Prikupljanje metadata podataka - udar na privatnost

Kustec, Jurica

Undergraduate thesis / Završni rad

2017

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Pula / Sveučilište Jurja Dobrile u Puli**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:137:726383

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**

Image not found or type unknown

Repository / Repozitorij:

Digital Repository Juraj Dobrila University of Pula

zir.nsk.hr

Image not found or type unknown

Sveučilište Jurja Dobrile u Puli

Odjel za komunikacijske – informacijske tehnologije

JURICA KUSTEC

PRIKUPLJANJE METADATA PODATAKA: UDAR NA PRIVATNOST

Pula, 2017.

Sveučilište Jurja Dobrile u Puli
Odjel za komunikacijske – informacijske tehnologije

JURICA KUSTEC

PRIKUPLJANJE METADATA PODATAKA

Završni rad

JMBAG: 0145015596, izvanredni student

Studijski smjer: Informatika

Predmet: Ekonomika informacijskog sustava

Mentor: doc. dr. sc. Ivan Pogarčić

Pula, rujan, 2017.

IZJAVA O AKADEMSKOJ ČESTITOSTI

Ja, dolje potpisani _____, kandidat za prvostupnika _____ovime izjavljujem da je ovaj Završni rad rezultat isključivo mojega vlastitog rada, da se temelji na mojim istraživanjima te da se oslanja na objavljenu literaturu kao što to pokazuju korištene bilješke i bibliografija. Izjavljujem da niti jedan dio Završnog rada nije napisan na nedozvoljen način, odnosno da je prepisan iz kojega necitiranog rada, te da ikoji dio rada krši bilo čija autorska prava. Izjavljujem, također, da nijedan dio rada nije iskorišten za koji drugi rad pri bilo kojoj drugoj visokoškolskoj, znanstvenoj ili radnoj ustanovi.

Student:

U Puli, 27. 10. 2017. _____

SADRŽAJ

1. UVOD ……..……...……1
2. OSNOVE META PODTAKA …...…….……..2
   2.1. Implementacija meta podataka ……3
   2.2. Statistički meta podaci ……4
   2.3. Registri meta podataka …… 4
3. KATEGORIZACIJA META PODATAKA .…… 5
   3.1. Opisni ……7
   3.2. Strukturalni …… 7
   3.3. Administrativni ……..... 8
   3.4. Atributi i značajke meta podataka ……..... 8
   3.5. Ostale značajke ……... 9
4. PRIMARNE FUNKCIJE META PODATAKA …..……..... 10
   4.1. Stvaranje, višestruko preusmjeravanje,
   ponovna uporaba, rekontekstualizacija informacijskog objekta ……... 10
   4.2. Organizacija i opis ……..... 10
   4.3. Provjera valjanosti ……... 11
   4.4. Pretraživanje i pronalaženje ……. 11
   4.5. Korištenje i očuvanje ……... 11
   4.6. Raspored …… 11
5. VAŽNOST META PODATAKA……..... 12
   5.1. Web pretraživači…….. 13
   5.2. Upotreba META tagova ……..…… 14
   5.3. Title tagovi ……..... 15
   5.4. Dark „web“ ……... 18
6. UDAR NA PRIVATNOST ……... 19
   6.1. Meta podaci i „Facebook“…… 20
   6.2. Primjeri zlouporabe podataka preko IP adrese …….. 22

6.3. Primjer zlouporabe podataka putem e-mail-a …..……22
6.4. Slabosti društvenih stranica …… 23
6.5. Zaštita podataka …… 23
6.6. Europska komisija …….... 24
7. ZAKLJUČAK ……...…… 26
8. LITERATURA ..……... 27
9. POPIS SLIKA …….…… 29
10. SAŽETAK ..…… 30
11. SUMMARY ..……... 31

1. UVOD

Danas je pretraživanje informacija odnosno postala uobičajena pojava za svaku starosnu dob. Stoga je pretraživanje sadržaja nezamislivo bez neke vrste meta podatka (eng. metadata), zato je ova tema vrlo važna za sve dobne skupine koje se koriste podacima na internetu.

Rad je podijeljen na sedam cjelina kroz koje će se objasniti što su to meta podaci, odnosno definicije, podjele, glavne specifikacije i funkcije te objasniti kako meta podaci utječu na sigurnost informacija vezanih za korisničke podatke te kako te podatke zaštiti od ranjivosti koje im prijete.

Kroz rad je opširnije opisano na koji način pretraživači upisane riječi ili slogove povezuju sa pretraženim informacijama. Za rad s meta podacima potrebno je poznavati njihovu funkcionalnost i poznavati način njihovog pravilnog korištenja kako bi se pozitivni učinci meta podataka optimizirali, a negativni izbjegli i da li postoji mogućnost zaštite od negativnih udara na korisničke podatke unutar baza podataka.

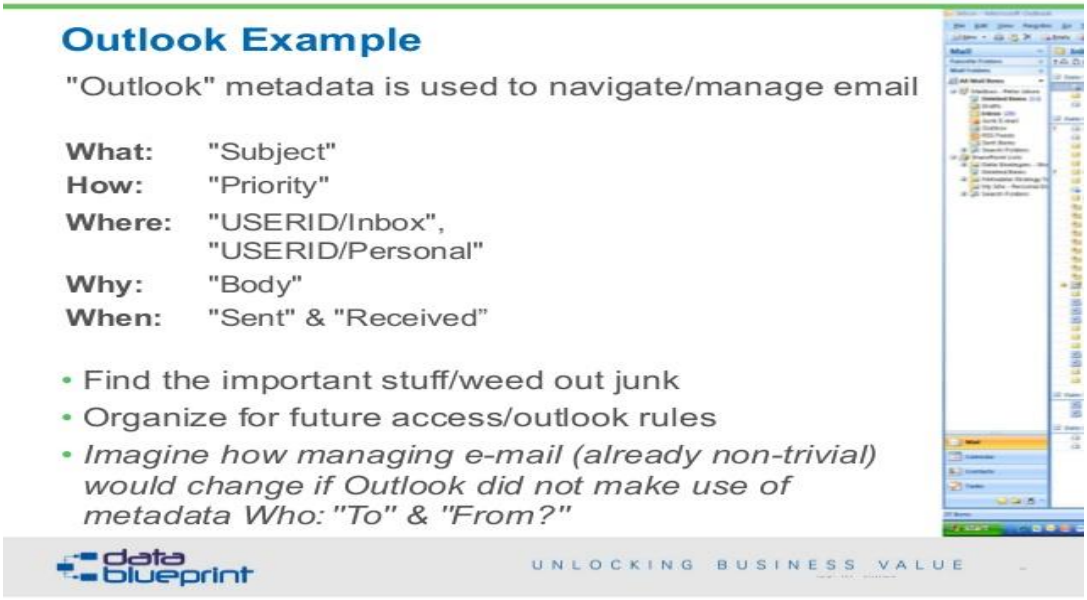
Cilj tematike je predstaviti meta podatke, razmotriti njihovu ulogu u svakodnevnom pretraživanju informacija i resursa na web stranicama i pretraživačima te prikazati kakav imaju utjecaj na sigurnost i privatnost korisnika te kako ih zaobići.

2. OSNOVE META PODATAKA

Meta podaci opisuju jedan podatak, cijelu skupinu podataka ili dio cjeline podataka odnosno oni su strukturirane informacije koje opisuju, objašnjavaju, smještaju ili na bilo koje druge načine lakše pronalaze, koriste ili upravljaju informacijskim resursima. U pravilu postoji niz raznih definicija, a najjednostavnija definicija je da su to informacije o podacima ili podaci o podacima.

Glavno obilježje meta podataka jest HTML (eng. HyperText Markup Language) dokument i zaglavlja slikovnih datoteka gdje se ugrađuju (Baca, M., 2008.), primjer je e-mail (slika 1.) čiji su meta podaci vrijeme i datum kada je poslan-primljen, naslov teksta, prioritet te primatelj i pošiljatelj mail-a.

Slika 1. – primjer meta podatka u e-mailu



https://www.slideshare.net/aberkowitz/2014-metadata-management-strategieskey-41655167

Meta podaci su iznimno važni kako bi se osigurali resursi za buduću upotrebu. Oni zamjenjuju stari način pohrane podataka pa se u budućnosti sve manje prakticira pohrana spisa i ostalih dokumenata u arhive odnosno pohranu u podrumska skloništa. Također služe u funkciji pronalaska resursa po relevantnim kriterijima, utvrđivanjem resursa, spajanjem

sličnih resursa i isto tako prepoznavanjem različitih resursa te otkrivanju lokacija za te resurse.

Zajednica može tumačiti meta podatke na razne načine pa tako postoje podaci koji služe za pohranu u bazu podataka i zapise koji opisuju elektroničke resurse. U bazama podataka se koriste kao bilo koja formalna shema opisa resursa. Tradicionalno katalogiziranje u bazama podataka je jedan od oblika meta podataka kao što je skup pravila odnosno shema MARC 21 (Baca, M., 2008.). Sve druge sheme meta podataka razvijene su kako bi opisale različite vrste tekstualnih i ne tekstualnih predmeta kao što su skupovi elektroničkih podataka, objavljene knjige, edukacijske i obrazovne dokumente, umjetničke predmete itd.

2.1. Implementacija meta podataka

Početci meta podataka vuču se od ranih '90-tih godina prošlog stoljeća gdje su poslovni menadžeri uvidjeli vrijednosti informacijskih objekata i podataka. Pa su tako počeli razvijati tada novije alate za njihov razvoj.

Potencijalne koristi tog perioda, uključivali su semantički sloj između poslovnih partnera i veliki sustav tvrtki, smanjivali ukupne troškove, stvaranje novih strategija i djelotvornih informacija te smanjenja neispravnih odluka.

Srednjih godina '90-tih pa na samom kraju '90-tih godina dolazi do spajanja podataka velikih baza podataka sa fokusom na širenje world wide web-a. U tom vremenu započeta je standardizacija definicije metadata podataka kao što su CDIF (Case Definition Interchange Facility), Dublin Core Metadata Elements, ISO 11179 standardi za specifikaciju i standardizaciju meta podataka (https://www.slideshare.net/aberkowitz/2014-metadata-management-strategieskey od 17.09.2017.), zatim CWM (Common Warehouse MetaModel) koji je dio grupe koja je zadužena za objektno programiranje te se javljaju mnogi drugi standardi.

Dvadeset i prvo stoljeće dovodi do daljnjeg rasta upotrebe metadata podataka i njihovoj implementaciji na web stranice. Prodavači prodaju i nude usluge stvaranja meta podatkovnih struktura, pojavljuju se organizacije koje otkupljuju ili stvaraju podatkovne baze. Danas se uglavnom ulažu napori kako implementirati meta podatke izvan tradicionalno strukturiranih izvora baza podataka.

2.2. Statistički meta podaci

Statistički podaci su oni koji su prikupljeni kroz razna anketiranja. Takva vrsta podataka prikupljena i obrađena kroz anketiranje zove se još mikro podaci, makro podaci i vremenske serije. Statističkim meta podacima kroz dokumentiranje definiramo i opisujemo statističke podatke kroz njihov vremenski tijek.

2.3. Registri meta podataka

Baza podataka koja se koristi za pohranu, organiziranje, upravljanje i dijeljenje meta podataka naziva se registar meta podataka. Osmišljen je za rješavanje problema upravljanja meta podatcima iz perspektive organizacije, a ne samo na malim programskim područjima. Također i osigurava potrebne metadata podatke za opis interesnih predmeta te entiteta potrebitih za standardizaciju i registraciju.

3. KATEGORIZACIJA META PODATAKA

Kako bi što bolje razumjeli koncepte meta podataka, korisno ih je razdvojiti u različite kategorije: administrativne, opisne, uporabne, meta podaci za čuvanje, te strukturne koji se reflektiraju na ključne funkcionalne aspekte.
Meta podaci za čuvanje i uporabu odnose se na očuvanje i odnose pri upravljanja zbirkama i informacijskih resursa odnosno tekstualne predmete.

Danas meta podatkovni standardi odnosno sheme (Baca M., 2008.) koje se primjenjuju u pretraživanja od velikog su značaja jer se upravo preko njih dolazi do potrebnih informacija na osnovu pretraživanja baza podataka, a ti standardi jesu:

- Meta TAG,
- MARC,
- Dublin Core,
- CDWA,
- EAD,
- RDF (Resource Description Framework) i
- Bountiful Harvest.

Meta tag-ovi u bazama podataka su riječi ili fraze pridružene slogu baze, a služe za opis podataka, identifikaciju, sortiranje i pronalaženje. Popularni su u bazama podataka za javnost kao na primjer YouTube, Flickr, G-mail, gdje korisnici postavljaju vlastite tagove radi lakše identifikacije ili pronalaženja podataka.

MARC je standard za prikaz i prijenos bibliografskih i drugih odgovarajućih podataka, osigurava protokol preko kojeg računala razmjenjuju, koriste te implementiraju bibliografske informacije.

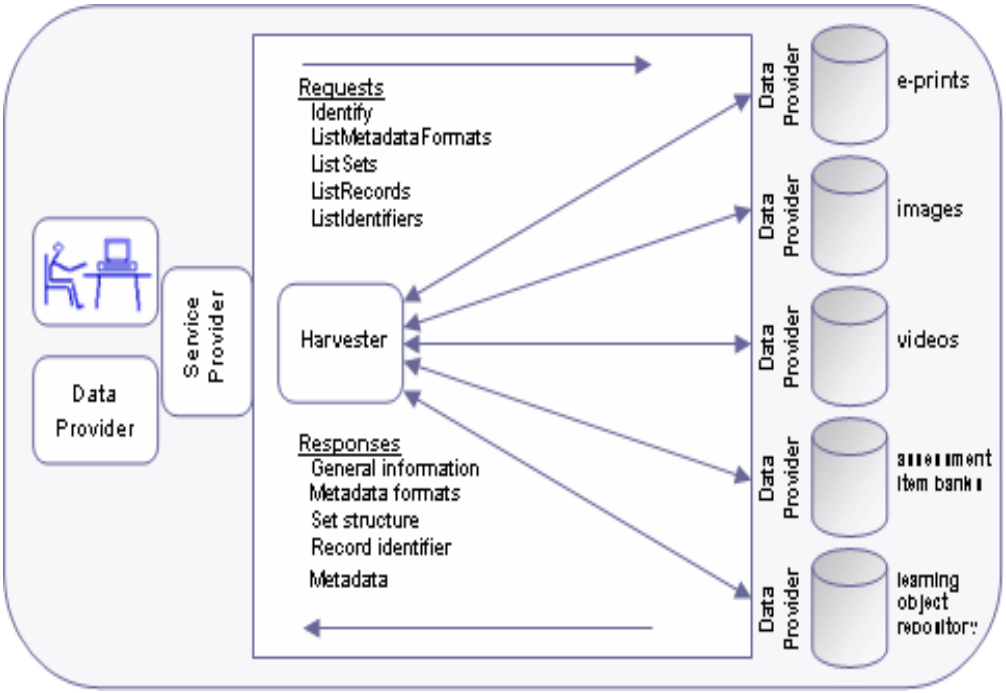
Meta podaci kao široko i generički iskoristivi podaci za opisivanje izvora iz različitih područja je ujedno i opis Dublin Core standarda koji je nastao iz Dublin Core Metadata Intiative organizacije koja se bavi izradom standarda i rječnika radi lakšeg upravljanja informacijama.

CDWA ili kategorija za opis umjetničkih djela, opisuju sadržaja baza podataka umjetnički djela artikulirajući konceptualni okvir za opisivanje i pristup informacijama o umjetničkim djelima, arhitekturama, drugim materijalnim kulturama te grupama i zbirkama djela te srodnim slikama.

Osnova semantičkog weba je jezik RDF (Resource Description Framework) koji je u počecima predstavljao jezik za prezentiranje informacija o web resursima. Te informacije su uključivale podatke o nastanku resursa, autoru, autorskom pravu i sličnim koje su se zapisivale u meta tagovima HTML stranica. RDF služi kao framework za izražavanje informacija o općenitim resursima koji mogu obrađivati različite aplikacije i nisu primarno namijenjene za direktan prikaz korisniku.

Bountiful Harvest je protokol za inicijativu otvorenoga arhiva za skupljanje meta podataka, pruža alternativnu metodu za izradu Deep web-a. Kod ove vrste standarda umjesto ugrađivanja meta podataka u stvarni sadržaj Web stranice, OAI-PMH (https://publishopenaccess.blogspot.hr/2010/11/open-archives-initiative-protocol.html) je skup jednostavnih protokola koji dopušta zapis meta podataka na predvidljiv način na webu tako da ih mogu dohvatiti i drugi računalni sustavi kompatibilni sa OAI-PMH protokolom. Na slici 2. može se vidjeti jedan način koji opisuje Bountiful Harvest protokol pri prikupljanju podataka od same strukturne baze podataka pa sve do krajnjeg korisnika.

Slika 2. – Metadata harvesting model



https://publishopenaccess.blogspot.hr/2010/11/open-archives-initiative-protocol.html

Po NISO-u (http://www.niso.org/publications/press/UnderstandingMetadata.pdf od 20.09.2017.]) postoje tri glavne kategorije meta podataka, a oni su: opisni, strukturalni i administrativni meta podaci.

3.1. Opisni

Opisni meta podaci opisuju izvor u svrhu pronalaska i identifikacije istog. Ovaj tip uključuje elemente poput autora, naslova, sažetka. Kao vrsta opisnih meta podataka često se spominju identifikacijski meta podaci.

3.2. Strukturalni

Strukturalni meta podaci ukazuju kako se sjedinjeni objekti spajaju, ukazuju na format podataka, način prezentacije dokumenta, hardware i software koji su potrebni za prikaz informacija.

3.3. Administrativni

Administrativni meta podaci sadrže informacije koje su potrebne za potpomaganje upravljanja dokumentom. To se odnosi na informacije o tome kada i kako je kreiran dokument, vrstu dokumenta, tko dokumentu može pristupiti, te druge tehničke informacije. U podtipove administrativnih meta podataka spadaju „rights management“ metadata – meta podaci relevantni za upravljanje pravima, koji se bave pravima intelektualnog vlasništva, te „preservation“ metadata – meta podaci vezani uz zaštitu, koji sadrže informacije potrebne kako bi se izvor arhivirao i očuvao.

3.4. Atributi i značajke meta podataka

Izvori internih meta podataka kreiraju se generiranjem agenta za stvaranje informacijskih objekata (nazivi datoteka, zatim informacija o zaglavlju, formatu datoteka te njihovim strukturama), zatim podataka ubačenih u objekte i djela (naslovi i podnaslovi u umjetničkim djelima te na naslovnim stranicama tiskanih knjiga) te vanjskih podataka koji su vezani za originalne stavke ili objektne podatke (URL-ove, registre i pravne informacije).

Metode kreiranja razlikuju se u tome da li ih je kreirao čovjek (računalni početnik, informatički profesionalac) ili su generirani putem računala (napredni alatni sustavi za stvaranje informacija o podacima).

Ostali atributi meta podataka su status, struktura, semantika te nivo. Status opisuje podatke kao statičke ili nepromjenjive, dinamičke koji se mogu modificirati , dugoročne koji opisuju neke pravne informacije i kratkoročne kao neke privremene informacije o lokaciji. Strukturni atributi bazirani su na standardizaciji te nestrukturirani odnosno slobodni podaci te napomene bez tekstualnih informacija.

Kod semantike razlikujemo kontrolirane meta podatke standardizirane pravilima te nestandardizirane odnosno nekontrolirane podatke, a nivo meta podataka opisuje jedan predmet informacijskog objekta ili grupe izvornih predmeta informacijskih objekata.

Važno je da se meta podaci povezuju s informacijskim objektom kroz dvosmjerne pokazivače ili hiperveze te da su njihovi odnosi dokumentirani u registrima meta podataka te je pri tom ključno da sadržaj i meta podatak koegzistiraju odnosno postanu sastavni dio objekta informacija to jest da je ugrađen u objekt.

Kako dolaze sve zahtjevnije strukture u modeliranju i dizajniranju baza podataka tako se razvija nužna evolucija i implementacija na sustavne promjene značajki meta podataka.

3.5. Ostale značajke

Meta podaci nisu nastali pojavom digitalnih datoteka. Oni postoje od upotrebe prvih zbiraka podataka na papiru kao što su katalozi u gradskim čitaonicama, poštarski dostavni nalozi, oznake datoteka itd. Povezuju i opisuju objekte u strukturama podataka te ukazuju na kontekst, upravljanje, obradu, očuvanje i korištenje resursa, dok muzejske primjerke, arhive i bibliotečne zapise povezuju sa terminom opisa i katalogizacije tih zapisa. Meta podaci u digitalnom izdanju su sigurniji naravno ako su pohranjeni na više izvora, a ostali papirnati primjerci ne mogu dugo ostati neoštećeni, osim ako nisu arhivirani u posebnim prostorijama.

4. PRIMARNE FUNKCIJE META PODATAKA

Primarne funkcije meta podataka povezane su značajkama koje su ovise jedna o drugoj kroz prepoznavanje, upravljanje, praćenje uporabe sadržaja i njegovim preuzimanjem (Baca M., 2008.) a sastoje se od:

- Stvaranja, višestrukog preusmjeravanja, ponovne uporaba, rekontekstualizacije informacijskog objekta
- organizacija opisa ili identifikacija resursa,
- provjera valjanosti,
- pretraživanja i pronalaženja,
- korištenja i očuvanja te
- rasporeda, odnosno informacije o lokacijama resursa.

4.1. Stvaranje, višestruko preusmjeravanje, ponovna uporaba, rekontekstualizacija informacijskog objekta

Objekt postaje digitalni informacijski sustav ako se stvori u digitalnom formatu ili njegovim pretvaranjem u digitalni format. Moguće je stvaranje više verzija istog objekta za očuvanje, izlaganje, istraživanje, proširenje pa čak i u svrhu razvoja. Rekontekstualizacija informacijskog objekta nalaže da neki administrativni i opisni meta podaci kreirani od strane moderatora ili oni koji su već bili u digitalnom formatu mogu i moraju biti uključeni u sustav upravljanja digitalnom imovinom radi mogućnosti ponovne uporabe.

4.2. Organizacija i opis

Bitna funkcija meta podataka je opis i organizacija izvornih objekata u spremišta ili zbirke kako bi se informacija objekta povezala sa originalnim objektom. Informacijski objekti mogu biti ručno ili automatski organizirani u strukturi digitalne informacije i može sadržavati opise koje generira moderator. Profesionalni informatičari mogu izraditi i dodatne meta podatke kroz registraciju , katalogizaciju i procesom indeksiranja, a mogu biti kreirani kroz druge oblike od strane korisnika tog informacijskog sustava.

4.3. Provjera valjanosti

Korisnici pregledavaju meta podatke i preuzete stavke i informacije kako bi utvrdili nadležnost i pouzdanost tih resursa.

4.4. Pretraživanje i pronalaženje

Dobro objašnjeni meta podatak neophodan je korisniku kako bi lakše došli do informacije i njegovog preuzimanja. Lokalni podaci koje korisnici pretražuju moraju biti lako dostupni kroz algoritme preuzimanja, korisničke transakcije i sustavno učinkoviti u skladištenju i pronalasku.

4.5. Korištenje i očuvanje

Informacijski objekti su podložni različitim vrstama upotrebe, modifikacije i reprodukcije tijekom njihovog ciklusa. Stoga digitalni podaci, a posebno oni već digitalno kreirani moraju biti podvrgnuti kontinuiranom režimu očuvanja, osvježavanja, migracije i provjere integriteta kako bi se osigurala njihova trajna dostupnost te je potrebna dokumentacija za svaku izmjenu koja se može pojaviti u informacijskom objektu tijekom procesa očuvanja.

4.6. Raspored

Metadata podaci su ključna komponenta u dokumentiranju rasporeda originalnih objekata i spremljenih stavki kao i informacijskih objekata koji se odnose na originale. Informacijski objekti koji su neaktivni ili nemaju nikakvu upotrebu poželjno je izbrisati.

5. VAŽNOST META PODATAKA

Bitan naglasak je na praktičnim načelima pri korištenju, stvaranju i održavanju meta podataka te upotreba meta podataka kroz povezivanje korisnika i informacija.

Način na koji se meta podaci koriste prikazati će jedan od primjera postavljanja opisa neke knjige na neki od web engine-a odnosno u bazu podataka same tražilice. Najčešće informacije vezane za knjigu su naslov, podnaslov i autor pa recimo ako se upiše samo naslov teško će se na taj način doći do potrebnih informacija. Upravo zato je potrebno „tagirati" odnosno označiti više atributa o toj knjizi kako bi ciljana publika došla do samog resursa.

Meta podaci su sastavljeni od elemenata (http://www.arma.org/bookstore/files/Franks-Kunde1.pdf od 15.09.2017.).
Oni opisuju karakteristike informacijskih objekata ili izvora. Dobro poznati elementi meta podataka su npr. autor, naslov i tema. Meta podaci imaju standardizirani izgled i sadržaj odnosno prezentaciju. Većina elemenata koristi dodatna pravila i standarde ili kodirane sheme za osiguranje prezentacije njihovog sadržaja. Elementi meta podataka ujedno opisuju izvor na standardiziran način, dosljednim rječnikom kako bi omogućili i olakšali pretraživanje i rukovanje informacijama. Drugim riječima meta podaci nameću strukturu nestrukturiranim informacijama (tj. dokumentima, mapama, audiovizualnom materijal, itd.) te nadodaju više strukture onima koji su već strukturirani (primjerice baza podataka). Struktura meta podataka na taj način je iskoristiva za svrhe pronalaska informacija, administracije, upravljanja spisima i čuvanja. Razvoj meta podataka rezultira pozitivnim promjenama u načinu na koji se informacije mogu pronalaziti i koristit.

Bitna stavka metadata podataka su tražilice (eng. search engine) ili pretraživači preko kojih moderatori ili web dizajneri unose podatke u strukture objekata od kojih korist imaju krajnji korisnici. Tražilice su jako korisne stranice gdje se može praktički pronaći bilo kakva vrsta informacija koja naravno sadrži meta podatke, a ako se pri tome koriste kolačići (eng. cookies) i ne brišu se iz povijest sa računala, mobitela i ostalih gadget-a mogu se pronaći informacije stare i do nekoliko mjeseci.

5.1. Web pretraživači

Koordinirani skup programa je definicija za pretraživač (Baca M., 2008.), a uključuju pauke ili botove koji se ubacuju na svaku stranicu i izvlače informacije koristeći hiper-tekstualne veze na svakoj stranici kako bi izvukao što više istoznačnih poveznica za podatak koji se pretražuje. Zatim sadrži i programe koji stvaraju velike indekse odnosno kataloge sa stranica koje su učitane. Program zaprima naš zahtjev za pretraživanje, a potom ga i uspoređuje sa katalogom te vraća korisniku. Alternativa tražilicama jest istraživanje strukturiranih datoteka povezanih sa odabranim temama. Neki od poznatih pretraživača su Google, Yahoo, a tražilice za kupnju predmeta su Amazon, Ebay i ostali.

Pretraživači Google-a i Yahoo-a (koji koristi bazu od Google-a) indeksiraju velik dio sadržaja sa weba te prikazuju ogromne količine podataka koje znaju zbuniti korisnika. Postoje i specijalizirane tražilice koje indeksiraju određenu vrstu podataka pa na primjer stranica Scribd je tražilica koja prikuplja, opisuje i nudi ogromnu količinu knjiga, raznih članaka, audio zapisa, novina pa i seminarskih radova, zatim Sportsdirect tražilica za kupnju sporteske obuće i odjeće, a tu su i Ebay i Amazon koji se baziraju na korisnicima koji pretražuju razne predmete i stvari za kupnju. Postoji niz različitih web mjesta koji su specijalizirani za ciljana indeksiranja ovisno o potrebama korisnika.

Načini pretraživanja podataka ovise o samom korisniku i njegovim potrebama pa ako na primjer pretraživanje na stranici Yahoo ispadne neuspješno uvijek je poželjno provjeriti po sličnim stranicama, a ponekad je vrlo vjerojatno moguće da ono što ne pronađete na osnovnim tražilicama uspijete pronaći na manje poznatim ali specijaliziranim stranicama.

5.2. Upotreba META tagova

Meta tagovi su informacije smještene u head prostor web tražilice ili stranice (https://www.greenfish.rs/clanci_meta_tagovi.php od 15.09.2017.). To su HTML-ovi koji pružaju informacije o sadržaju neke web stranice ili tražilice.

Meta tagovi imaju razne funkcije i mogu se koristiti na razne načine, na primjer za: opis stranice, ključne riječi koje pomažu bržem pronalasku podataka sa tražilice, koriste se u označavanju vlasnika stranice, autora stranice, karakternim kodovima i drugim funkcijama.

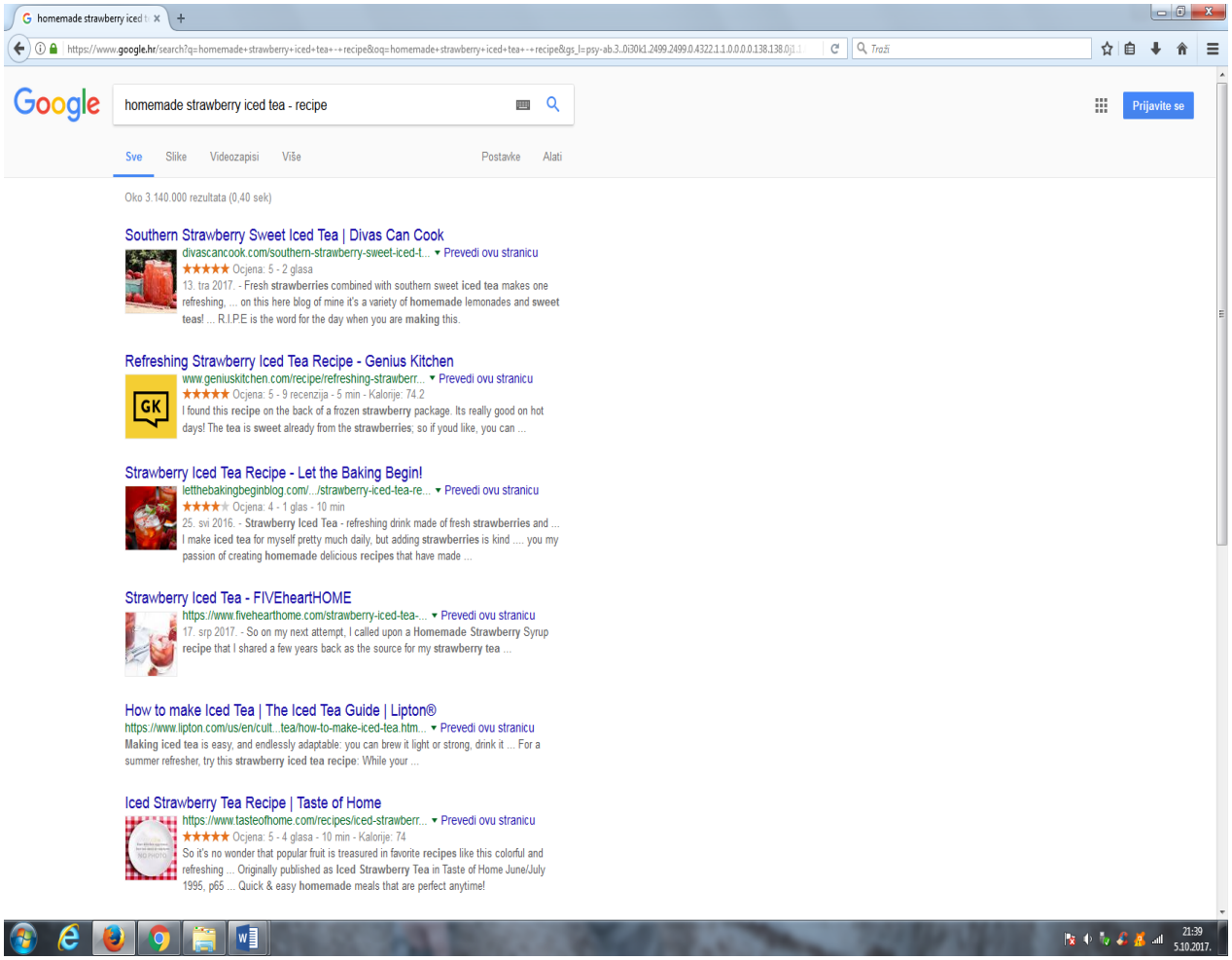
Korisnici nemaju previše koristi od tih tagova ali imaju korist od title taga. Osim title taga, informacije koje se nalaze u head prostoru ne mogu se vidjeti od strane korisnika te web stranice za razliku od samih moderatora ili web dizajnera koji naravno ih kreiraju i modificiraju prema potrebama sadržaja samih stranica.

Postoje i HTTP-EQUIV meta tagovi koji se koriste za zaglavlja stranica te služe korisniku kako bi odradio sve potrebne radnje prije nego počne sa title označavanjem. Ovim tagovima postiže se na primjer dužina postojanja stranice u cache memoriji tražilice.

5.3. Title tag

Title tag je značajan jer se u njemu nalazi naslov upita. Kako on nije meta tag dobro ga je spomenuti jer pravilno korištenje ovog taga omogućava vam preciznije rezultate pretraživanja. Preporuka je da title tag bude što kraći i da se najbitniji podaci upisuju na sam početak pretraživača, a da ne bude odviše dug. Na slici 3. je primjer jednog tagiranja po naslovu gdje je u tražilicu upisan recept za izradu sladoleda od jagode.

Slika 3. Title tag – recepti za izradu sladoleda od jagode

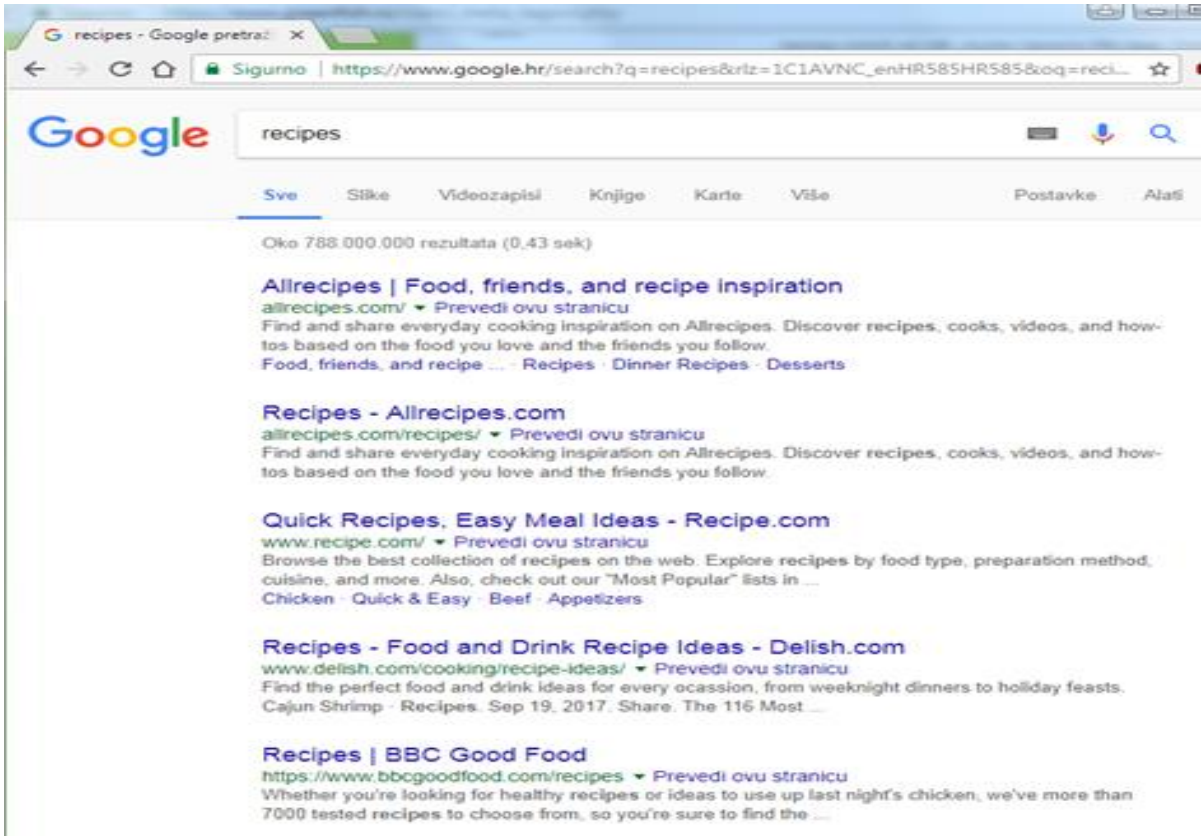


https://www.google.hr/search?q=homemade+strawberry+iced+tea+-+recipe&oq

Dakle tražilica je ponudila više recepta za sladoled od jagode jer se jasno označio podatak koji je korisniku bio potreban.

Kada bi pokušali pretražiti isti upit vezan za recepte, a da pritom želimo pronaći recept za izradu sladoleda od jagode te se upiše samo naslov recipe bez ostalih dodatnih tagova dobili bi veliku količinu sadržaja vezanih za skoro čitav sadržaj o receptima kojom je tražilica opisana, kao na slici 4.

Slika 4. Title tag – recepti



https://www.google.hr/search?dcr=0&source=hp&q=recepti&oq

Ovakav primjer prikazuje kako ponekad pretraživanje ciljanih informacija može biti naporno ako se pretraga bazira na nedorečenom naslovnom tagu pa ponekad samo dodavanjem jedne ili više naslovnih riječi odnosno tagova uvelike pomaže u ciljanoj pretrazi.

Postoje različiti primjeri tagiranja, a u nastavku su opisani oni koji su od velike koristi web dizajnerima i moderatorima web stranica ili pretraživača su:

- deskriptivni (eng. description),
- ključne riječi (eng. keywords),
- jezični tagovi (eng language),
- tagovi osvježavanja (eng. refresh),
- vremenski (eng. revisit),
- robot,
- autorski (eng. author),
- prava kopiranja (eng. copyright) i
- apstraktni meta tag (eng. abstract).

U deskriptivnom tagu se nalaze najvažniji opisi sadržaja, a podržavaju ga uglavnom manje poznati pretraživači, dok u manjem obimu Google, Yahoo jer oni uglavnom sami kreiraju opis sadržaja.

Tag ključne riječi sadrži serije ključnih fraza koje se nalaze na web stranicama. Mogu se postavljati specifične fraze koje bi korisnik poželio pretraživati ukoliko se stranica bazira na takvoj strukturi baze podataka.

Jezični tagovi definiraju na kojem jeziku je opisana stranica. Potreban u situacijama gdje se također mogu modelirati specifične riječi nekog jezika.

Tagovi osvježenja mogu biti od koristi za stranice gdje se mijenjaju imena same stranice ili prilikom premještanja stranice na drugu adresu. Osvježenje adrese stranice se može postaviti na određeni vremenski period, a najbolja korist je kada korisnik ispunjava neku formu na stranici pa se na taj način oslobađa prostor odnosno brišu zapisi sa stranice nakon osvježenja.

Tag koji definira kada će se korisnik ili posjetitelj stranice vratiti i to u danima je vremenski tag. Korisno za stranice na kojim se mijenja sadržaj početne stranice.

Robot tag definira koje će stranice indeksirati te koje će linkove pratiti.

Autorski tagovi sadrže imena autora stranica, ovaj tag je nepotreban ali može se iskoristiti.

Pravni tagovi se koriste se kod pravila kopiranja.

Tag sličan deskriptivnom je apstraktni tag gdje ga moderatori stranica ubacuju kao opis rezultata dobiveni pretraživanjem.

5.4. Dark Web

Dark web ili nevidljivi web je dio svjetskog web-a koji zahtjeva posebne aplikacije za pristup. Kada se pristupi ovoj vrsti web-a stranice i usluge se ne razlikuju od standardnih web stranica.

Kod nevidljivog neke web lokacije su skrivene i može im se pristupiti ako se posjeduje ispravna adresa ili je korisnik pozvan od strane drugog korisnika. Na takvim skrivenim web lokacijama djeluju posebna tržišta pod nazivom crna tržišta (eng. darknet) gdje se mogu pronaći nezakoniti proizvodi, a kupovina se vrši kripto-valutama kao što je bitcoin, ethereum, litcoin, minero i mnoge druge.

Jedna od aplikacija za otvaranje nevidljivih stranica zove se TOR (http://theconversation.com/explainer-what-is-the-dark-web od 17.09.2017.) i funkcionira na način da omogućava pregled stranica s time da zadržava anonimnost posjetitelja bez mogućnosti njegovog praćenja odnosno stranice nisu kreirane kako bi se tražilicama pristupilo nekom sadržaju osim ako moderatori ne implementiraju indeksiranjem poznatih web pretraživačima na nevidljivim web stranicama.

S sigurnosne strane ovakva vrsta weba jako je nesigurna jer ako posjetitelj ne može vidjeti moderatore tih stranica to ne mora značiti da oni ne vide posjetitelje, stoga ovakva vrsta stranica potpuno je nesigurna i trebalo bi je izbjegavati.

6. UDAR NA PRIVATNOST

Nakon definicije i važnosti meta podataka, nameće se pitanje sigurnosti ili bolje nesigurnosti koje one donose.

Kako je baza podataka na internetu velika i resursno povezana sa tražilicama, negativno je to što prilikom upisivanja nekih podataka u tražilicu ona će izvući sve što je povezano uz te podatke od imena i prezimena pa sve do lokacije gdje je zadnji put korisnik pustio trag.

Konkretan primjer je facebook koji je ujedno i baza podataka koja se svakim danom povećava i obnavlja te je jako ranjiva zbog privatnih informacija koje se spremaju u njenu bazu podataka.

Tu su primjeri web tražilica i mjesta koje preko kolačića spremaju podatke i informacije na računalo, a kasnije te informacije iskorištavaju kako bi olakšali ponovno i daljnje korištenje tih podataka ali uz to i omogućile drugim osobama uvid u te podatke odnosno moderatorima baza podataka uvid u vaše meta podatke, a potom i na privatne informacije što je udar na privatnost.

U pravilu moderatori odnosno kreatori baza podataka su oni koji imaju uvid u meta podatke korisnika praktički bez njihovog pristanka. Jedan od primjera su poslužitelji mobilnih usluga gdje mobiteli sadrže potrebne informacije kao bi mobitel korisnika ispravno funkcionirao.

Negativna posljedica upotrebe lokacijskih usluga je taj što poslužitelji uvijek imaju uvid gdje se korisnik mobitela nalazi zbog informacija koje sadrži taj mobitel, a to je IMEI broj i serijski broj SIM kartice koja je u GPS sustavu. Svako spajanje na mrežu bilo putem WI-FI-a ili mobilnih podataka pa i Bluetooth-om postaje podatak. Zatim rješavanje online kvizova i anketa putem mreže je također negativan primjer jer se tako otkrivaju osobne tendencije i interesi te bi ih se zato trebalo izbjegavati.

Velike kompanije također su podložne sigurnosnim udarima, a primjeri su broji napadi na korisničke račune unutar kompanija. Sve poslovne informacije mogu biti ugrožene zbog najjednostavnijeg napada na jedan korisnički račun. Zbog toga mnoge kompanije kako bi sačuvale svoje informacijske vrijednosti ulažu velike napore kako bi osigurale svoje podatke zaštitom internih servera te sigurnosnim kopijama (eng. back up) sustava, upotrebom izjava o

uvjetima korištenja te korporativnim odjelima zaduženim za sigurnost osobnih podataka pa tako i važnih podataka kompanije.

Elektronička pošta je zamjena za klasičnu poštu gdje su paketi u virtualnom svijetu oblik privitaka u elektronskoj pošti. Takva vrsta pošte je svaka komunikacija u obliku teksta, slike, zvuka, glasa, naravno prenesena elektroničko komunikacijskom mrežom gdje ne postoje zakoni definirani koji bi zaštitili korisnika od neželjene e-mail pošte odnosno prijevara putem istih.

Najčešći udari na privatnost korisnika je e-mail račun gdje se pošiljatelji pogrešno predstavljaju u namjeri da prevare primatelje. Postoje razni načini kako prevariti korisnike (https://www.lifewire.com/top-internet-email-scams-2483614 od 30.09.2017.), a oni su: phishing e-pošti, loto prijevare, traženje pozajmica, traženje novaca nakon elementarnih nepogoda, lanac mail poruka sa naslovima „Make money fast“ i sličnim.

Potrebno je spomenuti da većinu informacija koje se nalaze na web-u mogu doći u posjed američke službe za sigurnost (NSA) koja marljivo prikuplja informacije i pretražuje baze podataka sa interneta kako bi mogli što bolje i jasnije ući u svaki podatkovni trag.

6.1. Meta podaci i Facebook

Facebook kao društveno-socijalna aktivnost postiže svoj maksimum i na samom je vrhu mrežnih servisa.

Zašto je toliko bitan pokazuju tri važna segmenta kroz proces njegovog korištenja, a oni su: prikupljanje podataka, pohrana i algoritamski procesi i ciljana grupa korisnika.

Brojka od 1.4 milijardi mjesečno aktivnih korisnika u 2016. godini te 890 milijuna dnevno aktivnih korisnika (slika 5.), pokazuje koliko je facebook postao moćan mrežni servis sa ciljem povezivanja korisnika cijelog svijeta (https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/ od 05.10.2017.). Misija pa i vizija Zuckerberga inače tvorca ove popularne društvene mreže je povezivanje što većeg broja korisnika kako bi svijet postao otvoreniji i povezaniji. Mnogo je razloga zašto je ovom trendu potrebna tolika posveta, a onaj najvažniji je kontrola osobnih podataka koji se u ovako velikom otvorenom svijetu informacija jako teško može pratiti pa i očuvati.

Slika 5. Prosjek aktivnosti korisnika



https://labs.rs/wp-content/uploads/2016/08/FB-Research-Final-01.png

Facebook koristi različite načine za pohranu podataka pa se tako svaka korisnikova radnja pohranjuje u bazu podataka samog facebook-a, zatim prati svaki digitalni trag bilo preko mobilne aplikacije ili rada na računalima te prate pokreti i radnje korisnika putem kolačića koji su povezani i sa svim ostalim mrežnim aplikacijama i partnerima u facebook-ovoj mreži. Facebook okuplja preko stotinu partnera kao što su partneri za data algoritam, marketing, tehnološke procese i mnogi drugi. Recimo tvrtka Oculus Rift može prikupljati i ustupiti podatke kretanja i prilikom korištenja opreme za virtualnu stvarnost u Facebook mreži, isto

tako postoji cijeli niz aplikacija koje su usko povezane sa ovim mrežnim gigantom te otkrivaju svaki korisnikov pokret (https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/ od 05.10.2017.).

6.2. Zloupotreba podataka putem e-mail-a

Jedan od pokušaja slučajnog pronalazaka e-mail podataka sa web poslužitelja odvijao bi se na način da prevarant pronađe web mail poslužitelja koji nema određenu jačinu autentifikacije pa prevarant lagano dođe do informacije o lozinki. Adresa e-mail-a kao meta podatak je već vidljiva na raznim forumima kao i ime i prezime korisnika. Prevarant na osnovu podataka koje ima o oštećenom korisniku, a to je e-mail adresa te osobni podaci, ime i prezime koje također može iskoristiti sa foruma, pristupa recimo Facebook stranici te zatraži novu lozinku, a razlog navede da ju je zaboravio te prijavi e-mail račun koji je pobrao sa foruma te dobije novu lozinku pa tako ima uvid u tuđi Facebook račun, odnosno može napraviti veliku štetu na tom računu (https://kompjuteras.com/privatnost-na-internetu-praktican-primer-zloupotrebe/ od 25.09.2017.).

6.3. Zlouporabe podataka preko IP adrese

Haker ili prevarant koji želi doći do potrebnih podataka o nekom korisničkom računu ili bilo kakvoj informaciji može na razne načine. Primjer gdje korisnik određene forumske stranice ostavi trag u vidu IP adrese, a osoba koja želi doći do podataka na osnovi IP adrese korisnika, koja je vidljiva na svakoj forumskoj stranici može doći putem posebnih stranica do IP adrese poslužitelja usluge gdje se nalaze ključni podaci o korisniku. Preko tih podataka dolazi do adrese stanovanja korisnika, imena i prezimena pa haker već ima sve potrebne informacije koje može iskoristiti na način prateći svaki korisnikov pokret na forumu te može vrlo jednostavno djelovati kada se korisnik forumske stranice najmanje nada (https://kompjuteras.com/privatnost-na-internetu-praktican-primer-zloupotrebe/ od 25.09.2017.).

6.4. Slabosti društvenih stranica

Društvene stranice poput facebook-a, twiter-a, instagram-a pa i youtube-a, sadržavaju ogromne količine kako podataka, tako i meta data podataka. Samo na jednom računu facebook-a možete pronaći više desetaka meta podataka koji otkrivaju vaš identitet, a zlouporabom kao što je opisano u prethodnim poglavljima mogu korisnicima stvarati probleme.

6.5. Zaštita podataka

Jedno od gorućih pitanja je kako meta podatke pa tako i podatke koje ostaju zapisani u povijesti stranica odgovarajuće zaštiti od strane moderatora i vanjskih napada.

Pristupanjem odnosno kreiranjem računa na svakoj stranici nakon upisivanja osobnih podataka ili prije same registracije pojavljuje se izjava o uvjetima korištenja sa kojim se možete, a i ne morate složiti. Naravno ukoliko se ne složite ne možete pristupiti željenoj stranici. Takvim dokumentom poslužitelji se ograđuju od relativno velikog broja odredbi kako bi zaštitili vašu privatnost, a posebno vlastitu. Ali tu se nameće pitanje što ako se dogodi napad treće strane koja se ne nalazi u izjavi o uvjetima korištenja. Na koji način se korisnik može zaštiti od takvih napada, s najbolji odgovor bio bi odgovornim i svjesnim ponašanjem. Najviše napada koji pogađa korisnike društvenih mreža od strane trećih su e-mail poruke u koje je ugrađen sigurnosni napad kroz obične tekstualne poruke koje upravo imaju cilj napasti određenu grupu korisnika. Obrana od takvih napada je ne otvarati sadržaje koji su nepoznati, nemaju smisleni naslov, a najvažnije od svega je informirati se na vrijeme te ažurirati sigurnosne antivirusne aplikacije kako na računalima tako i na mobilnim aparatima ažurirati aplikacije.

Uporaba kolačića može biti još jedan u nizu sigurnosnih propusta ako se njima ne postupa na odgovarajuće načine. Kolačići se upotrebljavaju za praćenje uporabe interneta i izradu korisničkog profila, a potom i za prikaz prilagođenih Internet oglasa na temelju vaših preferencija.

Pa ako koristite kolačiće na nekoj stranici za kupnju prilikom odabira nekog artikla tražilica će vam ponuditi uz odabrani artikl na osnovi meta podataka iz kolačića drugih korisnika sve artikle koje su ti određeni korisnici kupovali uz odabrani artikl. Uz to zabilježiti će u tom

trenutku artikl koji je odabran i putem takvog odabira proširiti indeks te stranice pa svaki put kada se korisnik prijavi stranica će mu predložiti iste i slične proizvode koje je zadnji put koristio što znači da na neki način takva stranica prati korisnikove pokrete i navike pa tako narušava korisnikovu privatnost.

Dakle najbolji način obrane osobnih podatak skoro i ne postoji ali pristranim, savjesnim i odgovornim ponašanjem, a to je odavanje što manje podataka na web stranicama, redovitim ažuriranjem sustava, antivirusnim provjerama te upotrebom raznih aplikacija za brisanje informacija spremljenih u povijest tražilica, moguće se je obraniti od neželjenih napada i krađe podataka i identiteta.

6.6. Europska komisija

Europska Komisija pokrenula je inicijativu zaštite privatnosti na internetu postavljanjem pravila o visokoj razini zaštite privatnosti po pitanju svih elektroničkih komunikacija te ažuriranja pravila o zaštiti podataka za institucije EU-a (http://europa.eu/rapid/press-release_IP-17-16_hr.htm?locale=EN od 30.09.2017.).

Ovom direktivom Europska Komisija, a prijedlog prođe svi online servisi i web usluge morati će se prikloniti strožim europskim pravilima o zaštiti privatnosti i biti će izjednačeni s pravilima koja se odnose na telekom operatore. Tako će web poslužitelji imati određen limit kada je u pitanju elektronička komunikacija, odnosno protok podataka.

Predložena uredba o zaštiti privatnosti i elektroničkim komunikacijama povećati će zaštitu privatnih podataka građana i otvoriti nove mogućnosti poduzećima pa će se tako strože kontrolirati društvene stranice i web poslužitelje poput Facebook-a, Skype-a, Gmail-a, Viber-a, Messengera te sličnih socijalno-društvenih mrežnih usluga.

Kod sadržaja komunikacije i meta podataka jamčit će se privatnost i u pogledu sadržaja i u meta podacima koji proizlaze iz elektroničkih komunikacija (mjesto i datum slanja, vrijeme slanja i slično). Korisnici će imati mogućnost zabrane odavanja takvih podataka od strane web poslužitelja te se ti podaci neće ni prikazivati u slučaju zabrane.

Prema propisima Europske komisije određeno je da svaka Internet stranica na kojoj se upotrebljavaju kolačići, mora obavijestiti korisnika o tome, a korisnik mora dati pristanak. (Schenier, B, 2015.). Korisnik ima pravo izbora u postavkama Internet pretraživača i deaktivirati upotrebu kolačića.

Pravila bi se trebala početi primjenjivati od 25. svibnja 2018. godine.

7. ZAKLJUČAK

Baze podataka konstantno rastu pa je tako rastu i meta podatci koje nose svoje pozitivne i negativne učinke. Cijeli sustav Interneta je povezan, a svaki indeksirani podatak zapisan je u neku bazu podataka.

Pozitivna strana meta podataka je u tome što se u bilo kojem trenutku korisnik pretraživača može na brz i efikasan način doći do potrebne informacije, a sve to zbog velike baze podataka, ubrzanog indeksiranja i kvalitetno modeliranih meta i title tagova.

Sa negativnog aspekta meta podaci u neku ruku konstantno ukazuju na potrebu opreznosti pri njihovom korištenju. Opasnosti na pretraživačima nikada ne miruju ali odgovornim ponašanjem te ispravnim odlukama, zapreke se mogu izbjeći. U konačnici kvalitetnim i pravilnim upravljanjem informacijama značajno se povećava sposobnost razmjene i očuvanja informacija te zaštita od udara na privatnost.

8. LITERATURA

Baca, M. (2008.) *Introduction to Metadata*, Los Angeles, CA, J. Paul Getty Trust

Schenier, B (2015.) *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World.* New York, W.W. Norton & Company

Register, R. i Mcllroy, T. (2012.) *The METADATA Handbook*. Columbus, OH, a DATACURATE

Krsmanović, D. (2007.) *Upotreba META tagova* [Online] Dostupno na: https://www.greenfish.rs/clanci_meta_tagovi.php [15.09.2017.]

Kompjuteraš (2014.) *Privatnost na internetu* [Online] Dostupno na:

https://kompjuteras.com/privatnost-na-internetu-praktican-primer-zloupotrebe/ [25.09.2017.]

Franks P., Kunde, N. (2008.) *Why metadata matters* [Online] Dostupno na: http://www.arma.org/bookstore/files/Franks-Kunde1.pdf [15.09.2017.]

National Information Standards Organization. (2004.) *Understanding metadata.* [Online] Dostupno na: http://www.niso.org/publications/press/UnderstandingMetadata.pdf [15.09.2017.]

Data-Ed Online. (2014.) *Metadata Management Strategies.* [Online] Dostupno na: https://www.slideshare.net/aberkowitz/2014-metadata-management-strategieskey-41655167 [17.09.2017.]

The Conversation. (2017.) *Explainer: what is dark web?* [Online] Dostupno na: http://theconversation.com/explainer-what-is-the-dark-web-46070 [17.09.2017.]

Wikipedia. (2017) *Metadata*. [Online] Dostupno na: http://en.wikipedia.org/wiki/Metadata [17.09.2017]

Publish Open Access. (2010.) *The Open Archives Initiative Protocols.* [Online] Dostupno na: https://publishopenaccess.blogspot.hr/2010/11/open-archives-initiative-protocol.html [17.09.2017.]

Lifewire. (2017.) *The Top 10 Internet and Email Scams* [Online] Dostupno na: https://www.lifewire.com/top-internet-email-scams-2483614 [30.09.2017.]

European Commission. (2017.) *Press Release Database* [Online] Dostupno na: http://europa.eu/rapid/press-release_IP-17-16_hr.htm?locale=EN [30.09.2017.]

Share Lab. (2016.) *Immaterial Labour and Data Harvesting* [Online] Dostupno na:

https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/ [05.10.2017.]

9. POPIS SLIKA

Slika 1. Primjer meta podatka u e-mailu - https://www.slideshare.net/aberkowitz/2014-metadata-management-strategieskey-41655167

Slika 2. Metadata harvesting model - https://publishopenaccess.blogspot.hr/2010/11/open-archives-initiative-protocol.html

Slika 3. Title tag – recept za izradu sladoleda od jagode - https://blog.adeptmarketing.com/seo-title-tags-101/

Slika 4. Title tag – recepti - https://www.google.hr/search?dcr=0&source=hp&q=recepti&oq=recepti&gs_l=psy-ab.3..0i131k1j0j0i131k1j0.2559.5106.0.5649.9.8.1.0.0.0.146.768....0...1.1.64.psy-

Slika 5. Prosjek aktivnost korisnika - https://labs.rs/wp-content/uploads/2016/08/FB-Research-Final-01.png

10. SAŽETAK

Metadata podatak jesu informacije o podatku. Postoji više funkcija i atributa metadata podataka u baznim strukturama. Obilježje tih funkcija i atributa je da opisuju, upravljaju, pohranjuju i pronalaze odgovarajuće podatke u svakodnevnici, upotrebom meta tagova. Pri upotrebi meta podataka potrebno je dati naglasak na sigurnost upravljanja. Mogu se pronaći na svim Internet stranicama i pretraživačima, odnosno oni su sastavni dio meta podataka. Stoga, postoji čitav niz negativnih činjenica zašto su meta podaci opasni te mogu ugrožavati sigurnost korisnika, a da bi se takve situacije izbjegle potrebno je savjesno i odgovorno ponašanje prilikom korištenja podataka na internetu. S druge strane, oni su od velike koristi jer se preko pretraživača i naravno na internetu može pronaći bilo kakva informacija, naravno uz pravilno indeksiranje podataka.

Ključne riječi: meta podatak, funkcije, atributi, meta tagovi, sigurnost meta podataka.

11. SUMMARY

Metadata data are about data information. There are multiple functions and attributes of metadata data in base structures. The features of those functions and attributes are that they describe, manage, store and find relevant data in everyday life, using Meta tags. When using metadata, emphasis should be placed on the security of management. They can be found on all Internet sites and browsers, or they are an integral part of metadata. Therefore, there is a whole range of negative facts about why metadata is so dangerous and could endanger the security of users, and to avoid such situations, it is necessary to have conscientious and responsible behavior when using data on the Internet. On other hand, they are of great use because through the Internet and the search engines can find any information, of course, with correct indexing of data.

Key words: metadata, functions, attributes, Meta tags, metadata security.