

BAZE ZNANJA TEMELJENE NA SEMANTIČKOM WEBU - PRIMJER I PODRUČJA UPOTREBE

Vukomanović, Lovro

Undergraduate thesis / Završni rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Economics in Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Ekonomski fakultet u Osijeku**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:145:889421>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-08-18**



Repository / Repozitorij:

[EFOS REPOSITORY - Repository of the Faculty of Economics in Osijek](#)



Sveučilište Josipa Jurja Strossmayera u Osijeku
Ekonomski fakultet u Osijeku
Preddiplomski studij, smjer *Poslovna Informatika*

Lovro Vukomanović

**BAZE ZNANJA TEMELJENE NA SEMANTIČKOM WEBU –
PRIMJER I PODRUČJA UPOTREBE**

Završni rad

Osijek, 2021.

Sveučilište Josipa Jurja Strossmayera u Osijeku
Ekonomski fakultet u Osijeku
Preddiplomski studij, smjer *Poslovna Informatika*

Lovro Vukomanović

**BAZE ZNANJA TEMELJENE NA SEMANTIČKOM WEBU –
PRIMJER I PODRUČJA UPOTREBE**

Završni rad

Kolegij: Upravljanje informacijskim resursima

JMBAG: 0010226783

e-mail: lovrovukomanovic@gmail.com

Mentor: prof. dr. sc. Josip Mesarić

Osijek, 2021.

Josip Juraj Strossmayer University of Osijek
Faculty of Economics in Osijek
Undergraduate study (Poslovna Informatika)

Lovro Vukomanović

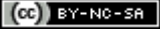
**SEMANTIC WEB KNOWLEDGE BASE – EXAMPLE AND
AREAS OF USE**

Finalpaper

Osijek, 2021.

IZJAVA

OAKADEMSKOJ ČESTITOSTI, PRAVU PRIJENOSA INTELEKTUALNOG VLASNIŠTVA, SUGLASNOSTI ZA OBJAVU U INSTITUCIJSKIM REPOZITORIJIMA I ISTOVJETNOSTI DIGITALNE I TISKANE VERZIJE RADA

1. Kojom izjavljujem i svojim potpisom potvrđujem da je završni (navesti vrstu rada: završni / diplomski / specijalistički / doktorski) rad isključivo rezultat osobnoga rada koji se temelji na mojim istraživanjima i oslanja se na objavljenu literaturu. Potvrđujem poštivanje nepovredivosti autorstva te točno citiranje radova drugih autora i referiranje na njih.
2. Kojom izjavljujem da je Ekonomski fakultet u Osijeku, bez naknade u vremenski i teritorijalno neograničenom opsegu, nositelj svih prava intelektualnoga vlasništva u odnosu na navedeni rad pod licencom *Creative Commons Imenovanje – Nekomercijalno – Dijeli pod istim uvjetima 3.0 Hrvatska*. 
3. Kojom izjavljujem da sam suglasan/suglasna da se trajno pohrani i objavi moj rad u institucijskom digitalnom repozitoriju Ekonomskog fakulteta u Osijeku, repozitoriju Sveučilišta Josipa Jurja Strossmayera u Osijeku te javno dostupnom repozitoriju Nacionalne i sveučilišne knjižnice u Zagrebu (u skladu s odredbama Zakona o znanstvenoj djelatnosti i visokom obrazovanju, NN br. 123/03, 198/03, 105/04, 174/04, 02/07, 46/07, 45/09, 63/11, 94/13, 139/13, 101/14, 60/15).
4. izjavljujem da sam autor/autorica predanog rada i da je sadržaj predane elektroničke datoteke u potpunosti istovjetan sa dovršenom tiskanom verzijom rada predanom u svrhu obrane istog.

Ime i prezime studenta/studentice: Lovro Vukomanović

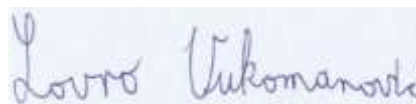
JMBAG: 0010226783

e-mail za kontakt: lovrovukomanovic@gmail.com

Naziv studija: Preddiplomski sveučilišni studij, Poslovna ekonomija, smjer *Poslovna Informatika*

Naslov rada: Baze znanja temeljene na semantičkom webu - primjer i područja upotrebe

U Osijeku 10. srpnja 2021. godine



Potpis _____

SAŽETAK

Semantički web može se definirati kao produžetak samoga WWW-a u kojem se daje definirano značenje informacijama, a ne bavi se strukturom podataka. Može se reći kako je semantički web prva nadogradnja koja je nužna za uvođenje kako bi World Wide Web (poznatiji kao WWW ili web) doživio evoluciju. Semantički web pruža sposobnost razumijevanja i predstavljanja podataka i informacija u oblicima koji odgovaraju korisniku, karakterizira ga jednostavnost i jasnoća što uvelike olakšava korištenje. Također, može se reći kako je semantički web nova generacija u fazi razvoja weba. Kod semantičkog weba važno je razumijevanje razmjene podataka i poznavanje baze znanja. Semantički web nije kreiran na ideji korištenja umjetne inteligencije ili stjecanja samosvijesti. Ono što ga karakterizira kao proširenje, jest ideja dodavanja metapodataka na web s ciljem stvaranja WWW-a strojno čitljivim, odnosno njegovo stjecanje alata za pronalaženje, razmjenu i ograničeno tumačenje informacija. Potrebno je usmjeriti se na upravljanje znanjem te poznavati alate za održavanje slabo strukturiranih podataka koji mogu predstavljati problem u rješavanju problema. Ovakvim pristupom korača se, polako ali sigurno, k budućoj generaciji koja će omogućiti komunikaciju strojevima u svrhu razmjene informacija te na kraju, čovjeku lakši pristup i kontekstualno razumijevanje svih prikupljenih informacija.

Ključne riječi: semantički web, podatci, informacije, baza znanja

ABSTRACT

Many people still don't have a clear idea of what the Semantic Web is. The word semantics itself is meaning or understanding. Simply put, the Semantic Web can be defined as an extension of the current network in which information is given a defined meaning rather than being concerned with data structures. The Semantic Web is arguably the first upgrade that the World Wide Web (known as the WWW or web) must adopt in order to evolve. The Semantic Web provides the ability to understand content and present it in a format that is useful to users. It's much easier to use because it's easier to understand and simpler. The Semantic Web also represents a new generation of web development phases. The Semantic Web not only enhances traditional search, but also provides simplicity, intelligence, and understandability. Sending data and knowing the knowledge base are very important in the Semantic Web. The way it works is that it collects and links various pieces of data found on the Internet and matches them against the original command. Knowledge management deserves special attention, and we need to know tools for maintaining unstructured data that can cause problems in solving problems. This approach is slowly but surely moving towards future generations that will enable communication with machines for the purpose of exchanging information, and ultimately facilitate access to all information collected.

Keywords: semantics, data, information, knowledge base

SADRŽAJ

1. UVOD	1
1.1 Struktura rada	1
2. METODOLOGIJA RADA	2
2.1. Predmet istraživanja.....	2
2.2. Metode istraživanja	2
2.3. Izvori istraživanja	2
2.4. Ciljevi istraživanja	3
3. KRATKA POVIJEST WEBA I SEMANTIČKOG WEBA	4
4. SLOJEVI SEMANTIČKOG WEBA	7
4.1. Semantički preglednici	9
4.1.1. Merit.....	9
4.1.2. Disco	9
4.1.3. LodLive	10
4.2. Semantičke tražilice weba	10
4.2.1. DuckDuckGo	12
4.2.2. Mojeek.....	13
4.3. Ontologija	14
4.4. URI	14
4.5. Linked Open Data.....	15
4.6. OWL.....	16
4.7. RDF	17
4.8. XML	17
4.9. SPARQL	18
5. IMPLEMENTACIJA SEMANTIČKOG WEBA NA STVARNOM PRIMJERU.....	19
6. RASPRAVA.....	21
7. ZAKLJUČAK.....	22
8. LITERATURA	22
9. POPIS SLIKA	24
10. POPIS TABLICA.....	25

1. UVOD

Kako bi se poboljšao način upravljanja podacima (strukturiranim i nestrukturiranim) raste broj organizacija koje se okreću k semantičkom webu. Graf baze podataka i ontologija samo su neki od semantičkih alata koji omogućuju učinkovitije upravljanje sadržajem s naglaskom na otkrivanje i pretraživanje podataka. Vraćaju na površinu skrivene, zaboravljene i nepersonalizirane informacije. Zbog jedinstvenih veza između resursa npr. A i resursa npr. B, podatke pretražuju izrazito brzo i točno. (Enterprise Knowledge, 2017.)

Može se, za semantički web, sa sigurnošću reći da je prva nadogradnja koju mora uvesti kako bi World Wide Web (WWW ili web) doživio evoluciju. Kronološki gledano, može se reći kako je World Wide Web tek dospio u svoju drugu generacijsku dob. Prethodna generacija bi bila okarakterizirana „ručnim“ pisanjem koda HTML stranica te statičkom generiranju stranica. Kroz godine, pojavili su se unaprijed napisani kodovi u određenom programskom jeziku koji su omogućavali generiranje dinamičkih stranica te je samim time Web ušao u drugu generaciju. Neki od korištenih jezika su PHP, ASP, CGI itd. Pod pokroviteljstvom World Wide Consortiuma (W3C) korača se, polako ali sigurno, k budućoj generaciji koja će omogućiti komunikaciju strojevima u svrhu razmjene informacija te na kraju, čovjeku olakšati pristup svim prikupljenim informacijama i interakciju s web-stranicama i drugim korisnicima.

1.1 Struktura rada

Rad se sastoji od 7 poglavlja. Prvim poglavljem čitatelja se uvodi u završni rad. U drugom poglavljju ili metodologiji rada promatraju se ciljevi, predmeti, izvori i metode istraživanja. U trećem poglavljju se prezentira povijest weba te semantičkog weba. Uvelike se spominje i tvorac istih. U četvrtom poglavljju predstavljena je hijerarhijska podjela semantičkog weba te se detaljizira svaki sloj navedene hijerarhije. U petom poglavljju se za primjer daje stvaran primjer semantičkog weba na temelju korporacije Chevron. Šesto i sedmo poglavljje su rasprava i zaključak rada. Na dnu završnog rada nalazi se popis literature, popis tablica te popis slika.

2. METODOLOGIJA RADA

U metodologiji rada obuhvaćeni su predmeti istraživanja, metode istraživanja, izvori istraživanja te ciljevi istraživanja.

2.1. Predmet istraživanja

Predmet istraživanja ovog rada su baze znanja temeljene na semantičkom webu. Baza znanja se veže za umjetnu inteligenciju. Predstavlja jedan od tri ključna dijela ekspertnog sustava. Baza znanja sadrži znanje koje je predstavljeno u obliku nestrukturiranog skupa činjenica i pravila. Semantički web postavlja se u kontekst weba i osigurava inteligentni pristup heterogenim, distribuiranim informacijskim sadržajima. Područje primjene neprekidno se širi i danas obuhvaća upravljanje znanjem, elektroničko poslovanje, razmjenu informacija u P2P sustavima, procesiranje prirodnih jezika itd.

2.2. Metode istraživanja

Najčešće korištene metode istraživanja bile su: induktivna metoda, deduktivna metoda, metoda analize, metoda generalizacije, metoda specijalizacije, metoda klasifikacije, metoda deskripcije. Valja naglasiti da ovaj rad nije istraživačkoga karaktera, nego predstavlja pokušaj sinteze aktualnih znanja u jednoj specifičnoj domeni što uključuje i analizu konkretnog slučaja i procjenu budućeg razvoja. Zbog toga će se u radu koristiti induktivne i deduktivne metode zaključivanja te dijelom klasifikacija, analiza i sinteza.

2.3. Izvori istraživanja

Izvori istraživanja su sekundarne i tercijarne publikacije (znanstveni članci u časopisima, knjige, enciklopedije) te stručni izvori, pretežito raspoloživi na internetu.

2.4. Ciljevi istraživanja

Cilj istraživanja u ovom radu je: kroz usvojene definicije i shvaćanja pojmova semantičkog weba i baza znanja prikazati aktualno stanje u području baza znanja koje se formiraju na semantičkom webu, prikazati način funkcioniranja na konkretnom primjeru i procijeniti pravce budućeg razvoja ovog kompleksnog područja.

3. KRATKA POVIJEST WEBA I SEMANTIČKOG WEBA

Semantika, također se naziva i semiologija, semiotika ili semaziologija je termin za znanstveno i filozofsko izučavanje značenja u umjetnim, ali i prirodnim jezicima. Pojam je nastao iz skupine engleskih riječi nastalih iz korijena grčke riječi *sēmainō* što znači „značiti“ ili „označavati“, no teško je oblikovati zasebnu definiciju za svaku izvedenicu ovog izraza zbog preklapanja u literaturi unatoč individualnim preferencijama. Semantički web također se naziva proširenjem World Wide Weba (WWW) gdje se podacima daje značenje (semantika) kako bi računalo moglo pretražiti i definirati odgovor na korisnički zahtjev, upit ili pretraživanje kao takvo. Jedan od glavnih zagovornika semantičkog weba je Sir Tim Berners-Lee, britanski izumitelj WWW-a i direktor konzorcija World Wide Weba (W3C) koji ujedno i nadgleda standarde koji su potrebni za projekt. Ideju o semantičkom webu osmislio je sir Tim Berners-Lee 1989. godine, iste godine kada je i dovršio WWW te je istu ideju predstavio na prvoj međunarodnoj WWW konferenciji održanoj 1994. godine što je rezultiralo stvaranjem W3C (World Wide Web Consortium). Konzorcij za Word Wide Web (W3C) je međunarodna zajednica u kojoj članovi, stalno zaposleno osoblje i javnost, rade zajedno na razvoju web-standarda. Vođena je već navedenim direktorom Timonom Berners-Leejem i izvršnim direktorom Jeffreyem Jaffeom. Misija ove zajednice je voditi web do razvoja njegovog punog potencijala.

Na slici 1 prikazana je osnovna ideja semantičkog weba kako ju je osmislio Tim Berners-Lee.



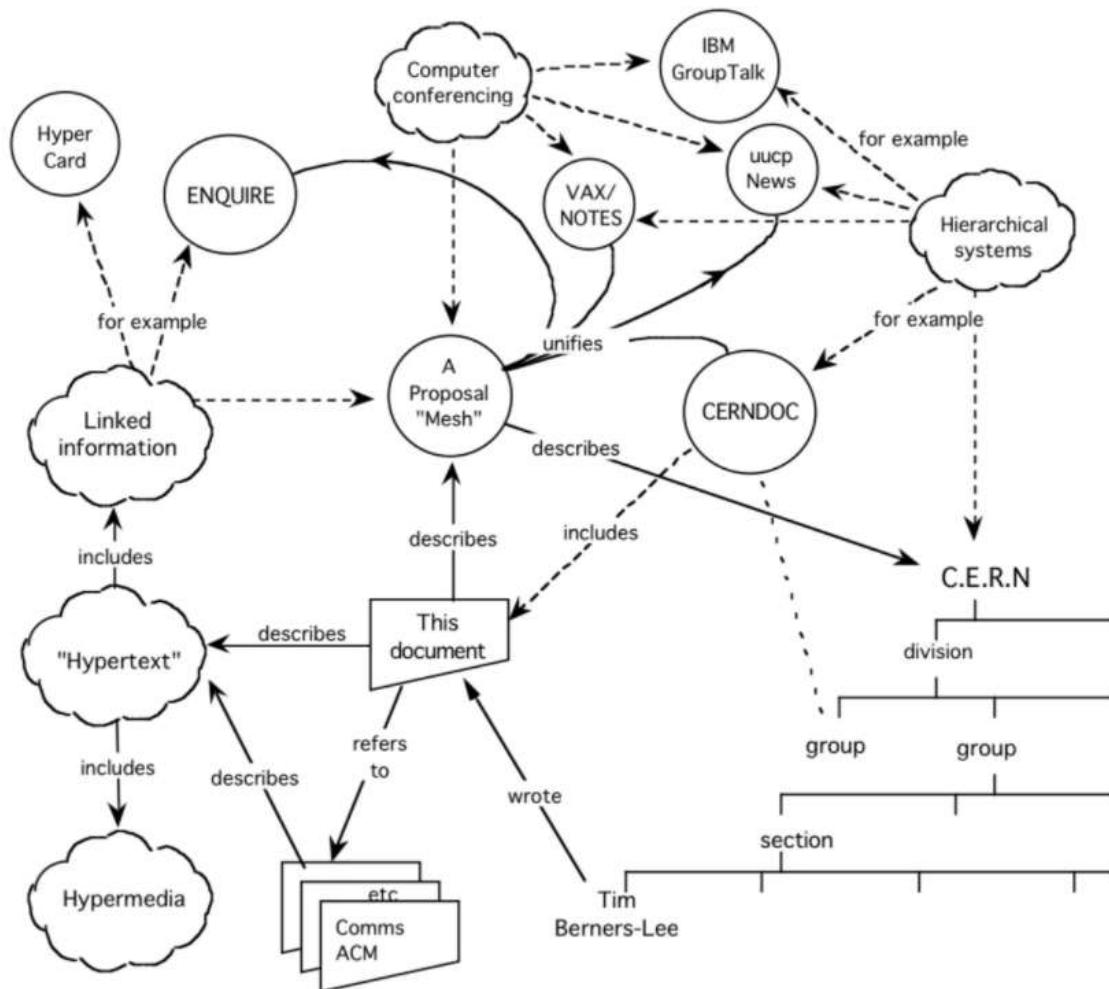
Slika 1. Fotografija sir Tim Berners-Lee-a 1990. godine

(Izvor:<https://www.vanityfair.com/news/2018/07/the-man-who-created-the-world-wide-web-has-some-regrets>, pristupljeno 27. 6. 2021.)

Berners-Lee je uvidio dvije ključne točke u razvoju ovog korisnog repozitorija za informacije potrebne za spremanje podataka, a to je omogućavanje metapodataka ili informacija o ostalim podacima koje se mogu pronaći na webu. Cilj je omogućiti „strojevima“ čitanje, razumijevanje i dodavanje vrijednosti na hiperveze odnosno *hyperlinkove* kako bi iste računala mogla koristiti za usmjeravanje pretraživanja. (Brooker, 2018) Dodavanje metapodataka na web-stranice se u početku smatralo previše zahtjevnim i korisniku s ograničenom procesorskom snagom se time nije izlazilo u susret u pogledu njegovih zahtjeva. Godine 2008. američka tvrtka za tražilice Yahoo! Inc. prihvaća ideju hijerarhijskog ustrojavanja web-stranica te se počinje isticati od ostalih preglednika po svojoj hijerarhijskoj strukturi pretraživanja.

Termin Web 1.0 opisuje rani razvoj World Wide Weba. Osobne internetske stranice su bile sveprisutne i većinski su uključivale statičke stranice održavane na besplatnim web-domenama ili web-serverima kontroliranim od strane internetskih poslužitelja (engl. *Internet service provider (ISP)*). Tijekom Weba 1.0 bilo je striktno regulirano i zabranjeno korištenje reklama (engl. *Ads*) dok Web 2.0. prikazuje trend koji se bazira na socijalizacijskoj noti, trend pod nazivom World Wide Web tehnologija (WWW). Web 2.0. korisnicima omogućuje kreiranje sadržaja weba; neke od njegovih karakteristika su otvorenost, sloboda i inteligencija te se mogu koristiti na svim raspoloživim internetskim platformama i uređajima te na svakom web-pregledniku.

Semantički web sljedeća je evolucijska stepenica razvoja World Wide Weba (WWW) te je u posljednje vrijeme prešao veliki put i postao najkorištenija usluga na internetskim preglednicima. Također, on predstavlja najpopularniji servis te način prikazivanja informacija koje su od velikih važnosti. On također predstavlja nadopunu WWW-a i omogućuje bolju suradnju između svakog računala i korisnika.



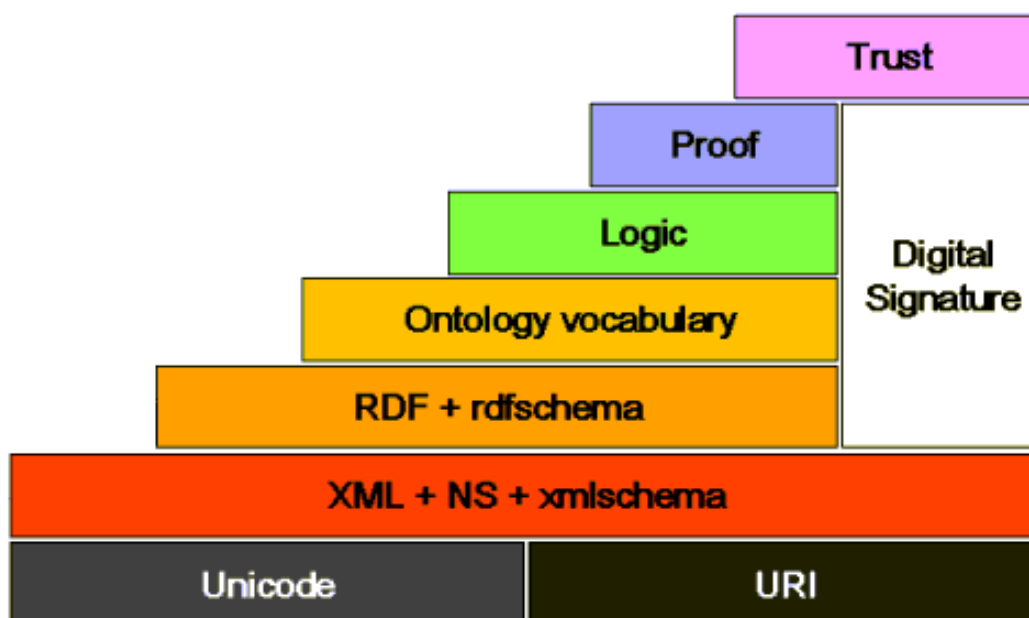
Slika 2. Originalni dijagram koji prikazuje kako je Sir Tim Berners-Lee predstavio World Wide Web 1989. godine

(Izvor: <https://www.smashingmagazine.com/2020/10/developing-semantic-web/>, pristupljeno 27. 6. 2021.)

4. SLOJEVI SEMANTIČKOG WEBA

U ovom poglavlju će se analizirati prezentacija semantičkog weba Tim Berners Leea te slojevi koje je predstavio. Razvoj se odvija u koracima. Svaki korak je jedan sloj koji se stavlja na drugi sloj. Principijelno opravdavanje za ovaj pristup je to da se ovime lakše postiže konsenzus minijaturnim koracima. Obično postoje nekoliko razvojno-istraživačkih skupina koje se kreću različitim pravcima te je ova ideja glavni pokretački pogon svakog znanstvenog napretka. Na slici 3 prikazan je slojeviti pristup semantičkom webu.

Unicode i *URI* su osnovni/početni slojevi koji služe za korištenje internacionalnog skupa znakova te predstavlja jedinstveno mjesto gdje je znakovni resurs nastao.



Slika 3. Slojeviti pristup semantičkom webu (Izvor: <https://www.w3.org/RDF/Metalog/docs/sw-easy>, pristupljeno 27. 6. 2021.)

Na sljedećem sloju nalazi se XML (engl. *Extensible Markup Language*) jezik koji omogućuje pojedincu pisanje strukturiranog dokumenta za web s prethodno definiranim korisničkim vokabularom. Djelomično se koristi za slanje dokumenata preko weba. Jednostavni model podataka se u ovoj shemi prikazuje u RDF sloju (engl. *Resource Description Framework*) odnosno hrvatski: Okvir za opis podataka). RDF prikazuje standardni model pomoću kojeg se razmjenjuju podatci na webu te sadrži značajke pomoću kojih se olakšava spajanje podataka bez obzira na razliku u shemi. On proširuje samu strukturu weba. RDF model podataka ne ovisi

o XML-u no ima XML baziranu sintaksu te je zbog toga u hijerarhiji stavljen iznad XML-a. RDF schema osigurava oblikovanje za hijerarhijsku organizaciju u web-objektima. Glavni oblikovni segmenti su svojstva i klase, podsvojstva i podklase te raspon i ograničenje domena. RDF schema se temelji na RDF-u te se shema može shvatiti kao jezik za pisanje ontologija. Prilikom hijerarhijske podjele na slojeve, Tim Berners-Lee je zamislio kompleksniju potrebu ontoloških jezika koji će ustvari u višim slojevima proširiti RDF shemu te zajamčiti reprezentaciju kompleksnijih veza/odnosa kod web-objekata. (Working group, 2013.) Logički sloj služi kako bi pojednostavio i nadogrudio ontološki jezik te dodatno dopustio i/ili omogućio pisanje aplikacijski baziranih načelnih znanja. Sloj dokaza točnije *Proof layer*, provodi provjeru istinitosti dokaza iz nižih slojeva te reprezentira informacije kroz procese, na kraju validacijom prikazuje dokaze. Na samome vrhu se nalazi sloj istinitosti odnosno povjerenja tzv. *Trust layer* koji dolazi do izražaja sa svojim „digitalnim potpisima“ koji služe kao validacija znanja koja je došla do zadnjeg, najvišeg sloja. Samim time što je ovaj proces na vrhu piramide, sloj istinitosti je ključni aspekt ove cjeline. Web će postići svoj puni značaj tek kada korisnik ima povjerenje u kvalitetu informacije koja mu je prezentirana. Iz perspektive inženjera postoji težnja standardizaciji, ako se dio istraživača ne slaže, odnosno slaže oko određenih problema i/ili pitanja utvrđuje se da ima smisla ispraviti dijelove dogovora. Prilikom slojevite izgradnje semantičkog weba trebaju se poštivati dva principa. Prvi princip koji se mora poštivati je takozvana „kompatibilnost prema dolje“. Osobe koje su u potpunosti svjesne i upoznate sa slojevima trebale bi znati interpretirati informacije napisane na nižim dijelovima sheme, odnosno ako je osoba svjesna semantike OWL (engl. *Web Ontology Language*) koja će kasnije biti razrađena u radu, osoba također mora biti svjesna te sposobna iskoristiti informacije napisane u RDF schema i RDF-u. Drugi princip koji se mora poštovati je „kompatibilnost prema gore“. Ovaj princip se bazira na tome da osobe koje su u potpunosti svjesne i upoznate sa slojem trebaju moći uzeti dio informacija s viših dijelova/slojeva, odnosno osoba koja je svjesna samo RDF i RDF scheme semantike potencijalno bi mogla iskoristiti informacije iz OWL-a (*Web Ontology Language*) djelomično, zanemarujući dijelove koji sežu iznad RDF i RDF scheme. (Hall i dr., 2009.)

4.1. Semantički preglednici

Semantički preglednici ili *hyperdata* preglednici su tipovi preglednika koji omogućavaju pretraživanje semantičkog weba, koji je ujedno ekstenzija World Wide Weba. Samim time semantičkim preglednicima može se pretraživati i regularan web koji zatim spaja različite podatke i izvore podataka kao regularan web-link tekst. Kako bi navedena teza bila moguća, možemo klasificirati semantičke preglednike na 3 tipa. Prvi tip je preglednik kao samostalna aplikacija, drugi tip je preglednik koji služi kao nadogradnja na već postojan web-preglednik te treći tip koji se može interpretirati kao preglednik koji se nalazi u web-pregledniku (najbolji primjer za treći tip bio bi web-stranica). Možemo ih podijeliti na dva načina prikaza podataka. Prvi tip prikazuje podatke u tekstualnom obliku. Ovi preglednici koriste tablične prikaze i liste kao prikaz jedinki, interakcija i obilježja. Suprotno prvom tipu koji prikazuje podatke u tekstualnom obliku su preglednici bazirani na vizualnoj prezentaciji. Ovi preglednici koriste slike, mape i grafičke prikaze kako bi vizualizirali veze između podataka. (EK Team, 2017.)

4.1.1. Merit

Merit je preglednik koji omogućava prikupljanje, pretraživanje i prikaz podataka koji su povezani informacijski RDF-om i Mikroformatom. Izrađena na Sindic i Yahoo BOSS pretraživačima koristi OKKAM web-poslužitelja i Any23 knjižnicu koju interpretira po komponentama koje razlaže na manje dijelove za lakši prijevod na druge jezike. Any23 (engl. *Anything To Triples*) prikazuje knjižnicu koja sadrži web-usluge i alat koji se izdvaja u strukturalne podatke RDF formata iz različitih web-dokumenata.

4.1.2. Disco

Disco je jednostavan *hyperdata* preglednik koji služi za navigaciju semantičkim webom. Može se koristiti bez prethodne instalacije na računalo zato što podatke povlači iz vlastitog servera. Preglednik daje korisniku mogućnost predaje svih informacija koje web-sjedište posjeduje na svojoj HTML stranici te potiče korisnika na istraživanje i posjećivanje povezanih pojedinih resursa (u *hyperdata* obliku). Filtrira svoja pretraživanja na temelju podataka koji su povezani statičnim RDF podacima negdje na webu ili RDF grafovima. (Kolb Peter, 2008.)

4.1.3. LodLive

LodLive naziv je projekta koji je nastao u svrhu korištenja podataka koji se koriste u svrhu pretraživanja i pregled RDF resursa. Program pretražuje podatke kroz RTF i SPARQL. Glavni princip dokaz je da resursi i podaci koji se nalaze u SPARQL standardu mogu biti lako dostupni. Ovime se pokušava natjerati i potaknuti vlasnike velikih količina podataka da dodaju svoje informacije u LodLive i podjele ih s ostatkom weba. SPARQL prikazuje semantički jezik koji se bavi upitima baze podataka te može upravljati svim podacima koji su pohranjeni u opisu svih resursa povezanih s RDF-om. Svojim korisnicima dopušta pisanje uputa, koje se mogu pisati prema podacima koji su povezani s RDF-om.

4.2. Semantičke tražilice weba

Značenje riječi semantičko pretraživanje nečega označava značenje sa smislom, potpuno drukčije od načina na koji pretraživači danas funkcioniraju. Danas pretraživači poput Googlea leksički pretražuju web te pokušavaju pronaći doslovno podudaranje riječi koju pretražujemo, ili barem dio te riječi, bez da, ustvari, preglednik stvarno shvati značenje upita koji mu je postavljen. Semantičke tražilice pokušavaju unaprijediti spomenuti problem na način da preciznost traženja usmjere k razumijevanju pretraživane namjere i kontekstualnog značenja relevantnih rezultata. Korisničkim pretraživanjem semantička tražilica fokusira se na korisničku namjeru i filtrira slične teme koje bi korisnika mogle zanimati u budućnosti. Semantičke tražilice weba možemo podijeliti na tražilice koje pokreće Microsoft Bing, Google, Metasearch pretraživače i pretraživače čija je web-stranica ujedno i pretraživač. (tablica 1.) Ovo je samo površna podjela, ukoliko se ista detaljizira, pretraživač ostvaruje još specifičnije filterske podjele kao na primjer: prema geografskoj lokaciji, poslovne pretraživače, pretraživače koji pretražuju samo *Dark Web*, edukacijske, pretraživače koji pretražuju samo događaje, hranu, poslove, zdravlje, novosti, ljude, nekretnine, televiziju, videoigre, mape, multimediju, e-mail, forume itd.

Ime pretraživača	Jezik pretraživača
Ask.com**	Višejezičan
Baidu	Kineski
Dogpile***	Engleski
DuckDuckGo*	Višejezičan
Ecosia*	Višejezičan
Exalead	Višejezičan
Excite*●	Višejezičan
Gigablast	Engleski
GoogleSearch	Višejezičan
HotBot*	Engleski
Lycos*	Višejezičan
MetaCrawler***	Engleski
Microsoft Bing	Višejezičan
Mojeek	Višejezičan
Qwant*	Višejezičan
Searx***	Višejezičan
Sogou	Kineski
Startpage.com**	Engleski
Swisscows*	Višejezičan
WebCrawler*	Engleski
Yahoo! Search*●	Višejezičan

YandexSearch	Višejezičan
Yippy***	Engleski

Tablica 1. Naziv i jezik na kojem se pretraživač bazira te način na koji tražilica funkcionira
(Izvor: vlastita izrada)

Legenda: * Tražilicu pokreće Microsoft Bing

** Tražilicu pokreće Google

*** Metasearch pretraživač

● Glavna web-stranica je ujedno i pretraživač

U nastavku će se dati kratak prikaz nekih od navedenih semantičkih tražilica.

4.2.1. DuckDuckGo

DuckDuckGo osnovao je Gabriell Weinberg u rujnu 2008. godine. Sjedište DuckDuckGo-a nalazi se u Pennsylvaniji u mjestu Paoli. DuckDuckGo je dostupan na većini pretraživača uključujući i najpopularnije pretraživače poput Google Chromea, Safarija i Firefoxa. DuckDuckGo ili DDG je pretraživač dizajniran za zaštitu korisničke privatnosti, glavna namjena mu je izbjegavanje „iskrivljenih“ pretraživanja koja se događaju zbog personaliziranih pretraga. Često se ova algoritamska pristranost koja iskrivljuje pretragu naziva i mjehur filtera (engl. *Filterbubble*). DuckDuckGo ne prati korisničku IP adresu te ostale informacije koje idu uz nju također nisu zabilježene na server. Za razliku od Googlea, DDG nema prilagođena pretraživanja poput lokacije na kojoj se korisnik nalazi, povijesti korisničkih pretraga i ostalih preferencija. Za pretraživanje DDG koristi svoj hibridni pretraživač koji povlači rezultate iz različitih izvora uključujući svoj program iliti svog „bota“ (engl. *crawler*) koji omogućava lakše indeksiranje podataka sa ostalih web sjedišta i pretraživača. Neki od partnera DDG-a su Yahoo!, Yelp, Yandex i Bing te ujedno od tih partnera DDG indeksiranjem povlači podatke s

navedenih pretraživača. Jedna od najčešće korištenih opcija koja se na DDG-u koristi je *!bang* opcija. Navedena opcija pretražuje samo podatke na stranici koju smo napisali poslije „!“ simbola. Može se pretraživati podatke samo na domeni npr. Twittera, upisujemo „!Twitter“ te pojam koji nas zanima. Ovime ograničavamo preglednik da pretražuje podatke koji se ne nalaze na toj stranici. (Gilmore, 2015.)

4.2.2. Mojeek

Mojeek je tražilica sa sjedištem u Ujedinjenom Kraljevstvu. Rezultati pretraživanja koje pruža Mojeek potječu iz vlastitog indeksa web-stranica, stvorenog pretraživanjem i indeksiranjem weba. Godine 2004., tražilica Mojeek započela je kao osobni projekt Marca Smitha u Sussex Innovation Centru. Tehnologija pretraživanja stvorena je od temelja koristeći uglavnom programski jezik C, a većim dijelom svog ranog života poslužitelji su se vodili iz Marcove spavaće sobe. Godine 2006., Mojeek je postao prva tražilica koja nema politiku privatnosti bez praćenja. Ova politika ostaje na snazi do danas. Nakon ulaganja, Mojeek je službeno registriran kao društvo s ograničenom odgovornošću 2009. Dana, 26. siječnja 2011., istaknuto je kao alternativna tražilica sa sjedištem u Velikoj Britaniji tijekom parlamentarne rasprave o internetskim tražilicama u Velikoj Britaniji zbog navoda o manipulaciji Googleovim rezultatima pretraživanja, osobito nepovoljan tretman njegovih neplaćenih i sponzoriranih rezultata te povlašteni plasman njihovih vlastitih usluga. Iste te godine premješta nove poslužitelje u podatkovne centre (po prvi put) te se novi indeks započinje „crawlati“ i izrađivati od nule. Nužno je napomenuti kako je 2015. godine postao jedina tražilica u Ujedinjenom Kraljevstvu i jedna od nekolicine u svijetu koja indeksira više od milijardu stranica. Kao novije zabilježeni podatci, Mojeek 2018. bilježi prelazak 2 milijarde stranica te je pokrenuto pretraživanje slika i tzv. *knowledge box*. Godine, 2019. pokrenut je model web-stranice dizajniran kako bi pomogao korisniku i pretraživaču navigaciju na stranici, kreiran je novi logo te je dodano još 100 web-poslužitelja. U 2021. godini Mojeek indeksira preko 4 milijarde stranica te doseže petnaest godina pružanja neovisne pretrage bez praćenja. Danas, Mojeek se može pohvaliti prelaskom 5 milijardi stranica, dodavanjem prijedloga pretraživanja te pokrenutom Mojeek zajednicom. (Mojeek, 2022.)

4.3. Ontologija

Jedna od najbitnijih komponenti, kako semantičkog weba tako i informativno poslovnih sustava, je ontologija. Ontologija sadržava koncepte, relacije i opis programa unutar neke hijerarhijske strukture u sustavu pojmova. Prilikom razvoja semantičkog weba posebno su važni standardi, programska rješenja te ontologijski rječnici. Ako su ove stavke ispunjene, web će se sastojati od ontologijskih opisa te će tim podacima komunicirati djelomičnim, ponekad i potpunim razumijevanjem korisničkih upita. Inženjeri weba služe se ontologijama prilikom registriranja karakteristika i/ili relacija promatranih objekata ili resursa. S obzirom na porast interesa za ontologijskim informacijama u znanosti, prikuplja se sve veća količina informacija koju treba obraditi. Drugim riječima, ontologije su domenski specificirani dijeljeni vokabulari koji pridodaju značenje podacima te ih čine fleksibilnijima. (Margaret Adolphus, 2009.) Ontologije ili taksonomije su opisane na sljedeći način: „shema koja formalno definira hijerarhije i odnose između različitih resursa. Semantičke mrežne ontologije sastoje se od taksonomije i skupa pravila zaključivanja iz kojih strojevi mogu donositi logične zaključke.“ (Altova, 2009.) Glavni tipovi opisa vokabularske terminologije su putem Web Ontology Languagea (OWL) koji nadodaje vokabular za opis svojstva i klase, i Simple Knowledge Organisation System (SKOS). Simple Knowledge Organisation System se koristi kako bi se dizajnirao sustav organizacije znanja i kako bi sama informacija imala jasniju primjenu. (O'Brien, 2020.)

4.4. URI

URI ili Universal Resource Identifier (negdje definiran i kao *Uniform Resource Identifier*) kako navodi HubSpot pruža jednostavan i proširivi način za identificiranje izvora a predstavlja kompaktni niz znakova koji razlikuje jedan resurs od drugoga. Navedena definicija može se objasniti kroz primjer; npr. `foo://example.com:8042/over/there?name=ferret#nose` je URI koji sadrži naziv sheme, ovlaštenje, putanju, upit i fragment. URI ne mora sadržavati sve te komponente. Sve što treba je naziv sheme i putanja datoteke, koja može biti prazna. Koriste se za identificiranje svega, od ljudi, koncepata, knjiga, stranica, mjesta itd. URI se može označavati na dva načina, identitetom i adresom. URL (*Uniform Resource Locator* ili web adresa) ili lokacija resursa označava gdje se točno resurs može pronaći na webu, no problem nastaje ukoliko se veza promjeni, tada resurs postaje nebitan. URI može biti URL no, URL ne

može biti URI. Vodile su se velike rasprave oko toga kakva je priroda URI-a, čak se postavljalo filozofsko pitanje koji je jedinstveni identifikator čovjeka te kako ga napraviti/nazvati. URI se može podijeliti na URN i URL. „URL (engl. *Uniform Resource Locator*) može se definirati kao niz znakova koji se odnosi na adresu. To je najraširenije korišten način lociranja resursa na webu. On osigurava metodu za dohvaćanje prezentacije fizičkog mjesta opisivanjem njegovog mrežnog mjesta ili primarnog mehanizma pristupa.“ (Gadget-info, 2019.) „URL je nepostojana vrsta URI-ja. Kasnija kategorija URI-ja je URN koja je postojana čak i kada resurs postane nedostupan. URN je potreban da postoji globalno jedinstven i ima globalni opseg.“ (Gadget-info, 2019.)

4.5. Linked Open Data

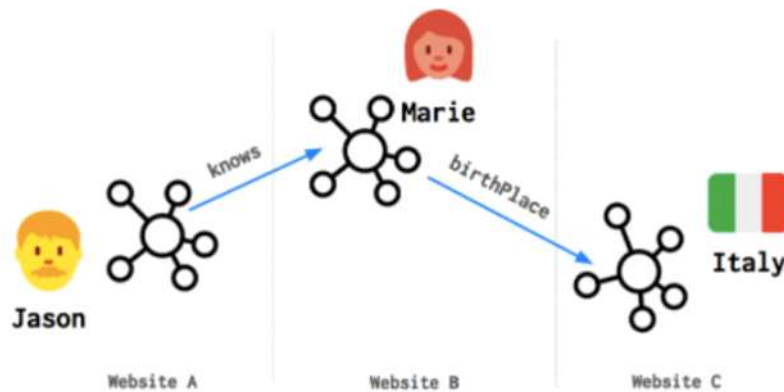
Prema Ogrizek – Biškupić i dr., kako bi mreža podataka postala stvarnost, potrebno je da velika količina podataka na internetu bude dostupna u standardnom formatu. Semantičkom webu nije potreban samo standardni pristup podacima, već je potreban odnos među podacima, kako bi se stvorila mreža istih. Takav skup podataka na webu nazivamo povezani podatci, odnosno *Linked Data*. Za funkcioniranje povezanih podataka, trebao bi postojati zajednički format, u ovom slučaju RDF, koji pretvara ili pruža trenutni pristup postojećim bazama podataka. Vrlo je bitno moći postaviti krajnje točke za lakši pristup tim podacima. Danas se nudi široka paleta različitih tehnologija za pristup podacima, a neke od njih su: RDF, PRAH, RIF, GRDDL i slične. *Linked Data* nalazi se u centru semantičkog weba. Skoro sve aplikacije temelje se na dostupnosti i povezanosti podataka na različitim razinama složenosti. Slika 4 predstavlja pojednostavljeni i razumljiviji prikaz povezanih podataka. *Linked Open Data* je put koji informacija mora prijeći kroz labirint podataka i informacija kako bi došla do odredišta.



Slika 4. Put podataka kroz labirint podataka

(Izvor: <https://www.ontotext.com/knowledgehub/fundamentals/what-is-the-semantic-web/>, pristupljeno 29. 6. 2021.)

Slika 5 predstavlja pojednostavljeni i razumljiviji prikaz povezanih podataka. Web-stranica A predstavlja Jasona i činjenicu da on poznaje Marie. Web-stranica B prikazuje informacije i Marie, a na web-stranici C prikazano je sve o njenom rodnom mjestu. Svaka od tih web-stranica sadrži strukturirane podatke koji opisuju određene entitete, u ovom slučaju (Jason, Marie i Italija) i vezu do entiteta koji je opisan na drugoj stranici.



Slika 5. Opis veze između entiteta

(Izvor: <https://wordlift.io/blog/en/entity/linked-data/>, pristupljeno 29. 6. 2021.)

4.6. OWL

Ontology Definition Language (OWL) u semantičkome webu definira ontologije. Kako bi ontologije mogle biti strojno čitljive, bilo je potrebno postaviti neke standarde i pronaći programska rješenja koja pružaju sustave kodiranja pogodne za strojnu obradu. Takva programska rješenja obično se nazivaju jezicima za prikaz ontologija i dijele se u tri kategorije: jezike temeljene na logici prvoga reda (npr. *Knowledge interchange format*), jezike temeljene na logici sustava okvira (npr. *Knowledge Language One* i *Frame Representational Language*) i mrežne jezike (HTML, XML, RDF...)" (Ogrizek - Biškupić, I., Banek - Zorica M., 2014.)

Trenutna verzija OWL-a, često se o njoj priča pod nazivom OWL 2 načinjena je od strane W3C OWL grupe (danas više grupa nije u pogonu) 2009. godine kao ekstenzija verzije iz 2004. pod nazivom OWL također razvijene od strane istog razvojnog tima.

4.7. RDF

Resource Description Framework (RDF) je standardni model za izmjenu podataka na webu. S naglaskom na olakšavanje spajanja podataka, čak i kad postoji razlika između osnovnih shema, nema potrebe da podatci koje korisnik koristi budu mijenjani. Proširenjem strukture weba koristi se URI za definiranje odnosa između resursa, kao i njihovih početaka i krajeva veza (najčešće definirano kao „trostruka veza“). Ovaj jednostavan, a opet složen model dozvoljava mogućnost miješanja strukturiranih i polustrukturiranih podataka te zatim njihovo izlaganje i dijeljenje s različitim aplikacijama. (World Wide Web Consortium, 2014.) Johnston navodi kako je središnje mjesto u RDF-u ideja izvora koji može biti bilo što se želi opisati - dokument, fizički objekt, osoba, imaginarno biće, koncept i ideja identificiranja resursa pomoću jedinstvenih identifikatora izvora (URI) (ili točnije URI reference). U RDF-u, URI reference su, jednostavno rečeno, imena za stvari.

4.8. XML

XML, odnosno *Extensible Markup Language*, jezik je sličan HTML-u, no razlika je da nema unaprijed definirane oznake za upotrebu, već se dizajniraju oznake koje su posebne za vlastite potrebe. XML je vrlo moćan način za pohranu podataka jer je u formatu koji se može pohraniti, dijeliti i pretraživati. S obzirom na to da je XML standardiziran primatelj i dalje može raščlaniti podatke zbog standardizirane XML sintakse. XML je razvila radna skupina koja je izvorno poznata kao SGML urednički odbor za pregled te koja je osnovana pod pokroviteljstvom World Wide Weba 1996. godine, navodi Knox i dr. XML je profil aplikacije ili ograničeni oblik standardnog generaliziranog jezika za označavanje te su po svojoj konstrukciji dokumenti XML u skladu sa SGML dokumentima. XML dokumenti se sastoje od jedinica za pohranu, odnosno entitetima te sadrže raščlanjene i neraščlanjene podatke. Raščlanjeni podatci sastoje se od znakova, od kojih neki označavaju, a neki od njih čine podatke o znakovima. Postoje brojne prednosti korištenja XML-a. Koristi ljudski, a ne računalni jezik, vrlo je čitljiv i razumljiv te je lakši za kodiranje od HTML-a. XML može koristiti vaše podatke, bez obzira na

platformu jer ga bilo koja aplikacija može obraditi. Također, XML se može proširiti te je potpuno prenosiv. Slika 6. prikazuje razliku između primjera HTML-a i XML-a. Može se primijetiti kako imena HTML oznaka ne otkrivaju ništa o njihovom značenju, kratice nisu čitljive, teža je razmjena podataka, dok su XML oznake čitljivije i razumljivije te je struktura podataka korisnija i olakšava razmjenu podataka.



Slika 6. Usporedba koda HTML-a i koda XML-a (Izvor:

<https://www.ibm.com/docs/en/i/7.1?topic=introduction-advantages-xml>, pristupljeno

30. 6. 2021.)

4.9. SPARQL

SPARQL, kao skup standarda za grafičku bazu podataka koju je objavio W3C, a naziv najčešće korišten za označavanje jezika upita, omogućava korisniku ispitivanje svih baza podataka koje mogu biti povezane s RDF-om. Standardni upitni jezik i protokol za gore naveden *Linked Open Data* i RDF baze podataka. Glavna namjena mu je ispitivanje raznih vrsta podataka a ujedno omogućava efektivno odvajanje informacija „skrivenih“ u nejednolikim podacima koji se spremaju u raznim formatima. (Ontotext, 2019.) S fokusom na dizajn i podršku, W3C (World Wide Web Consortium) putem SPARQL-a pomaže korisnicima skrenuti fokus na ono što

korisnik želi znati, a ne ono kako je baza podataka organizirana i strukturirana. Primjer na slici 7 pojašnjava način na koji SPARQL postavlja upite bazi podataka s kojom raspolaže. Varijable se označavaju prefiksom simbola „?x/y“. Upiti za „?capital“ i „?country“ bit će vraćeni korisniku u obliku teksta.

```
PREFIX ex: <http://example.com/exampleOntology#>
SELECT ?capital
       ?country
WHERE
{
  ?x ex:cityname      ?capital ;
     ex:isCapitalOf  ?y .
  ?y ex:countryname  ?country ;
     ex:isInContinent ex:Africa .
}
```

Slika 7. Primjer SPARQL upitnog koda (Izvor: <https://en.wikipedia.org/wiki/SPARQL>, pristupljeno 30. 6. 2021.)

5. IMPLEMENTACIJA SEMANTIČKOG WEBA NA STVARNOM PRIMJERU

Korporacija Chevron, koja se više od stotinu godina bavi bušenjem i izvlačenjem nafte iz zemlje, implementirala je semantičke resurse u svoje poslovanje kako bi minimizirala rizik svakodnevnog posla. Danas im je neusporedivo jednostavnije poslovati i izvlačiti sirovine zbog ogromnih količina podataka i resursa koje im semantički web pruža. Naftne platforme koje Chevron ima po cijelom svijetu (većinu kod Meksika) danas rade brže, efikasnije i sigurnije

zbog terabajta podataka koje prikupljaju svaki dan. Ti podaci im pružaju važne informacije koje, kada se pravilno iščitaju, mogu spriječiti ne planirane strojna zakazivanja i anomalije. Svakoga dana kada se greška previdi, korporacija gubi stotine milijuna dolara. Unazad nekoliko godina postojao je konstantan rizik od eksplozija (i danas se događaju eksplozije, no u puno manjem broju), no tada nisu samo u pitanju stotine milijuna dolara nego i ljudski životi. Semantički web danas inženjerima olakšava predviđanje operacija na polju naftne industrije te je njegova primjena u tom segmentu široka. Svakoga se dana generira velika količina podataka iz raznih izvora, seizmičkim mjerenjima, broj bušenja, broj transporta, zapisi o bušenjima, volatilitet dionica i marketinška statistika samo su neki od podataka koji svakodnevno ova firma filtrira radi unaprjeđenja produktivnosti. (Wynings, D., 2017.)

Konstrukcija semantičkog weba ovisi o uspostavi nove razine interoperabilnosti koja se temelji na otvorenim standardima, a osim definicije kvalitete obrasca sintakse dokumenta trebala bi pružiti i osnovu za sintaktički sadržaj. Semantički web temelji se na ideji korištenja znanja, načela i tehnologija. Ta znanja, principi i tehnologije osnova su običnih mreža. Kada se koriste u novim mrežama, postat će univerzalni medij za razmjenu podataka, informacija i znanja. Stoga je semantički web oblikovan u hijerarhijsku strukturu temeljenu na korištenju prethodno razvijene informacijske tehnologije. Mrežnom protokolu, kao postojećoj mrežnoj infrastrukturi, dodani su sljedeći protokoli: sintaksa podataka (XML/XML shema), semantički mrežni model podataka (RDF), definicija RDF rječnika (RDF shema), ontologija (OWL) te logika i mehanizmi za osiguravanje povjerenja (logika, dokaz i povjerenje). Logički sloj podržava pisanje pravila, dokazni sloj ih izvršava i ocjenjuje te radi sa slojem povjerenja kako bi ponovno provjerio povjerenje u točnost podataka. Da bi se dokument pretvorio u XML format, moraju biti ispunjeni sljedeći uvjeti: mora se koristiti DTD (ili XML shema), sve vrijednosti atributa moraju se navoditi, sve oznake moraju imati početne i završne elemente, osim ako je određena oznaka prazan element, ako je oznaka neovisna trebala bi imati simbol zatvaranja prije kraja oznake, sve oznake trebale bi biti ispravno ugniježdene, svaki XML dokument bez DTD-a trebao bi imati atribut tipa CDATA, a određene elemente treba navoditi. Inače, kako se ne bi uništila struktura dokumenta, a može postojati samo izvorni element. XML je osmišljen kao univerzalna oznaka (opisni jezik) koja je neovisna o svom konačnom predstavljanju za opisivanje i stvaranje strukture dokumenta. Karakterizira ga mogućnost prikazivanja istih informacija na različite načine, s razlikama u obliku ili suštini. Čak i ako se usvoji dogovorena struktura podataka, aplikacija može naići na dokumente koji koriste različite strukture verzija pa XML koristi svoj mehanizam za jasno navođenje strukture u dokumentu.

XML dokument sastoji se od niza oznaka o izvoru, pod uvjetom da svaka oznaka (ugniježđeni element), može imati bilo koji broj atributa. XML dokumenti mogu predstaviti višestruke strukture na više načina. Izgled dokumenta s „.xml“ ekstenzijom može se opisati kao stablo oznaka gdje je svaka oznaka pridružena imenovanom čvoru u modelu datoteke, a svaka ugniježđena oznaka predstavlja nasljednika oznake. Pogodan je za definiranje strukture dokumenta, ali ne i za definiranje njegova značenja. Najčešće se koristi za stvaranje novih vlasničkih jezika za označavanje, jezika za semantičko označavanje za web-stranice, mijenjanje jedinstvenog formata podataka i odvajanje podataka od prezentacija. XML kao format koristi se samo za definiranje strukturnih resursa podataka, a ne i njegovog značenja. Nesposobnost da se podrži bilo kakvo tumačenje podataka sadržanih u dokumentu njegov je glavni nedostatak. (Knox, R. i dr., 2006.)

6. RASPRAVA

Postoji bogatstvo znanja u zajednici semantičkog weba i postavlja se pitanje na koji način pristupiti tom znanju i tim podacima u svrhu razumijevanja postojećeg znanja i kreiranja novog znanja. Mogućnosti koje semantički web pruža su neograničene. U ovom radu nije toliko bilo detaljizirano kojim nečinima, nego o omogućavanju tih načina. Razni označivački jezici poput RDF-a, OWL-a, Schema itd. su pokušali kroz godine postati standard, no skala koju su postavili HTML i CSS jednostavno je bila previsoka te se semantički web i dalje koristi u jako malom broju. Iz godine u godinu sve više stranica počinje s implementacijom semantičkog weba na svoja web-sjedišta te se samim time ojačava pritisak i konkurenciju na obični web.

Semantički web ima veliki potencijal za koegzistenciju jer ima neosporne prednosti, no za potpuni prelazak na semantički web valja napomenuti da se ulaže mnogo truda u razvoj tehnologija koje će automatski izdvajati značenje iz nesemantičkog teksta pa će zbog toga prelazak na semantički web postati znatno lakši.

7. ZAKLJUČAK

Ono što danas semantički web treba je faza konsolidacije u svim primarnim i sekundarnim poljima koje su usmjerene na dokumentiranje svojih ciljeva, prednosti i nedostatka. Možemo za primjer uzeti popularni uređivački ontološki editor Protégé, OWL API uređivač Wikibase baza podataka koji za zadatak ima pokretati Wikidata ili ELK reasoner. Koliko god da ovaj softver ima potencijala, unatoč tome što svi koriste RDF i OWL za serijalizaciju, odnosno proces koji je zadužen za spremanje struktura podataka u čitljivi format, još će mnogo vremena proći kako bi jednostavno mogao raditi s ostatkom softverskih inačica koje su za to programirane. Tko bi mogao utjecati na ubrzavanje ovog procesa? U današnjem svijetu postoji nažalost izričito nizak financijski izdatak za razvojem i održavanjem softvera koji će biti stabilan i jednostavan za upotrebu. Tim Berners-Lee izjavio je: „Semantički web je kultura koliko je i tehnička prepirka“. Ta rečenica se danas pokazuje točnijom nego ikada prije. Web i povezanost na njemu će biti onoliko dobar koliko je u njega uloženo. U, za sada, dalekoj budućnosti ljudi ciljaju na neku novu vrstu weba u kojoj će se semantički potpuno integrirati s običnim webom. Ako se to ikada dogodi očekuje se da će strojevi moći procesuirati informacije na višem nivou te komunicirati među sobom bez ljudskih „ručnih“ kodiranja.

8. LITERATURA

1. Adolphus, M. (2009). : *The Semantic Web – a new tool for libraries?* Dostupno na: https://www.emeraldgroupublishing.com/archived/librarians/info/viewpoints/semantic_web.htm [pristupljeno 28. 6. 2021.]
2. Altova. (2009). : *What is the Semantic Web.* Dostupno na: http://www.altova.com/semantic_web.html [pristupljeno 28. 6. 2021].
3. Brooker, K. (2018). *Tim Berners-Lee*, Dostupno na : <https://www.vanityfair.com/news/2018/07/the-man-who-created-the-world-wide-web-has-some-regrets> [pristupljeno: 28.6.2021]
4. EK TEAM (2017). *Semantic Web Starter Kit.* Dostupno na : <https://enterprise-knowledge.com/semantic-web-starter-kit/> [pristupljeno: 26. 6. 2021]

5. Gadget-info. (2019). Difference between url and uri. Dostupno na: <https://hr.gadget-info.com/difference-between-url> [pristupljeno 1. 8. 2022.]
6. Gilmore, I. (2015). *DuckDuckGo*. Dostupno na : [https://whatis.techtarget.com/definition/DuckDuckGo#:~:text=DuckDuckGo%20\(DG\)%20is%20a%20general,other%20information%20are%20not%20logged](https://whatis.techtarget.com/definition/DuckDuckGo#:~:text=DuckDuckGo%20(DG)%20is%20a%20general,other%20information%20are%20not%20logged). [pristupljeno 30. 6. 2021.]
7. Hall, W. ; O' Hara, K. (2009). *Semantic Web*. Dostupno na : https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-30440-3_478 [pristupljeno 26. 6. 2021.]
8. HubSpot (2021). URI vs. URL: What's the Difference? Dostupno na: <https://blog.hubspot.com/website/uri-vs-url> [pristupljeno 1. 8. 2022.]
9. Johnston, P. (2005.) XML, RDF, and DCAPs. Dostupno na: <http://www.ukoln.ac.uk/metadata/dcmi/dc-elem-prop/#sec3> [pristupljeno 1. 8. 2022.]
10. Knox, R.; Fredman, T. ; Thompson, J. (2006). Sharing Semantics Across Applications, Dostupno na: <https://www.w3.org/2001/sw/sweo/public/UseCases/Chevron/> [pristupljeno 9. 8. 2022.]
11. Mojeek (2022). About Mojeek- building the world's aternative search engine. Dostupno na: <https://www.mojeek.com> [pristupljeno 2. 8. 2022.]
12. O' Brien, F. (2020). *Developing For The Semantic Web*. Dostupno na : <https://www.smashingmagazine.com/2020/10/developing-semantic-web/> [pristupljeno 1. 7. 2021.]
13. Ogrizek - Biškupić, I.; Banek - Zorica, M. (2014). *Web tehnologije*, Zaprešić : Visoka škola za poslovanje i upravljanje s pravom javnosti Baltazar Adam Krčelić
14. SPARQL Working Group (2013). *SPARQL QueryLanguage for RDF*. Dostupno na: <https://www.w3.org/2001/sw/wiki/SPARQL> [pristupljeno 1. 7. 2021]
15. World Wide Web Consortium (2014.) Resourde Desription Framework. Dostupno na: <https://www.w3.org/RDF/> [pristupljeno 2. 8. 2022.]
16. Wynings, D. (2017). *RIP: The Semantic Web*. Dostupno na: <https://blog.diffbot.com/rip-the-semantic-web/> [pristupljeno 27. 6. 2021.]

9. POPIS SLIKA

Slika 1 - Fotografija sir Tim Berners-Leea 1990. godine	4
Slika 2 - Originalni dijagram koji prikazuje kako je Sir Tim Berners-Lee predstavio World Wide Web 1989. godine	6
Slika 3 - Slojeviti pristup Semantičkom webu	7
Slika 4 - Put podataka kroz labirint podataka	15
Slika 5 - Opis veze između entiteta.....	15
Slika 6 - Usporedba koda HTML-a i koda XML-a.....	17
Slika 7 - Primjer SPARQL upitnog koda.....	18

10. POPIS TABLICA

[Tablica 1](#) - prikazuje jezik na kojem se pretraživač bazira, te način na koji tražilica funkcionira (vlastita izrada)..... 11