostopa do informacij, ki jih najde na spletu. Toda uporabnikova korist se s (pre)pogostim označevanjem predmetov premo sorazmerno manjša, saj je strošek iskanja pravega vira med vedno večjim številom lastnih oznak in zaznamkov prav tako vedno večji. Koristi imajo tako predvsem drugi uporabniki, saj so folksonomije lahko uporabljene kot odlična alternativa spletnih iskalnikov. Za razliko od Googla, ki v določenem trenutku ponuja najbolj relevantne zadetke, Delicious namreč omogoča tudi spremljanje najbolj svežih in priljubljenih spletnih virov, ki jih vsaj na prvi strani vrnjenih zadetkov na Googlu še ni. Individualna korist označevanja v folksonomijah je na videz torej zapostavljena na račun družbene, kar sicer ni slabo in je v skladu z naslovom tega podpoglavja, vendar pa se s tem odpirajo vprašanja, kako, kdaj in zakaj jih uporabljati.

Z informacijami ter znanjem se danes tudi na osebni ravni ukvarjamo bolj kot kadarkoli prej v zgodovini človeštva. Neodvisno od urednikov urejamo enciklopedije, ustvarjamo svoje bibliografske sezname in pišemo svoje dnevnike, ki ji lahko bere globalno občinstvo. Po eni strani smo torej priče izjemni in nepredstavljeni demokratizaciji informacij, po drugi strani pa smo pri iskanju specifičnega podatka še vedno bolj ali manj obsojeni na spletne iskalnike in na mučno sestavljanje celotne slike po koščkih. Imamo ogromno število razdrobljenih informacij, za katere se zdi, da smo jih s folksonomijami razdrobili še bolj.

Splet 2.0 je sprostil proces ustvarjanja informacij, semantični splet pa jih bo moral povezati, da bodo postale uporabne. Zdi se, da bodo izjemno pomembno vlogo v njem imele prav semantične folksonomije in na novo konceptualizirano družbeno označevanje.

6 Sklep

Družbeno označevanje in folksonomije v celoti sledijo filozofiji spleta 2.0, kjer vsebino ustvarjajo uporabniki. Že zdaj pa je moč slutiti, da bo vsaj ideja, če ne že sam način označevanja, tudi pomemben del naslednje generacije spleta. Povzemimo na tem mestu nekatere ključne ugotovitve naloge.

Osebni pristop k organizaciji informacij

Označevanje informacij zahteva določen miselni napor, poznavanje tematike, pogosto tudi zavzemanje stališča. Če parafraziramo Roszakovo (1994, 12) misel o izrinjenosti človeka s polja razumevanja informacij, se s personalizacijo informacij v obliki folksonomij človek tja ponovno vrača. Iz množice nepovezanih, razdrobljenih informacij izbira zanj relevantne in jim pripisuje pomen. Gre ponovno za kvalitativni pristop k obravnavi informacij, pri čemer z dobrim in smiselnim družbenim označevanjem uporabnik pomaga tudi drugim pri iskanju relevantnih informacij. Težko bi si tudi zamislili lepši primer Engelbartove vizije (Bardini v Oblak in Petrič 2005) spodbujanja intelektualne dejavnosti in medčloveškega sodelovanja s pomočjo računalniške tehnologije.

Digitalni red je zmešnjava

Uporabna informacija je le taka, ki jo lahko najdemo, prepoznamo ali povežemo z drugimi informacijami. Metapodatki, podatki o podatkih, so zato integralni del informacije. Brez njih bi od informacije ostala le forma, saj za svoj obstoj poleg generatorja nujno potrebujejo tudi prejemnika. Oznake v folksonomijah so uporabniško ustvarjeni metapodatki spletnih informacij in znanilci novega digitalnega reda, ki pravzaprav temelji na neredu (Weinberger 2007, 192). Zmešnjava digitalnega reda obrača logiko v nedigitalnem okolju prevladujočih taksonomij; bolj kot je neurejeno prepletena, bolj smiselna postaja. Z dodajanjem metapodatkov informacijam, nastajajo med njimi namreč nova in nova razmerja, ki jih pojasnjujejo in z realizacijo divje hipertekstovnosti tudi povezujejo. Eden izmed stranskih učinkov digitalnega (ne)reda je tako tudi popolna razpršenost distribucije spletnih vsebin; vhodnih točk do iskanega vira je namreč vsaj toliko, kot je tudi oznak za ta vir, kar poleg lažjega dostopa, kjer je vsebina pomembnejša od lokacije, omogoča uporabnikom tudi participacijo pri ustvarjanju pomena informacij.

Informacij je na spletu preveč, da bi jih lahko kategorizirali po principu taksonomij, poleg tega ekspertni način razvrščanja informacij za tako dinamično okolje ni primeren. Pri tem velja

poudariti, da to ni kritika taksonomičnega sistema, noben način organiziranja informacij namreč ni sam po sebi najboljši, ampak je izbira odvisna tako od narave informacij, kot tudi od tipa uporabnikov ali okolja, v katerem se informacije nahajajo. Odlične rezultate pri tem dajejo zlasti sistemi, ki kombinirajo različne principe organiziranja informacij, saj lahko v polni meri izkoristijo dobre lastnosti posameznega sistema in hkrati z njegovo dopolnitvijo nadoknadijo tudi nekatere slabosti. Tipični primer so denimo označevalne ontologije. Po eni strani dopuščajo svobodno interpretacijo vsebine označenega vira, a omejujejo njegovo definicijo in tako odpravljajo ključno slabost folksonomij, dvoumnost, po drugi strani pa ustvarjajo ontologije posameznih področij, ki sicer niso najbolj profesionalne, a se zato v primerjavi s klasičnimi ontologijami ustvarjajo skoraj s svetlobno hitrostjo.

Semantični splet

Družbene oznake so verjetno najbolj primitivna oblika klasificiranja podatkov, ki za razliko od denimo bolj napredne kategorizacije vsebin po mapah, od uporabnika ne zahtevajo nikakršnega predznanja. Na povsem drugem koncu spektra po zahtevnosti klasificiranja podatkov pa so ontologije, ki zahtevajo natančne opise spletnih vsebin in uporabo dogovorjenih formalizmov za zapis znanja, z namenom, da ji računalniki lahko razumejo. Prihodnost semantičnega spleta vidimo prek označevalnih ontologij v kombiniranju obeh sistemov, pri čemer bodo uporabniki s svojimi oznakami najbrž skrbeli tako za opis spletnih vsebin kot tudi za definicijo teh oznak. Eden izmed najbolj logičnih virov za črpanje definicij je Wikipedia, ki deluje po principu sodelovalnega filtriranja. Definicije v njej ne bodo nikoli končne oziroma ontološko čiste, v najboljšem primeru bodo kvečjemu izražale nek globalni konsenz okrog določene vsebine. To seveda ni Berners-Leejev semantični splet z veliko začetnico (Semantic Web), ki se zaradi časovno zamudne in kompleksne formalizacije opisovanja vsebin ter skoraj utopične zahteve po eni univerzalni Definiciji, zdi vse bolj v prihodnost odmaknjen projekt. Z veliko večjo verjetnostjo lahko govorimo o skorajšnjem prihodu semantičnega spleta z malo začetnico (semantic web), ki bo s svojo neuniverzalnostjo in parcialnostjo marsikoga razočaral, za velik del spletne skupnosti pa bo najbrž, podobno kot družbeno označevanje, popoln v svoji nepopolnosti.

Literatura

Amazon. Dostopno prek: <http://www.amazon.com/> (8. september 2008).

Anderson, Chris. 2004. *The Long Tail*. Dostopno prek: <http://changethis.com/10.LongTail> (14. junij 2008).

Angeletou, Sofia, Marta Sabou, Lucia Specia in Enrico Motta. 2007. *Bridging the Gap Between Folksonomies and the Semantic Web: An Experience Report*. Dostopno prek: <http://kmi.open.ac.uk/people/marta/papers/semnet2007.pdf> (4. januar 2009).

Ban, Sandra. 2006. *Elektronsko trženjsko komuniciranje s poudarkom na spletnih dnevnikih in RSS (really simple syndication)*. Magistrsko delo. Ljubljana: Ekonomska fakulteta.

Berners-Lee, Tim, James Hendler in Ora Lassila. 2001. The Semantic Web. *Scientific American* (maj): 29-37. Dostopno prek: <http://www.sciam.com/article.cfm?id=the-semantic-web> (4. januar 2009).

Brejc, Igor in Marjan Heričko. 2005. Uporaba ontologij pri opisovanju in svetovanju o načrtovalskih vzorcih. *COTL* 11 (2). Dostopno prek: <http://cot.uni-mb.si/cotl/jesen2005/brejc.html> (4. januar 2009).

Brynjolfsson, Erik, Yu Jeffrey Hu in Michael D. Smith. 2006. From Niches to Riches: The Anatomy of the Long Tail. *Sloan Management Review* 47 (4): 67-71. Dostopno prek: <http://www.heinz.cmu.edu/~mds/smr.pdf> (20. april 2008).

Budimir, Gordana. 2006. Podatkovne XML-scheme v knjižničnih informacijskih sistemih. *Organizacija znanja* 11 (1): 2-18.

Cambridge Dictionary. Dostopno prek: <http://dictionary.cambridge.org/> (23. februar 2009).

Castells, Manuel. 1996. *The Rise of the Network Society*. Cambridge, Oxford: Blackwell.

Citizen Watch. 2009. *Citizen watches online*. Dostopno prek: <http://www.citizenwatch.com/>

COA/English/browse.asp (5. marec 2009).

Clark, Jeff. 2009. *TED Shaped Word Cloud*. Dostopno prek: <http://www.neoformix.com/2009/TEDShapedWordCloud.html> (14. maj 2009).

Delicious. Dostopno prek: <http://delicious.com/> (14. januar 2009).

Delicious. 2008. *Everyone's Bookmarks for: Wired 12.10: The Long Tail*. Dostopno prek: <http://delicious.com/url/1f4a4ea728c7e9ff20eaeda1dd1cc7c8> (28. november 2008).

developerWorks Interviews. 2006. *Tim Berners-Lee Originator of the Web and director of the World Wide Web Consortium talks about where we've come, and about the challenges and opportunities ahead*. Dostopno prek: <http://www.ibm.com/developerworks/podcast/dwi/cm-int082206txt.html> (13. maj 2008).

Dmoz. Dostopno prek: <http://www.dmoz.org/> (17. april 2009).

Dreyfus, Hubert. 2001. *On the Internet*. London; New York: Routledge.

Elliance. 2006. *Tagging explained*. Dostopno prek: <http://searchengineoptimization.elliance.com/pdfs/Tagging-Explained.pdf> (28. maj 2008).

Euzenat, Jérôme. 2001. *Report from the NSF-EU Workshop, Research Challenges and Perspectives of the Semantic Web*. Dostopno prek: <http://www.ercim.org/EU-NSF/semweb.html> (4. maj 2008).

Fay, Rob. 2006. *Egocentrically Altruistic Web 2.0*. Dostopno prek: <http://www.robfoy.com/2006/01/18/egocentrically-altruistic-web20/> (16. september 2008).

Golder, Scott in Bernardo A. Huberman. 2005. *The Structure of Collaborative tagging Systems*. Dostopno prek: www.hpl.hp.com/research/idl/papers/oznakas/tags.pdf (2. maj 2007).

Google. 2009. *Technology Overview*. Dostopno prek: <http://www.google.com/corporate/tech.html> (28. april 2009).

Green, Heather in Robert D. Hof. 2005. Picking Up Where Search Leaves Off. The time-saving trend of "tagging" is luring legions of Web surfers -- and Yahoo! *BusinessWeek*, 11. april. Dostopno prek: http://www.businessweek.com/magazine/content/05_15/b3928112_mz063.htm (30. maj 2008).

Gruber, Tom. 2007. Ontology of Folksonomy: A Mash-up of Apples and Oranges. *Int'l Journal on Semantic Web & Information Systems* 3 (1): 1-11. Dostopno prek: <http://tomgruber.org/writing/ontology-of-folksonomy.htm> (23. februar 2009).

Guy, Marieke in Emma Tonkin. 2006. Folksonomies: Tidying up Tags? *D-Lib Magazine* 12 (1). Dostopno prek: <http://www.dlib.org/dlib/january06/guy/01guy.html> (2. maj 2007).

Hammond, Tony, Timmo Hannay, Ben Lund in Joanna Scott. 2005. Social Bookmarking Tools (I): A General Review. *D-Lib Magazine* 11 (4). Dostopno prek: <http://www.dlib.org/dlib/april05/hammond/04hammond.html> (23. maj 2008).

Hannay, Timmo. 2005. *Introduction: Timo Hannay*. Dostopno prek: <http://tagsonomy.com/index.php/introduction-timo-hannay/> (14. junij 2008).

Hassan-Montero, Yusef in Victor Herrero-Solana. 2006. *Improving Tag-Clouds as Visual Information Retrieval Interfaces*. Dostopno prek: http://www.nosolousabilidad.com/hassan/improving_tagclouds.pdf (3. maj 2009).

Herlocker Jonathan L., A. Joseph Konstan, G. Loren Terveen in John T. Riedl. 2004. Evaluating Collaborative Filtering Recommender Systems. *ACM Trans. Inf. Syst* 22 (1): 5-53. Dostopno prek: http://web.engr.oregonstate.edu/~herlock/papers/eval_tois.pdf (8. september 2008).

Hodge, Gail. 2001. *Metadata Made Simpler*. Dostopno prek: http://www.niso.org/news/Metadata_simpler.pdf (10. september 2007).

Horrigan, B. John. 2006. *The Internet as a Resource for News and Information about Science*. Dostopno prek: http://www.pewinternet.org/~media/Files/Reports/2006/PIP_Exploratorium_Science.pdf (4. maj 2009).

Ingram, Mathew. 2008. *Delicious 2.0: Who bookmarks any more?* Dostopno prek: <http://www>.

mathewingram.com/work/2008/07/31/delicious-20-who-bookmarks-any-more/ (1. maj 2009).

Johnson, M. Christopher. 2001. A survey of current research on online communities of practice. *The Internet and Higher Education* 4 (1): 45–60. Dostopno prek: <http://www.learnloop.org/olc/johnsonOnlineCoP.pdf> (21. junij 2007).

Kavčič, Alenka. 2007. Učinkovita uporaba obstoječih gradiv za e-izobraževanje. V *Mednarodna konferenca Splet izobraževanja in raziskovanja z IKT – SIRIKT*, ur. Mojca Orel, 355-359. Ljubljana: Arnes.

Kelly, Kevin. 2005. We Are the Web. *Wired Magazine* 8 (13). Dostopno prek: <http://www.wired.com/wired/archive/13.08/tech.html> (18. junij 2007).

Kloos, Martin. 2006. *Communities of practice 2.0. How blogs, wikis, and social bookmarking offer facilities that support learning in practice in communities of practice*. Dostopno prek: <http://www.martinkloos.nl/thesis-M.Kloos.pdf> (24. februar 2009).

Koutrika, Georgia, Frans A. Effendi, F.A., Zoltan Gyöngyi, Paul Heymann in Hector Garcia-Molina. 2007. *Combating Spam in Tagging Systems*. Dostopno prek: http://airweb.cse.lehigh.edu/2007/papers/paper_97.pdf (4. januar 2009).

Landow, George. 2006. *Hypertext 3.0: Critical Theory and New Media in an Era of Globalization*. Baltimore, London: The Johns Hopkins University Press.

Lavbič, Dejan, Marko Bajec in Marjan Krisper. 2006. Pravila na semantičnem spletu. *Elektrotehniški vestnik* 73 (5): 249-254. Ljubljana.

Locke, Matt. 2005. *A taxonomy of humour - what nurses can teach us about classification*. Dostopno prek: <http://test.org.uk/2005/02/25/a-taxonomy-of-humour-what-nurses-can-teach-us-about-classification/> (25. maj 2008).

MacManus, Richard. 2005. *Web 2.0 is dead. R.I.P.* Dostopno prek: http://www.readwriteweb.com/archives/web_20_is_dead.php (17. september 2007).

Many Eyes. Dostopno prek: <http://manyeyes.alphaworks.ibm.com> (16. januar 2009).

Marlow, Cameron, Mor Naaman, Danah Boyd in Marc Davis. 2006. *Position Paper, Tagging, Taxonomy, Flickr, Article, ToRead*. Dostopno prek: <http://www.danah.org/papers/WWW2006.pdf> (6. maj 2007).

Mathes, Adam. 2004. *Folksonomies—Cooperative Classification and Communication Through Shared Metadata*. Dostopno prek: <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.pdf> (21. maj 2008).

McGonigal, Jane. 2005. *The ministry of reshelving*. Dostopno prek: <http://blog.avantgame.com/2005/08/ministry-of-reshelving.html> (17. marec 2007).

Mika, Peter. 2007. *Social tagging and ontologies will co-exist*. Dostopno prek: www.semantic-web.at/1.36.resource.88.peter-mika-x22-social-tagging-and-ontologies-will-co-exist-x22.htm (6. maj 2008).

Milgram, Stanley. 1967. The small world problem. *Psychology Today* 67 (1): 61–67.

Millard, David E. in Martin Ross. 2006. *Web 2.0: Hypertext By Any Other Name?* Dostopno prek: <http://eprints.ecs.soton.ac.uk/13085/01/Web2-short-final.pdf> (6. september 2007).

Oblak, Tanja. 2000. Mitske podobe o "Življenju na mreži": od poljubnih identitet do izgubljenih skupnosti v kibernetnem prostoru. *Teorija in praksa* 37 (6): 1052-1068.

Oblak, Tanja in Gregor Petrič. 2005. *Splet kot medij in mediji na spletu*. Ljubljana: Fakulteta za družbene vede.

OCLC, Online Computer Library Center. 2003. *Summaries DDC, Dewey Decimal Classification*. Dostopno prek: <http://www.oclc.org/dewey/resources/summaries/deweysummaries.pdf> (23. junij 2007).

O'Reilly, Tim. 2004. *Web 2.0 Conference*. Dostopno prek: <http://conferences.oreillynet.com/> (2. marec 2009).

--- 2006. *Web 2.0 Compact Definition: Trying Again*. Dostopno prek: <http://radar.oreilly.com/>

archives/2006/12/web_20_compact.html (17. september 2007).

--- 2007. *Today's Web 3.0 Nonsense Blogstorm*. Dostopno prek: <http://radar.oreilly.com/archives/2007/10/web-30-semantic-web-web-20.html> (5. marec 2009).

Petrič, Gregor. 2003. *Družbeno delovanje v omrežju svetovnega spleta: individualni in strukturni vidiki*. Doktorska disertacija. Ljubljana: Fakulteta za družbene vede.

Pirc, Tomazs. 2007. *SOF in oglaševanje 2.0*. Dostopno prek: <http://www.pirc.cc/dnevnik/sof-20/> (7. maj 2008).

Quintarelli, Emanuele. 2005. *Folksonomies: power to the people*. Dostopno prek: <http://www.dimat.unipv.it/biblio/isko/doc/folksonomies.htm> (14. junij 2007).

Reitz, Joan. 2004. *Dictionary for Library and Information Science*. Dostopno prek: <http://lu.com/odlis/> (23. februar 2009).

Rheingold, Howard. 2003. *Smart mobs: the next social revolution*. Cambridge: Perseus Publishing.

Roszak, Theodore. 1994. *The Cult of Information: a neo-Luddite treatise on high tech, artificial intelligence, and the true art of thinking*. Berkeley; London: University of California Press.

Russell, Terrell. 2005. *Contextual Authority Tagging: Cognitive Authority through Folksonomy*. Dostopno prek: <http://www.terrellrussell.com/projects/contextualauthoritytagging/conauthtag200505.pdf> (22. junij 2007).

Schmitz, Christoph, Miranda Grahl, Andreas Hotho, Gerd Stumme, Ciro Cattuto, Andreas Baldassari, Vitorio Loretto in Vito D.P. Servedio. 2007. *Network Properties of Folksonomies AI Communications*. Dostopno prek: http://www2007.org/workshops/paper_13.pdf (4. januar 2009).

Shaw, Russell. 2005. *Web 2.0? It doesn't exist*. Dostopno prek: <http://blogs.zdnet.com/iptelephony/?p=805> (17. september 2007).

Shirky, Clay. 2005. *Ontology is Overrated: Categories, Links, and Tags*. Dostopno prek:

http://www.shirky.com/writings/ontology_outrated.html (14. junij 2007).

Sinha, Rashmi. 2006. *A cognitive analysis of tagging*. Dostopno prek: http://www.rashmisinha.com/archives/05_09/tagging-cognitive.html (5. maj 2007).

Spivack, Nova. 2005. *Folktologies-Beyond the folksonomy vs. ontology distinction*. Dostopno prek: http://novaspivack.typepad.com/nova_spivacks_weblog/2005/01/whats_after_fol.html (23. februar 2009).

--- 2007. *Harnessing the collective intelligence of the World-Wide Web*. Dostopno prek: http://novaspivack.typepad.com/nova_spivacks_weblog/2008/04/great-collectiv.html (23. maj 2008).

Surowiecki, James. 2005. *The wisdom of crowds: why the many are smarter than the few*. London: Abacus.

Tauberer, Joshua. 2008. *What is RDF and what is it good for?* Dostopno pre: <http://www.rdfabout.com/intro/> (23. maj 2008).

Trček, Franc. 2003. *Problem informacijske (ne)dostopnosti*. Ljubljana: Center za prostorsko sociologijo, Fakulteta za družbene vede.

Twine. 2008. *About Twine*. Dostopno prek: <http://www.twine.com/about> (4. december 2008).

Twine. Dostopno prek: <http://www.twine.com/> (4. januar 2009).

Voss, Jakob. 2007. *Tagging, Folksonomy & Co - Renaissance of Manual Indexing*. Dostopno prek: <http://arxiv.org/abs/cs/0701072> (5. maj 2007).

Wal, Vander. 2005. *Explaining and Showing Broad and Narrow Folksonomies*. Dostopno prek: <http://www.vanderwal.net/random/entrysel.php?blog=1635> (7. september 2008).

--- 2007. *Folksonomy Coinage and Definition*. Dostopno prek: <http://vanderwal.net/folksonomy.html> (2. marec 2009).

Walker, Jill. 2005. *Feral Hypertext: When Hypertext Literature Escapes Control*. Dostopno

prek: <http://jilltxt.net/txt/FeralHypertext.pdf> (8. julij 2008).

Weinberger, David. 2007. *Everything is miscellaneous: the power of the new digital disorder*. New York: Time books.

Weinroth, Adam. 2005. *Tag Team*. Dostopno prek: <http://adam.easyjournal.com/entry.aspxeid=2632426> (2. maj 2007).

Wesch, Michael. 2007. *The Machine is Us/ing Us*. Dostopno prek: <http://mediatedcultures.net/mediatedculture.htm> (2. marec 2009).

Wikipedia. 2007. *Web 2.0*. Dostopno prek: http://en.wikipedia.org/wiki/Web_2.0 (10. oktober 2007).

Winer, Dave. 2005. *Busted*. Dostopno prek: <http://www.scripting.com/2005/12/19.html#busted> (17. junij 2007).

WorldExplorer. Dostopno prek: <http://tagmaps.research.yahoo.com/worldexplorer.php> (28. marec 2008).

Zigtag. Dostopno prek: <http://www.zigtag.com/> (4. januar 2009).