

Semantičke grafičke baze podataka: izrada grafa znanja

Mataija, Katarina

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Humanities and Social Sciences / Sveučilište Josipa Jurja Strossmayera u Osijeku, Filozofski fakultet**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:142:297101>

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

Download date / Datum preuzimanja: **2024-11-26**



Repository / Repozitorij:

[FFOS-repository - Repository of the Faculty of Humanities and Social Sciences Osijek](#)



Sveučilište J.J. Strossmayera u Osijeku

Filozofski fakultet

Dvopredmetni diplomski studij informatologije i informacijske tehnologije

Katarina Mataija

Semantičke grafičke baze podataka: izrada grafa znanja

Diplomski rad

Mentor: prof. dr. sc. Boris Bosančić

Osijek, 2024.

Sveučilište J.J. Strossmayera u Osijeku
Filozofski fakultet Osijek
Odsjek za informacijske znanosti
Dvopredmetni diplomski studij informatologije i informacijske tehnologije

Katarina Mataija

Semantičke grafičke baze podataka: izrada grafa znanja

Diplomski rad

Društvene znanosti, informacijske znanosti, informacijski
sustavi i informatologija

Mentor: prof. dr. sc. Boris Bosančić

Osijek, 2024.

Sadržaj

Sažetak	4
1. Uvod.....	1
2. Semantički web i povezani podaci.....	2
2. 1. Slojevi semantičkog weba.....	2
2. 2. Povezani podaci	5
2. 3. Tehnologije semantičkoga weba i povezanih podataka.....	7
3. Računalne ontologije	10
4. Semantičke grafičke baze podataka i grafovi znanja.....	13
4.1. Relacijske vs. grafičke baze podataka	13
4.2. Semantičke grafičke baze podataka	15
4.3. Grafovi znanja.....	15
5. Primjer izrade grafa znanja	18
5.1. Metodologija	18
5.2. Izrada skupa povezanih podataka	18
5.3. Izrada ontologije	23
5.4. Izrada grafa znanja.....	27
6. Zaključak.....	35
Popis literature	36

Sažetak

Ovaj rad bavi se temom semantičkih grafičkih baza podataka, s posebnim naglaskom na izradu grafa znanja. Cilj rada je predstaviti skup povezanih podataka, ontologiju i semantičke grafičke baze podataka, te izraditi prototip grafa znanja na temelju tih elemenata. U radu se detaljno razmatraju koncepti semantičkog weba, povezanih podataka, ontologija i tehnologija koje omogućuju integraciju i organizaciju informacija, čime se postiže njihovo bolje razumijevanje i korištenje. Posebna pažnja posvećena je izradi grafa znanja, temeljenog na skupu povezanih podataka i izgradnji ontologije. Praktični dio rada prikazuje postupak izrade grafa znanja kroz nekoliko faza, uključujući definiranje i kreiranje skupa povezanih podataka, kreiranje ontologije te integraciju podataka u semantičku grafičku bazu podataka. Ovaj rad pruža uvid u važnost i potencijal grafova znanja u organizaciji i analizi informacija, kao i ulogu semantičkih grafičkih baza podataka u modernim informatičkim sustavima.

Ključne riječi: semantički web, semantičke grafičke baze podataka, povezani podaci, ontologije, grafovi znanja.

1. Uvod

U suvremenom digitalnom dobu, količina informacija dostupnih na internetu neprestano raste, što pretraživanje, organizaciju i razumijevanje tih podataka čini sve složenijim zadatkom. Klasični pristupi organizaciji podataka, poput relacijskih baza podataka, često se suočavaju s ograničenjima u pogledu fleksibilnosti i semantičkog razumijevanja podataka. Kao odgovor na te izazove, razvijene su semantičke grafičke baze podataka koje omogućuju dublje i smislenije povezivanje informacija. Ove baze podataka pružaju okvir za integraciju raznovrsnih izvora podataka, omogućujući njihovu lakšu analizu i korištenje.

Semantički web, kao proširenje postojećeg World Wide Weba, omogućava bolje razumijevanje podataka putem jasnih i preciznih opisa resursa. Ključnu ulogu u ovom procesu imaju grafovi znanja koji strukturiraju podatke na način koji omogućuje njihovo smisleno pretraživanje i povezivanje. Svrha ovoga rada je istražiti koncept semantičkih grafičkih baza podataka i izraditi prototip grafa znanja koji ilustrira praktičnu primjenu ovih tehnologija. Ciljevi rada uključuju kratki prikaz semantičkoga weba, povezanih podataka, ontologija i semantičkih grafičkih baza podataka te izradu prototipa grafa znanja na temelju navedenog.

Rad je podijeljen u nekoliko ključnih poglavlja. Nakon uvodnog dijela, drugo poglavlje pruža pregled semantičkog weba, uključujući njegove slojeve, povezane podatke te tehnologije koje omogućuju njihovu implementaciju. Treće poglavlje bavi se konceptom računalne ontologije, dok četvrto analizira različite tipove baza podataka, s naglaskom na relacijske, grafičke i semantičke grafičke baze podataka. U ovom poglavlju dan je osvrt i na koncept grafa znanja, dok peto poglavlje donosi prikaz praktične izrade grafa znanja, uključujući primijenjenu metodologiju, izradu skupa povezanih podataka, izradu ontologije i konačnu implementaciju grafa. Na koncu, u samome zaključku sažeti su glavni nalazi i preporuke za buduća istraživanja.

2. Semantički web i povezani podaci

Semantički web inspiriran je vizijom trenutnog weba koja postoji od njegovog nastanka, a na njega je utjecao rad Vannevara Busha iz 40-ih godina prošlog stoljeća, u kojem izlaže ideju o *memex* stroju. Tim Berners-Lee prvotno je zamislio da *World Wide Web* uključuje bogatije opise dokumenata i veze među njima. Međutim, u nastojanju da se osigura jednostavan, upotrebljiv i robustan radni sustav koji bi svi mogli koristiti „izvan kutije”, te su ideje ostavljene po strani, a rezultat je jednostavniji web prilagođen i za one bez tehničkog znanja a koji danas poznajemo.¹

Naprednija vizija pronašla je svoj izraz u članku koji su napisali Tim Berners-Lee, Jim Hendler i Ora Lassila 2001. godine. U ovom članku oni pružaju uvjerljivu viziju weba na kojem sam web, umjesto ljudi, može obavljati zadatke poput pronalaženja dokumenata, lociranja usluga ili obavljanja rutinskih poslova, poput zakazivanja termina kod liječnika i slično. To se može postići pružanjem dostatnog konteksta o resursima na webu te pružanjem alata za korištenje tih resursa kako bi strojevi mogli pronaći prave i točne izvore informacija i donositi odluke. Kao što su i sami autori naveli u članku: „... semantički web dovest će strukturu do smislenog sadržaja mrežnih stranica, stvarajući okruženje u kojem softverski agenti koji 'lutaju' od stranice do stranice mogu brzo obavljati sofisticirane zadatke za korisnike“.² To je ambiciozan dugoročni cilj: ništa manje od davanja smisla samome webu. To je drugi način pružanja smislenih načina za opisivanje resursa dostupnih na internetu i, možda još važnije, odgovor na pitanje zašto postoje poveznice koje ih povezuju. Jednom kada web bude imao mehanizam za definiranje semantike o resursima i vezama, tada će nastati mogućnost za automatsku obradu weba koju provode softverski agenti bez potrebe za ljudskim posredovanjem.³ U istom članku semantički web autori definiraju kao: „... proširenje postojećega weba u kojem se informacijama daje jasno definirano značenje, čime se računalima i ljudima omogućuje učinkovitija suradnja.“⁴

2. 1. Slojevi semantičkog weba

The Open Systems Interconnection model (OSI model) produkt je napora *Open Systems Interconnection*, međunarodne organizacije za standardizaciju. To je način dijeljenja

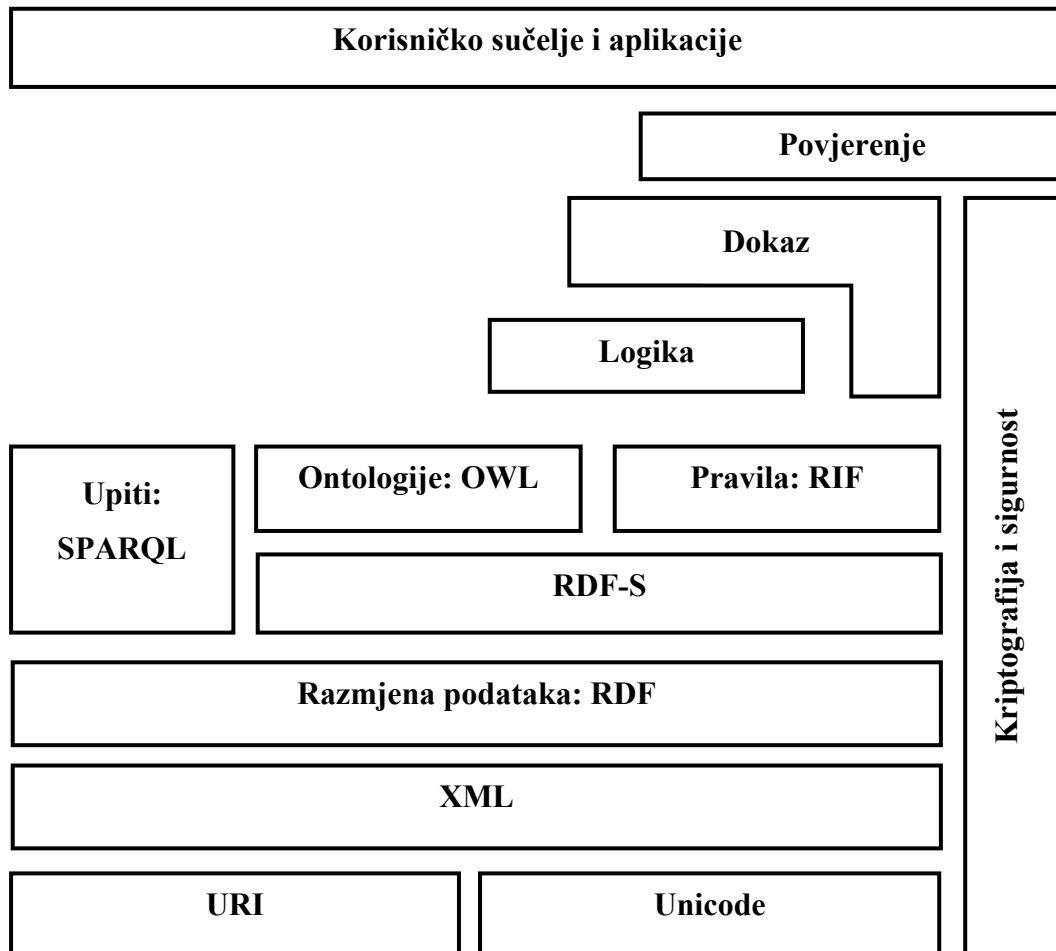
¹ Usp. Matthews, Brian. *Semantic Web Technologies*. // *E-learning and Digital Media* 6, 6(2005), str. 2. URL: https://www.researchgate.net/publication/30408878_Semantic_Web_Technologies (2024-07-04)

² Usp. Berners-Lee, T., Hendler, J., Lassila, O., et al. *The semantic web*. // *Scientific American* 284, 5(2001), str. 28–37. URL: 10.1038/scientificamerican052001-yL7Vw7HIOZ4iSjlnEeVsJ (2024-07-03)

³ Usp. Matthews, Brian. *Nav. dj.*, str. 3.

⁴ Usp. Berners-Lee, T., Hendler, J., Lassila, O., et al. *Nav. dj.*

komunikacijskog sustava na manje dijelove koji se nazivaju slojevi. Sloj je skup konceptualno sličnih funkcija koje pružaju usluge sloju iznad i primaju usluge od sloja ispod njega. Ovaj model široko se koristi u dizajniranju mrežnih arhitektura na globalnoj razini. Model započinje fizičkim slojem, a završava aplikacijskim slojem, koji uključuje protokole poput HTTP-a. Tim Berners-Lee osmislio je sličnu konceptualnu strukturu semantičkoga weba pod nazivom „*The Semantic Web Layer Cake*“.⁵



Slika 1 – Prikaz slojeva semantičkoga weba.

Najniži sloj odgovoran je za kodiranje znakova na bilo kojem jeziku te za jedinstveno prepoznavanje različitih resursa. Ovaj sloj je standardizirao tekst koristeći Unicode i URI/IRI (*Uniform Resource Identifier/Internationalized Resource Identifier*). Unicode služi za

⁵ Usp. J. Domingue, D. Fensel, J. A. Hendler, Handbook of Semantic Web Technologies. Berlin, Germany: Springer, 2011. Str. 18 URL: <https://link.springer.com/referencework/10.1007/978-3-540-92913-0> (2024-07-04)

predstavljanje i rukovanje tekstem u računalnom okruženju na mnogim jezicima. Semantički web omogućuje povezivanje dokumenata na različitim ljudskim jezicima.⁶

Drugi sloj obilježava XML koji se koristi za postizanje interoperabilnosti na webu. Ovaj sloj pruža osnovu za miješanje različitih elemenata iz različitih vokabulara za opis resursa. XML pruža elementarnu sintaksu za strukturu sadržaja unutar dokumenata, ali ne pridružuje semantiku značenju sadržaja koji je unutar njega. No, kako će se uskoro pokazati, XML nije nužna komponenta tehnologija semantičkog weba, jer se u većini slučajeva koriste alternativne sintakse poput Turtlea. Turtle je tako postao standard koji se široko primjenjuje u praksi, a da nije prošao formalni proces standardizacije.⁷

Prva dva sloja obuhvaćaju osnovne web tehnologije, dok se od trećega sloja prikazuju tehnologije semantičkog weba. *Resource Description Framework* (RDF) je W3C standard za opis mrežnih resursa. U početku, RDF je rabio XML sintaksu te je dizajniran tako da ga računala čitaju i razumiju, a ne za prezentiranje ljudima. RDF je jednostavan podatkovni model koji izražava odnose između objekata („resursa“). RDF model podataka može se predstaviti u raznim sintaksama, poput RDF/XML, N3, Turtle itd. RDF je, stoga, osnovni standard semantičkog weba te ga preporučuje i W3C (*World Wide Web Consortium*). *RDF Schema* (RDFS) proširuje RDF te predstavlja vokabular za opisivanje svojstava i klasa resursa temeljenih na RDF-u, sa semantikom koja omogućuje hijerarhiju istih.⁸

Nadalje, četvrti sloj obuhvaća ontologije koje se bave točnim opisom stvari i njihovih odnosa. Ponekad RDFS definicija pojedinog web resursa nije dovoljna, pa je potreban opsežniji ontološki rječnik. *Web Ontology Language* (OWL) jezik je za opis mrežnih informacija te također W3C standard. Naime, OWL predstavlja robusniji vokabular za opis klasa i svojstava, između ostalog i odnosa između klasa, kao što je primjerice nepovezanost, kardinalost, jednakost, simetriju itd. RDF i OWL u velikoj mjeri su vrlo slični, no OWL prednjači s većom strojnom interpretacijom s većim rječnikom i snažnijom sintaksom od RDF Scheme.⁹ Također, ovome sloju pripada *Simple Protocol and RDF Query Language* (SPARQL) te *Rule*

⁶ Usp. Matthews, Brian. *Semantic Web Technologies*. // *E-learning and Digital Media* 6, 6(2005), str. 6. URL: https://www.researchgate.net/publication/30408878_Semantic_Web_Technologies (2024-07-04)

⁷ Isto.

⁸ Usp. Gopal, Pandey. *The Semantic Web: An Introduction and Issues*// *International Journal of Engineering Research and Applications*, 1(2012), str. 780. URL: https://www.academia.edu/1959581/IJERA_www_ijera_com_ (2024-07-04)

⁹ Usp. Wilson. M.; Matthews, B. *The semantic Web: prospects and challenges*. // *7th International Baltic Conference on Databases and Information Systems*, 2006, str. 28. URL: 10.1109/DBIS.2006.1678469. (2024-07-21)

Interchange Format (RIF). Kriptografija je u OSI modelu važna za osiguranje i provjeru autentičnosti informacija te za potvrdu da one dolaze iz pouzdanog izvora, što se može postići odgovarajućim digitalnim potpisom RDF izjava.¹⁰

Peti sloj naziva se i logičkim slojem (engl. *Logic*) koji je potreban za izražavanje pravila te se ovaj sloj koristi za dodatno obogaćivanje ontoloških jezika.¹¹

Nadalje, šesti sloj naziva se i dokaznim slojem (engl. *Proof*) koji uključuje stvarni proces dedukcije, kao i prikaz dokaza na web jezicima. Primjerice, kada netko pošalje mrežnoj stranici X dokaz da je ovlašten koristiti lokaciju, stranica X mora biti u mogućnosti provjeriti taj dokaz. Ovaj proces se odvija pomoću odgovarajućeg mehanizma za zaključivanje.¹²

Sedmi sloj naziva se i sloj povjerenja (engl. *Trust*) u kojem se koriste digitalni potpisi. Primjerice, tvrtka X pošalje informacije tvrtki Y, ali ne postoji način na koji bi tvrtka Y mogla biti sigurna da te informacije uistinu dolaze od tvrtke X. U takvom slučaju, zaključuje se da se tvrtki X ne može vjerovati te jedino preostaje odbaciti te informacije. Povjerenje u informacije gradi se tako da se osigura dokaz kako one dolaze iz pouzdanih izvora i da se koristi jasna logika za pronalaženje novih informacija.¹³

Posljednji sloj sastoji se od korisničkoga sučelja (engl. *UI – User Interface*) i aplikacija koji omogućuju pristup aplikacijama semantičkoga weba.¹⁴

2. 2. Povezani podaci

Posljednjih godina web se razvio iz globalnog informacijskog prostora povezanih dokumenata u prostor u kojem su i dokumenti i podaci povezani. Temelj ove evolucije je skup najboljih praksi za objavljivanje i povezivanje strukturiranih podataka na webu, poznat kao koncept povezanih podataka (engl. *linked data*). Uvođenje najboljih praksi u pogledu povezanih podataka dovelo je do proširenja weba s globalnim podatkovnim prostorom koji povezuje podatke iz različitih područja, kao što su poslovanje, knjige, znanstvene publikacije, filmovi i slično. Postoje generički preglednici povezanih podataka koji korisnicima omogućuju da započnu pregledavanje u jednom izvoru podataka, a zatim se kreću po vezama do drugih izvora povezanih podataka. Postoje i tražilice povezanih podataka koje pretražuju web praćenjem

¹⁰ Gopal, Pandey. Nav. dj., str. 781.

¹¹ Isto.

¹² Wilson. M.; Matthews, B. Nav. dj., str. 28.

¹³ Gopal, Pandey. Nav. dj., str. 782.

¹⁴ Isto.

veza između izvora podataka, omogućujući složene upite nad agregiranim podacima, slično kao što se danas postavljaju upiti lokalnoj bazi podataka. Također se pojavljuju i nove mogućnosti za aplikacije specifične za pojedinu domenu znanja. Za razliku od *Web 2.0*, hibridnih web aplikacija koje rade s fiksnim skupom izvora podataka, aplikacije koje koriste povezane podatke funkcioniraju povrh nevezanog, globalnog podatkovnog prostora. To im omogućuje da pružaju potpunije odgovore kako se novi izvori podataka pojavljuju na webu.¹⁵

Povezani podaci jednostavno se odnose na korištenje weba za stvaranje tipiziranih odnosno semantičkih poveznica između podataka iz različitih izvora. One mogu biti jednako raznolike kao baze podataka kojima upravljaju dvije organizacije na različitim zemljopisnim lokacijama ili heterogeni sustavi unutar jedne organizacije koji, povijesno gledano, nisu jednostavno međusobno interoperabilni na razini podataka. Tehnički, povezani podaci odnose se na strojnočitljive podatke objavljene na webu, s jasno definiranim značenjem, koji su povezani s drugim vanjskim skupovima podataka. Dok su primarne jedinice hipertekstualnog weba HTML (*HyperText Markup Language*) dokumenti povezani neoznačenim hipervezama (engl. *untyped hiperlinks*), povezani podaci oslanjaju se na dokumente koji sadrže podatke u RDF formatu. Međutim, umjesto jednostavnog povezivanja tih dokumenata, povezani podaci koriste RDF za izradu semantičkih izjava koje povezuju proizvoljne entitete u svijetu. Rezultat koji iz toga proizlazi znanstvenici nazivaju webom podataka (*The Web of Data*).¹⁶

Tim Berners-Lee predstavio je skup načela ili principa za objavljivanje podataka na webu na način da svi objavljeni podaci postaju dio jedinstvenog globalnog podatkovnog prostora. Prvo načelo odnosi se na korištenje URI/IRI-ja kao naziva za 'stvari' (engl. *things*) ili entitete. Drugo načelo nalaže da je za nazive 'stvari', potrebno koristiti HTTP URI/IRI. Slijedeće načelo glasi: "Kada se pristupi URI/IRI-ju, treba dobiti korisne informacije." To znači da prilikom pristupa URI/IRI-ju korisnik ili aplikacija trebaju dobiti relevantne podatke o entitetu koji taj URI/IRI predstavlja, koristeći standarde poput RDF-a ili SPARQL-a. Naposljetku, četvrto načelo se odnosi na uključivanje poveznica na druge URI/IRI-je kako bi se otkrile druge 'stvari', odnosno izvori informacija. Ova pravila postala su poznata kao „načela povezanih podataka“ te pružaju osnovni „recept“ za povezivanje i objavljivanje podataka na webu poštujući njegovu arhitekturu i standarde.¹⁷

¹⁵Usp. Berners-Lee, Tim; Bizer, Christian; Heath, Tom. Linked Data: The Story so Far. // International journal on Semantic Web and information systems, 5(2009), str. 1.

¹⁶ Berners-Lee, Tim; Bizer, Christian; Heath, Tom. Nav.dj., str. 2.

¹⁷ Usp. Berners-Lee, Tim. Linked data, 2006. URL: <https://www.w3.org/DesignIssues/LinkedData.html> (2024-07-04)

2. 3. Tehnologije semantičkoga weba i povezanih podataka

Povezani podaci temelje se na dvjema ključnim tehnologijama za web, a to su jedinstveni identifikator resursa (URI/IRI) i protokol za prijenos hiperteksta (HTTP). Iako su URL-ovi kao posebna vrsta URI/IRI-ja postali poznati kao adrese lokacija na kojima su dokumenti smješteni, oni se koriste i za identifikaciju bilo kojeg entiteta koji postoji u svijetu. Ako su URI/IRI-ji identificirali entitete koji koriste HTTP protokol, ti se entiteti mogu potražiti samo uklanjanjem reference na URI/IRI putem HTTP protokola. Na taj način HTTP protokol pruža jednostavan, ali univerzalan mehanizam za dohvat resursa koji se mogu serijalizirati kao tok bajtova (kao što je primjerice fotografija mačke) ili dohvat opisa entiteta koji se sami po sebi ne mogu slati preko weba na taj način (kao što je sama mačka).¹⁸

URI/IRI i HTTP nadopunjuju se tehnologijom koja je ključna za web podataka – RDF-om. Dok HTML pruža sredstvo za strukturiranje i povezivanje dokumenata na webu, RDF pruža generički podatkovni model zasnovan na grafikonima s kojim strukturira i povezuje podatke koji opisuju stvari u svijetu. RDF model kodira podatke u obliku trostruke veze objekta, subjekta i predikata. Subjekt i objekt su URI/IRI-ji koji upućuju na različite resurse, dok predikat upućuje na njihov odnos. Dakle, predikat određuje kako su subjekt i objekt povezani, a predstavlja ga također URI/IRI. Na primjer, RDF triplet (engl. *RDF triple*) može navesti da su dvije osobe, A i B, svaka identificirana URI/IRI-jem, povezane činjenicom da A zna B. Na sličan način RDF triplet može povezati osobu C sa znanstvenim člankom D u bibliografskoj bazi podataka navodeći da je C autor djela D. Dva tako povezana sredstva mogu se povući iz različitih skupova podataka na internetu, čime se omogućuje povezivanje podataka u jednom izvoru podataka s podacima u drugom izvoru, čime se i ostvaruje vizija semantičkog weba.¹⁹

RDF je temelj i za obradu metapodataka te osigurava interoperabilnost među aplikacijama koje razmjenjuju strojno razumljive informacije na webu. Metapodaci su „podaci o podacima” ili u kontekstu RDF-a „podaci koji opisuju web resurse”. Razlika između podataka i metapodataka nije apsolutna, ona se prvenstveno definira kontekstom ili specifičnom primjenom. Izvorno RDF je rabio XML kao sintaksu enkodiranja metapodataka koja je dobila naziv RDF/XML. Resursi koje opisuje RDF općenito se odnose na sve što se može identificirati i imenovati pomoću URI/IRI-ja, kao što je već rečeno. Opći cilj RDF-a je

¹⁸ Usp. Berners-Lee, Tim; Bizer, Christian; Heath, Tom. Nav. dj., str. 3.

¹⁹ Isto.

definirati mehanizam za opisivanje resursa koji ne daje pretpostavke o određenoj aplikacijskoj domeni niti definira semantiku bilo koje aplikacijske domene.²⁰

RDFS i OWL osnova su za stvaranje vokabulara koji se mogu koristiti za opisivanje entiteta u svijetu te načina kako su povezani. Oni se izražavaju u RDF-u koristeći izraze iz RDFS-a i OWL-a, koji pružaju različite stupnjeve izražajnosti u modeliranju područja interesa. Korištenjem HTTP URI/IRI-ja za identifikaciju resursa, HTTP protokola kao mehanizma za dohvaćanje, te RDF podatkovnog modela za predstavljanje opisa resursa, povezani podaci izravno se temelje na općoj arhitekturi weba.²¹

Semantički web se stoga može promatrati kao dodatni sloj koji je čvrsto isprepleten s klasičnim webom te kao klasični web može sadržavati bilo koju vrstu podataka. Nadalje, bilo tko na semantičkome webu, može objavljivati podatke na bilo kojem vokabularu kojim predstavlja podatke. Entiteti su povezani RDF vezama, stvarajući globalni podatkovni grafikon koji obuhvaća izvore podataka i omogućuje otkrivanje novih izvora podataka.²²

S obzirom na to da je RDF podatkovni model, potreban je i jezik upita nad podacima opisanim tim modelom. Kao što je SQL sredstvo za izražavanje upita u relacijskim bazama podataka, SPARQL (*Simple Protocol And RDF Query Language*) je jezik upita koji se rabi u okviru grafičkog podatkovnog modela RDF-a.²³ SPARQL nudi sintaktički jezik sličan SQL-u za postavljanje upita RDF grafovima putem uspoređivanja uzoraka, kao i jednostavan komunikacijski protokol koji klijenti mogu koristiti za izdavanje SPARQL upita prema krajnjim točkama.²⁴ Kao što je već spomenuto, SPARQL je sintaktički sličan SQL-u za postavljanje upita nad RDF grafovima putem uspoređivanja uzoraka, ali je puno robusniji od SQL-a, budući da je dizajniran za otvoren, decentraliziran i fluidan web. S obzirom na to da se podaci iz RDF-a kojima se opisuju isti objekti u stvarnom svijetu mogu objaviti na više web lokacija, specifikacije SPARQL-a uključuju ne samo jezik upita, već i jednostavan

²⁰ Usp. Miller, Eric; Schloss, Bob. Resource Description Framework (RDF): Model and Syntax, 1997. URL: <https://www.w3.org/TR/WD-rdf-syntax-971002/> (2024-07-14)

²¹ Usp. Berners-Lee, Tim; Bizer, Christian; Heath, Tom. Nav.dj. str. 4.

²² Isto.

²³ Usp. Domingue, John; Fensel, Dieter; Hendler, James A. Introduction to the Semantic Web Technologies. // Handbook of Semantic Web Technologies / Domingue, John; Fensel, Dieter; Hendler, James A. New York: Springer Publishing Company, 2011. Str. 17. URL: <https://link.springer.com/referencework/10.1007/978-3-540-92913-0> (2024-07-21)

²⁴ Valle, Emanuele Della; Ceri, Stefano. Querying the Semantic Web: SPARQL. // Handbook of Semantic Web Technologies / Domingue, John; Fensel, Dieter; Hendler, James A. New York: Springer Publishing Company, 2011. Str. 301. URL: <https://link.springer.com/referencework/10.1007/978-3-540-92913-0> (2024-07-21)

komunikacijski protokol koji klijent može upotrebljavati za izdavanje SPARQL upita na udaljene krajnje točke, čime se dobivaju odgovori u RDF-u ili XML-u.²⁵

SPARQL se sastoji se od skupa RDF tripleta u kojima svaki element - subjekt, predikat i objekt - može biti varijabla, odnosno zamjenski znak (engl. *wildcard*). Rješenja varijabli nalaze se tako da se uzorci RDF tripleta u upitu podudaraju s RDF tripletima u skupu podataka. SPARQL sadrži četiri osnovne vrste upita:

1. ASK – ovom naredbom postavlja se upit postoji li barem jedno podudaranje uzorka upita u RDF podacima;
2. SELECT – naredba kojom se odabiru sva ili neka od tih podudaranja u tabličnom obliku (uključujući agregiranje i uzorkovanje);
3. CONSTRUCT – izrada RDF grafa zamjenom varijabli u pronađenim podudaranjima u skupu predložaka RDF tripleta;
4. DESCRIBE – naredba kojom se opisuju sva podudaranja pronađena izradom odgovarajućeg grafa.²⁶

²⁵ Isto, str. 302.

²⁶ Usp. Ontotext. What is SPARQL. URL: <https://www.ontotext.com/knowledgehub/fundamentals/what-is-sparql/> (2024-08-01)

3. Računalne ontologije

U filozofiji, ontologija istražuje prirodu bića i postojanja. Pojam „ontologija” vuče porijeklo iz grčkih riječi „*onto*”, što znači „biće“ ili „ono što jest“ i „*logos*”, što znači „riječ“, „govor“ ili općenito pisani ili izgovoreni diskurs.²⁷ U računalnim i informacijskim znanostima ponovna uporaba znanja olakšana je uporabom eksplicitne ontologije, za razliku od implicitne ontologije, tj. znanja kodiranog u softverskim sustavima. Stoga su potrebni odgovarajući ontološki jezici kako bi se ostvarila formalna ontologija zasnovana na tri važna aspekta: konceptualizaciji, vokabularima ili rječnicima te aksiomatizaciji. S obzirom na konceptualizaciju, jezik bi trebao odabrati odgovarajući referentni model, kao što je model entiteta i odnosa ili objektno orijentiran model, te osigurati odgovarajuće ontološke konstrukcije koje predstavljaju činjenično znanje, poput definiranja entiteta i odnosa u domeni te uspostavljanje odnosa među entitetima. Osim semantike, jezik bi trebao obuhvatiti i sintaksu kao što su dodjela simbola (tj. dodjeljivanje simbola konceptima) i gramatike. Kako bi se obuhvatila semantika za donošenje zaključaka, uz činjenično znanje potrebna su i pravila i ograničenja. Na primjer, pravila se mogu koristiti za stvaranje novih činjenica iz postojećeg znanja i za potvrđivanje dosljednosti znanja. S druge strane, aktivnosti dijeljenja znanja temeljene na webu zahtijevaju da se ljudski i/ili strojni agenti dogovore o zajedničkoj i eksplicitnoj ontologiji kako bi razmijenili znanje i ostvarili ciljeve suradnje.²⁸

Pojmovne generalizacije individualnih entiteta koji sadrže određena zajednička obilježja u ontologiji se nazivaju klase (engl. *class*). Klase koje sadrže zajednička obilježja mogu se uvrstiti u zajedničke 'natklase' ili superklase (engl. *superclass*). Obilježja entiteta klasa mogu se primijeniti primarno na taj entitet ili opisati njegov odnos s drugim entitetima. Karakteristike se u ontologijama zovu svojstva (engl. *properties*). Svojstva se mogu podijeliti na podatkovna i objektna svojstva. Podatkovna svojstva opisuju značajke entiteta, dok objektna svojstva opisuju odnose među konceptima. Nadalje, individualni entiteti u ontologiji nazivaju se instance (engl. *instances*). Ontologije na razini apstrakcije povrh individualnih instanci uključuju dvije osnovne komponente: klase/pojmove/koncepte i svojstva/relacije. Ontologije uključuju i pojedinačne instance koje opisuju razinu specifičnih konstrukata. Instance

²⁷ Usp. Ding, L. et al. Using Ontologies in the Semantic Web: A Survey. // Springer, 14(2007), str. 79. URL: https://doi.org/10.1007/978-0-387-37022-4_4 (2024-07-10)

²⁸ Isto, str. 80.

pripadaju određenim klasama, a svojstva izražavaju odnos prema drugim pojedinačnim instancama.²⁹

Kako bi se znanje podijelilo u različitim zajednicama ili područjima, pri razvoju eksplicitne ontologije trebalo bi uzeti u obzir tri zahtjeva, i to su proširivost, pristupačnost te primjenjivost. U kontekstu semantičkog weba, kreatori ontologija trebali bi moći postupno razvijati ontologiju, ponovno koristeći što više postojećih popularnih koncepata prije nego što stvore novi koncept. Na primjer, pojam „žena” može se definirati kao potklasa (engl. *subclass*) postojećeg pojma „osoba”. Ovaj zahtjev zahtijeva ekspresivan zajednički referentni model kao i raspodijeljene mehanizme razlučivanja simbola. Nadalje, samo objavljivanje znanja na internetu ne jamči da ga strojevi ili korisnici mogu lako razumjeti. Kako bi znanje bilo vidljivo i pristupačno na webu, potrebna je dodatna zajednička ontološka osnova sintakse i semantike između izdavača informacija i potrošača. Ovaj je zahtjev posebno važan za strojeve, jer nisu sposobni razumjeti znanje zapisano na nepoznatom jeziku. Ontologija ne služi samo prikazu, odnosno numeriranju znanja o činjeničnim područjima, nego i izračunu, odnosno omogućavanju logičkog zaključivanja o činjenicama postupkom aksiomatizacije. Stoga bi ontologija na webu trebala osigurati konstrukte za učinkovito povezivanje s logičkim osnovnim zaključcima i opcijama koje podržavaju različite zahtjeve u pogledu izražajnosti i složenosti računanja. Proširivost je osigurana ne samo odgovarajućim rječnikom pojmova za opis resursa i njegovih odnosa s drugim resursima, već i jednostavnim grafičkim podatkovnim modelom okvira za opis resursa, kao što je RDF. Vidljivost odnosno pristupačnost ontologije ostvaruje se putem web izdavačkog mehanizma (koji počiva na izjavi „svatko može objavljivati izjave o bilo kojem resursu”) koji koristi URI/IRI, XML sintaksu, RDF podatkovni model i ontološke jezike poput RDFS-a i OWL-a. Primjenjivost se osigurava također putem RDFS-a i OWL-a povezivanjem izjava o znanju s logičkim zaključcima na različitim razinama ekspresivnosti i računalne složenosti. Formalno definirana semantika RDF-a i OWL-a igra ključnu ulogu u zaključivanju.³⁰

Ontologije mogu imati ključnu ulogu u omogućavanju obrade, dijeljenja i ponovne upotrebe znanja na webu među aplikacijama. Ontologija obično sadrži hijerarhiju koncepata unutar domene i opisuje ključna svojstva svakog koncepta putem mehanizma atribut-vrijednost. Dodatni odnosi između koncepata mogu se opisati dodatnim logičkim rečenicama.

²⁹ OWL Web Ontology Language Guide. // W3C. URL: <https://www.w3.org/TR/owl-guide/> (2024-08-27)

³⁰ Ding, L. et al. Nav. dj., str. 80–81.

Naposljetku, konstante (kao što je „siječanj”) pridružuju se jednom ili više pojmova (kao što je „mjesec”) kako bi se definirala njihova odgovarajuća kategorija.³¹

³¹ Usp. Decker, Stefan et al. The Semantic Web: the roles of XML and RDF. // IEEE Internet Computing 4, 5(2000), str. 63-64. URL: 10.1109/4236.877487 (2024-07-28)

4. Semantičke grafičke baze podataka i grafovi znanja

4.1. Relacijske vs. grafičke baze podataka

Baza podataka je bilo koja zbirka podataka ili informacija koja je posebno organizirana za brzo pretraživanje i dohvat podataka računalnim putem. Baze podataka strukturirane su kako bi olakšale pohranu, dohvat, izmjenu i brisanje podataka zajedno s različitim postupcima obrade podataka.³²

Relacijske baze podataka posebna su vrsta tabličnih baza podataka koje već desetljećima pružaju podršku za pohranu podataka kroz implementacije poput Oraclea, MySQL-a i drugih. Na početku, baza podataka se koristila samo za pohranu tabličnih podataka kao što su izvješća o kupnji i financijski zapisi. Relacijske baze podataka bile su savršeno rješenje, jer su se transakcije u tablici financija povezivale sa stavkom u tablici kupnje. Ali u današnjem okruženju, relacijski prikazi nisu učinkoviti u izvođenju mnogih operacija, pa tako *World Wide Web* pokazuje puno složenije mreže odnosa nego što se to očekivalo kada je SQL bio dizajniran. Mreža hiperlinkova koja povezuje sve stranice na *World Wide Webu* vrlo je složena i gotovo ju je nemoguće učinkovito modelirati u relacijskoj bazi podataka. Slična pitanja pojavljuju se i u modeliranju društvenih mreža, poput Facebooka. Rješenje takvih problema putem relacijskih baza podataka uključuje velik broj poveznica, čije je izračunavanje skupo i zahtjevno. Isto tako, relacijske baze podataka nisu korisne kada se podatkovni model razvija tijekom vremena, što znači da relacijske baze podataka ovise o krutoj shemi te otežavaju dodavanje novih odnosa između objekata. S obzirom na intenzivan porast korištenja interneta koji je doveo do potrebe za pohranom velikih količina međusobno povezanih podataka, pojavila se potreba za pohranom podataka koja je prilagođena specifičnostima grafičkih baza podataka. Drugim riječima, ograničenja relacijskih baza podataka dovela su do razvoja grafičkih baza podataka.³³

Tehnologija grafičke baze podataka učinkovit je skup alata za modeliranje, pohranu, manipulaciju, postavljanje upita i upravljanje grafičkim podacima usredotočujući se na attribute, vrijednosti i odnose između entiteta u podatkovnom modelu. Ova vrsta baze podataka jednostavna je za razumijevanje, jer se njezin koncept temelji na teoriji grafova, koja koristi

³² Usp. Database. // Encyclopedia Britannica. URL: <https://www.britannica.com/technology/database> (2024-08-01)

³³ Usp. Batra, Shalini; Tyagi, Charu. Comparative Analysis of Relational And Graph Databases. // International Journal of Soft Computing and Engineering (IJSCE) 2, 2(2012), str. 509. URL: https://www.academia.edu/2891354/Comparative_Analysis_of_Relational_And_Graph_Databases (2024-07-14)

matematičke strukture za modeliranje odnosa između objekata. Objekti su predstavljeni čvorovima, dok su odnosi između objekata predstavljeni bridovima između čvorova. Može se reći da su u grafičkoj bazi podataka odnosi između objekata jednako važni kao i sami objekti. Grafičke baze podataka sada su održiva alternativa sustavima relacijskih baza podataka i imaju svoje primjene u mnogim područjima znanosti i inženjerstva poput biologije, kemije, društvenog umrežavanja te semantičkoga weba.³⁴ Postoji nekoliko implementacija grafičkih baza podataka poput Neo4j, Amazon Neptune i drugih.

Grafičke baze podataka su vrlo praktična vrsta baze podataka kako za korištenje tako i za upravljanje, što se očituje u jednostavnosti prikaza pohrane u njihovim strukturama te brz pristup podacima. One su optimizirane za pohranu te postavljanje upita strukturama grafa. Problem grafičkih baza podataka je u tome što ponekad nisu posebno učinkovite u svim željenim operacijama, primjerice prikazu podataka iz relacijskih modela. Naime, grafičke baze podataka ne predstavljaju zamjenu za relacijske baze podataka, već su učinkovito rješenje u upravljanju ogromnim količinama podataka koji sadrže međusobno povezane podatke. Grafičke baze podataka učinkovite su u okviru nekoliko područja, od telekomunikacija do financijskih usluga, ugostiteljstva, logistike i zdravstvene skrbi. Također, grafičke baze podataka ključne su u glavnim područjima, kao što su analiza ponašanja na društvenim mrežama, upravljanje podacima i studije povezane s popisom stanovništva. Grafičke baze podataka imaju mnoge prednosti u odnosu na relacijske baze podataka; primjerice, pretraživanje informacija je znatno poboljšano, jer se podaci nalaze blizu jednog ili više glavnih čvorova. Jedna od prednosti grafičke baze podataka jest njena intuitivnost zbog svoga prirodnog oblika prikaza informacija, a to je graf. U usporedbi s relacijskim bazama podataka, grafičke baze podataka pokazuju se učinkovitije u provedbi vrlo dubokih pretraga. Primjerice, eBay, velika mrežna trgovina, također je počela primjenjivati grafičke baze podataka kako bi razumjela navike, afinitete i ponašanje online kupaca te im u stvarnom vremenu pružila personalizirane preporuke. Dakle, pomoću grafičkih baza podataka tvrtke mogu velikom brzinom povezati ogromnu količinu kompleksnih podataka o proizvodima i kupcima kako bi stekle uvid ne samo u potrebe kupaca, već i u trendove proizvoda.³⁵

³⁴ Usp. Hadi Hor, A.E. et al. A Semantic Graph Database for Bim-Gis Integrated Information Model for An Intelligent Urban Mobility Web Application. // ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4, 4(2018), str. 90. URL: <https://ui.adsabs.harvard.edu/abs/2018ISPRAn.4.4...89H/abstract> (2024-07-15)

³⁵ Usp. Guia, José; Gonçalves Soares, Valéria; Bernardino, Jorge. Graph Databases: Neo4j Analysis. // International Conference on Enterprise Information Systems, 2017, str. 351-352. URL:

4.2. Semantičke grafičke baze podataka

Semantičke grafičke baze podataka su vrsta grafičkih baza podataka koje dodatno uključuju semantičku komponentu, temeljenu na tehnologijama poput RDF-a i OWL-a. Ove baze podataka ne samo da pohranjuju podatke u obliku grafa, već koriste ontologije kako bi definirale značenje i odnose među podacima, omogućujući strojno razumijevanje tih odnosa. Semantičke grafičke baze podataka se koriste za složenije scenarije gdje je potrebno formalizirati znanje i omogućiti napredno pretraživanje i zaključivanje. Primjeri uključuju grafičke baze podataka koje podržavaju semantički web, poput Stardoga, OpenLink Virtuosa ili GraphDB-a koji će se koristiti u praktičnom dijelu rada.

Dok su grafičke baze podataka usmjerene na strukturu i odnose među podacima, semantičke grafičke baze podataka dodaju sloj značenja i formalne semantike, omogućujući dublje razumijevanje i manipulaciju podacima na način koji podržava semantičke tehnologije i ontologije. Može se zaključiti da semantičke grafičke baze podataka predstavljaju specijaliziranu podskupinu grafičkih baza podataka.

4.3. Grafovi znanja

Od ranih dana, istraživači semantičkoga weba zalagali su se za grafički prikaz znanja, a to su činili poticanjem korištenja RDF-a. U takvoj grafičkoj prezentaciji znanja, kao što je već spomenuto, entiteti kao čvorovi grafa povezani su odnosima, koji su predstavljeni bridovima grafa. Drugim riječima, graf je podatkovna struktura koja se sastoji od bridova i čvorova. U mnogim slučajevima, skupovi entiteta i odnosa organizirani su u obliku računalne ontologije koja definira njihove klase, svojstva, instance i ograničenja uporabe. Dolaskom povezanih podataka predloženo je međusobno povezivanje različitih skupova podataka u semantičkome webu. Kroz međusobno povezivanje skupovi podataka mogu se shvatiti i kao jedan veliki graf znanja, iako vrlo heterogen po prirodi. Svaki prikaz nekog znanja temeljen na grafikonu i serijalizaciji u RDF-u može se smatrati grafom znanja.³⁶

Naime, objekt i subjekt RDF izjave smatraju se čvorovima grafa i povezani su predikatom koji označava svojstvo, a sam predikat je predstavljen strelicom. Čvor može ostati

<https://www.semanticscholar.org/paper/Graph-Databases%3A-Neo4j-Analysis-Guia-Soares/c867c372ed78e27c53feaf788e74aa7506302bbe#citing-papers> (2024-07-24)

³⁶Usp. Paulheim, Heiko. Knowledge graph refinement: A survey of approaches and evaluation methods. // Semantic Web 8, 3(2017), str. 490-491. URL: <https://content.iospress.com/articles/semantic-web/sw218> (2024-07-10)

prazan (engl. *blank nodes*), u obliku slovne vrijednosti (engl. *literal*) ili u obliku URI/IRI-ja. Čvorovi prikazani URI/IRI-jima zovu se izvorima prvoga reda te su obilježeni ovalnim oblikom. Naime, subjekt i predikat RDF izjave dani su u obliku neslovne vrijednosti (engl. *non-literal*), identificirani URI/IRI-jima, dok objekt može biti i neslovna i slovna vrijednost. U slučaju slovne vrijednosti, objekt RDF izjave u grafu se predočava pravokutnikom. Subjekti RDF izjave nikada ne mogu poprimiti slovne vrijednosti, ali mogu postati objekti druge RDF izjave. Ključno je naglasiti kako slovne vrijednosti ne mogu biti subjekti, već samo završne točke samoga grafa. RDF izjave putem čvorova koji su identificirani kao URI/IRI-ji mogu se međusobno neograničeno povezivati. Tim načinom na webu se može izgraditi globalni graf u kojemu su izvori međusobno direktno povezani.³⁷

Graf znanja DBpedije prvi je put objavljen 2007. godine i jedan je od najpoznatijih grafova znanja. DBpedia sama po sebi nije semantička grafička baza podataka, već *izvor podataka* prezentiran u obliku grafa znanja temeljen na tehnologijama semantičkog weba. To je središnji skup podataka na semantičkome webu koji je povezan s mnogim drugim skupovima podataka, što je vidljivo na dijagramu Oblaka otvorenih povezanih podataka (*The Linked Open Data Cloud*), skupnoj inicijativi koja se odnosi na praksu objavljivanja strukturiranih podataka oslanjajući se na tehnologije semantičkog weba.³⁸ Graf znanja DBpedije nastao je temeljem strukturiranih podataka Wikipedije. DBpedijina ontologija sadrži 760 klasa i oko 3000 svojstava.³⁹ Također, DBpedija ima mnoga područja primjene, kao što su obrada prirodnog jezika te istraživanje i obogaćivanje znanja.⁴⁰

S druge strane, Googleov graf znanja (*Google Knowledge Graph*) predstavljen 2012. godine, upotrebljava se za identifikaciju i raščlanjivanje subjekata u tekstu, obogaćivanje rezultata pretraživanja semantički strukturiranim sažecima te za pružanje poveznica na povezane subjekte u istraživačkom pretraživanju radi poboljšanja mogućnosti tražilice i poboljšanja korisničkog iskustva pri pretraživanju.

Nedugo nakon pojave Googleovog grafa znanja, mnoge druge tvrtke počele su razvijati vlastite grafove znanja. Na primjer, Microsoftov pretraživač Bing proveo je integraciju sa sličnim grafom znanja pod nazivom Satori. Danas grafovi znanja predstavljaju uobičajeno

³⁷ RDF 1.1 Concepts and Abstract Syntax. // W3C. URL: <https://www.w3.org/TR/rdf-concepts/#section-rdf-graph> (2024-08-28)

³⁸ *The Linked Open Data Cloud*. URL: <https://lod-cloud.net/> (2024-07-24)

³⁹ DBpedija. URL: <https://www.dbpedia.org/> (2024-08-27)

⁴⁰ Usp. Zou, Xiaohan. A Survey on Application of Knowledge Graph. // *Journal of Physics: Conference Series* 1487, 4(2020), str. 7. URL: <https://iopscience.iop.org/article/10.1088/1742-6596/1487/1/012016> (2024-07-10)

korisničko sučelje semantičkih grafičkih baza znanja, kao što su YAGO, Wikidata ili Freebase.⁴¹

⁴¹ Zou, Xiaohan. Nav. dj., str. 1.

5. Primjer izrade grafa znanja

5.1. Metodologija

U ovom poglavlju demonstrirat će se kreiranje grafa znanja iz skupa povezanih podataka i odgovarajuće ontologije. Prilikom izrade grafa znanja koristit će se metoda prototipa te metoda modeliranja. Metoda prototipa očituje se u kreiranju grafa znanja u okviru proizvoljno odabranog segmenta knjižničarske struke. Metoda modeliranja upotrebljava se u procesu izbora vrste podataka, klasa, svojstava, i instanci koji su prikazani u ontologiji te skupu povezanih podataka. Korišteni programi u procesu kreiranja grafa znanja su MS Excel, OpenRefine, Protégé te GraphDB.

Primjer korišten za izradu grafa znanja tematski je vezan uz knjižničnu struku, konkretno, uz količinu knjiga razvrstanih po vrsti knjižne građe u dvanaest hrvatskih narodnih knjižnica. Podaci su preuzeti s mrežnih stranica knjižnica, odnosno iz knjižničnih kataloga. Za izradu grafa znanja, najprije je bilo potrebno kreirati skup povezanih podataka, a potom i odgovarajuću ontologiju.

5.2. Izrada skupa povezanih podataka

Za opis vrsta građe koje posjeduju narodne knjižnice odabrane su sljedeće vrste podataka koje su unesene u tablicu u programu MS Excel: naziv knjižnice, URI/IRI knjižnice, vrste građe te količina. U prvom polju tablice nalazi se identifikator svakog zapisa putem kojeg se vrši konverzija linearnih podataka u RDF format pomoću programa OpenRefine. Nadalje, bilo je potrebno transformirati identifikator zapisa u HTTP URI/IRI s obzirom na to da identifikator mora biti jedinstven. Također je preporučljivo leksički označiti svrhu podatkovnog zapisa.

Vrsta gradje u knjižnicama	Naziv knjižnice	URI knjižnice	Vrsta gradje	Kolicina
https://oziz.ffos.hr/knjiznice-vrste-gradje#1	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	knjiga	76664
https://oziz.ffos.hr/knjiznice-vrste-gradje#2	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	članak	6686
https://oziz.ffos.hr/knjiznice-vrste-gradje#3	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	nakladnički niz	3541
https://oziz.ffos.hr/knjiznice-vrste-gradje#4	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	vizualna građa	1833
https://oziz.ffos.hr/knjiznice-vrste-gradje#5	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	e-knjiga	181
https://oziz.ffos.hr/knjiznice-vrste-gradje#6	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	igračka	158
https://oziz.ffos.hr/knjiznice-vrste-gradje#7	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	elektonička građa	114
https://oziz.ffos.hr/knjiznice-vrste-gradje#8	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	glazbena građa	101
https://oziz.ffos.hr/knjiznice-vrste-gradje#9	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	časopis/periodika	72
https://oziz.ffos.hr/knjiznice-vrste-gradje#10	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	tematski broj	63
https://oziz.ffos.hr/knjiznice-vrste-gradje#11	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	notna građa	47
https://oziz.ffos.hr/knjiznice-vrste-gradje#12	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	društvena igra	36
https://oziz.ffos.hr/knjiznice-vrste-gradje#13	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	zvučna građa	24
https://oziz.ffos.hr/knjiznice-vrste-gradje#14	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	zvučna e-knjiga	22
https://oziz.ffos.hr/knjiznice-vrste-gradje#15	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	kartografska građa	19
https://oziz.ffos.hr/knjiznice-vrste-gradje#16	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	grafička građa	6
https://oziz.ffos.hr/knjiznice-vrste-gradje#17	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	novine	6
https://oziz.ffos.hr/knjiznice-vrste-gradje#18	Gradska knjižnica i čitaonica Vinkovci	www.gkvk.hr	ostalo	7
https://oziz.ffos.hr/knjiznice-vrste-gradje#19	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	knjiga	55013
https://oziz.ffos.hr/knjiznice-vrste-gradje#20	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	članak	4992
https://oziz.ffos.hr/knjiznice-vrste-gradje#21	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	nakladnički niz	2531
https://oziz.ffos.hr/knjiznice-vrste-gradje#22	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	e-knjiga	514
https://oziz.ffos.hr/knjiznice-vrste-gradje#23	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	časopis/periodika	105
https://oziz.ffos.hr/knjiznice-vrste-gradje#24	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	elektonička građa	77
https://oziz.ffos.hr/knjiznice-vrste-gradje#25	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	notna građa	72
https://oziz.ffos.hr/knjiznice-vrste-gradje#26	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	igračka	50
https://oziz.ffos.hr/knjiznice-vrste-gradje#27	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	tematski broj	50
https://oziz.ffos.hr/knjiznice-vrste-gradje#28	Gradska knjižnica i čitaonica Đakovo	www.djknjiznica.com	vizualna građa	40

Slika 2 – Prikaz tablice s podacima o vrstama građe u narodnim knjižnicama.

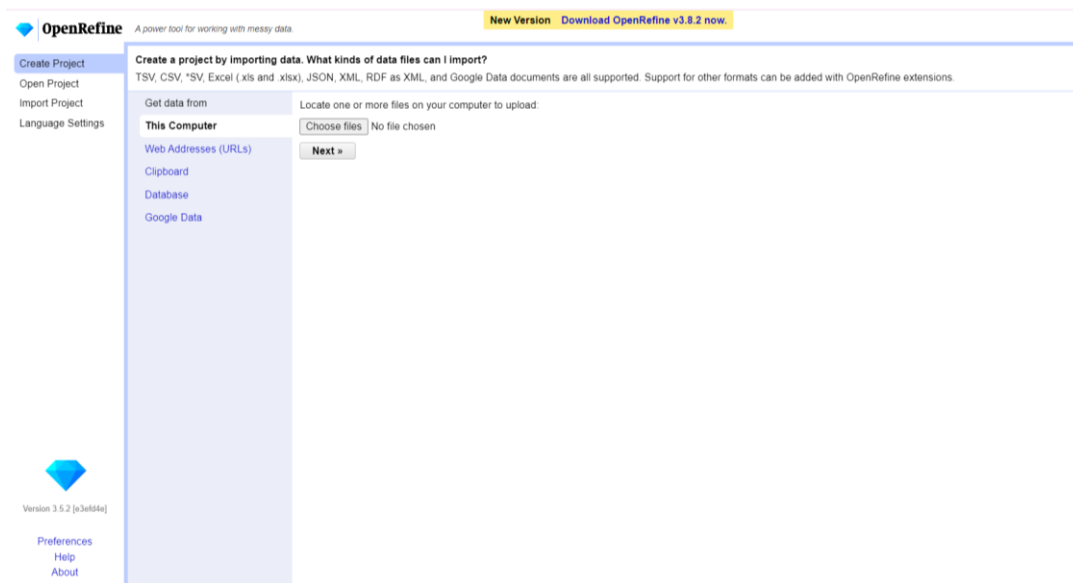
Na Slici 2. predočena je tablica u programu MS Excel o vrstama građe u odabranih dvanaest hrvatskih narodnih knjižnica. Tablica sadrži 192 zapisa, a podaci su prikupljeni u srpnju 2024. godine. Jedan od načina konverzije tabličnih skupova podataka u RDF format nudi program OpenRefine.⁴² Za to je bilo neophodno kreirati predložak za transformaciju tabličnih podataka u RDF podatke, odnosno zapis (Slika 3).

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns="https://oziz.ffos.hr/knjiznice-vrste-gradje#"
  xmlns:diplomski="http://oziz.ffos.hr/diplomski#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" |
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <vrste rdf:about="{jsonize(cells["Vrsta_gradje_u_knjiznicama"].value).replace("'", '')}">
    <naziv_knjiznice rdf:datatype="http://www.w3.org/2001/XMLSchema#string">{jsonize(cells["Naziv_knjiznice"].value).replace("'", '')}</naziv_knjiznice>
    <uri rdf:resource="{jsonize(cells["URI_knjiznice"].value).replace("'", '')}">
    <vrsta_gradje rdf:datatype="http://www.w3.org/2001/XMLSchema#string">{jsonize(cells["Vrsta_gradje"].value).replace("'", '')}</vrsta_gradje>
    <kolicina rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">{jsonize(cells["Kolicina"].value).replace("'", '')}</kolicina>
  </vrste>
</rdf:RDF>
```

Slika 3 – Predložak za transformaciju podataka u RDF zapis.

⁴² OpenRefine je besplatna aplikacija otvorenog koda za upravljanje svim vrstama podatkovnih dokumenata. OpenRefine je započeo kao projekt pod nazivom Freebase Gridworks, koji je kupio Google i preimenovao u GoogleRefine 2010. godine. Googleova službena podrška za projekt završila je 2012., a OpenRefine je prešao na istoimenu aplikaciju otvorenog koda. OpenRefine je moćan alat koji olakšava rad s takozvanim neurednim podacima, čisti ih, pretvara iz jednog formata u drugi, proširuje i integrira s web servisima i vanjskim podatkovnim vezama. Ovaj program štiti podatke korisnika, čuvajući ih privatnima na osobnom računalu korisnika sve dok korisnik ne odluči podijeliti podatke ili surađivati s drugim korisnicima ili institucijama. Vidi: OpenRefine. Documentation. URL: <https://openrefine.org/documentation.html> (2022-08-05).

U prvom dijelu predložka navode se imenski prostori korištenih rječnika. Zatim je bilo potrebno pravilno oblikovati sintaksu za izradu redaka u RDF datoteci. Iz prvog stupca tablice preuzet je URI/IRI subjekta zapisa, dizajniran kao <https://oziz.ffos.hr/knjiznice-vrste-gradje#> s dodanim rednim brojem. Definirani su podaci poput naziva knjižnice, URI/IRI knjižnice, vrste građe i količine uz pomoć JSON (*JavaScript Object Notation*) sintakse, što je prikazano na Slici 3. Nakon izrade predložka, isti se učitava u OpenRefine zajedno s Excel tablicom za stvaranje skupa povezanih podataka.

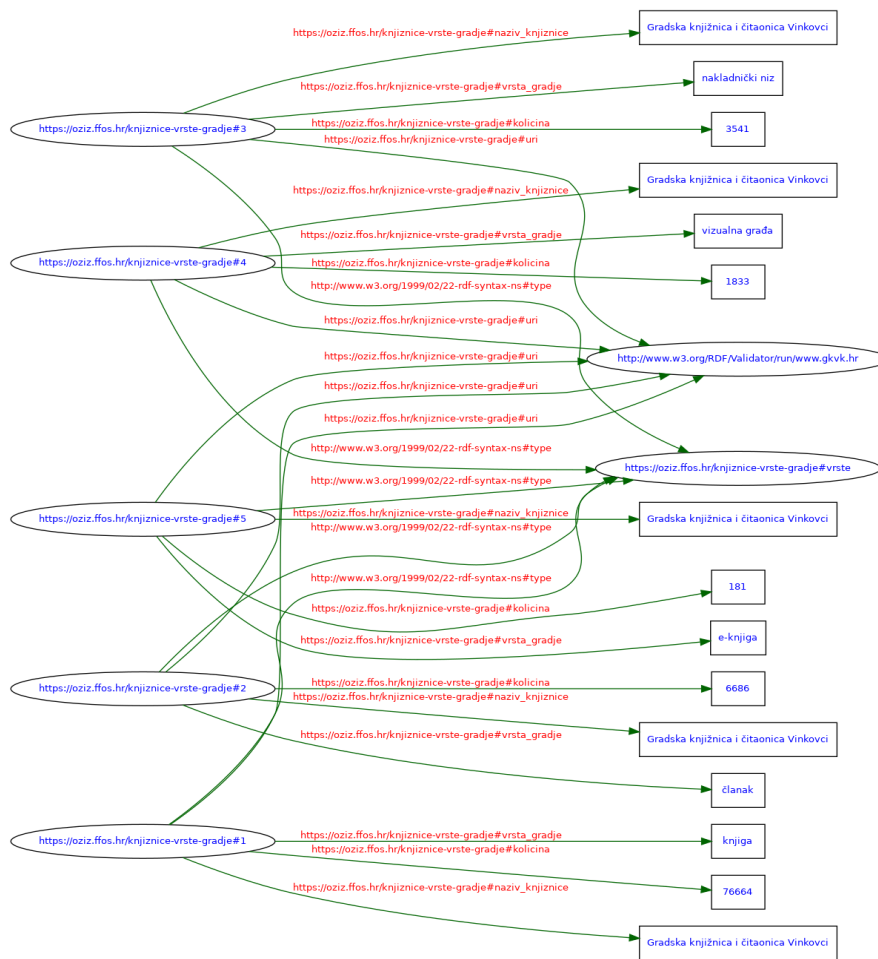


Slika 4 – Početna stranica programa OpenRefine.

Na mrežnim stranicama programa OpenRefine nalaze se upute za instalaciju te dokumentacija za korištenje samoga programa. Nakon učinkovite instalacije, program se pokreće u mrežnome pregledniku te se prikaže početna stranica programa, što je vidljivo na Slici 4. Učitavanjem datoteke prikazuje se zaslom sa svim podacima iz tablice te nakon provjere valjanosti svih podataka kreira projekt.

Kao što je vidljivo na Slici 6., segmenti predloška upisuju se u odgovarajuća polja. Sukladno tomu, deklaracijski kôd u kojem su definirani imenski prostori rječnika upisuje se u polje prefiks (engl. *prefix*). Zatim se u sljedeće polje upisuje kôd koji utvrđuje princip oblikovanja RDF tripleta u samome skupu povezanih podataka. Nadalje, polje za razdvajanje redova treba ostati prazno te se ovdje briše zarez koji je prvotno postojao pri otvaranju obrasca. Naposljetku, u posljednje polje ključno je unijeti završnu oznaku RDF dokumenta. Na desnoj strani obrasca prikazuje se rezultat, ukoliko su svi podaci ispravno upisani. Obrazac prikazuje transformirane tablične podatke u RDF. Svaki zapis u tablici pretvara se u semantičku izjavu s vlastitom semantičkom vrijednošću. U ovom trenutku važno je provjeriti tablicu rezultata, odnosno provjeriti jesu li svi podaci ispravno prikazani. Pogreške se mogu pojaviti ako predložak nije ispravno formatiran ili ako svi podaci nisu uneseni u tablicu. Nakon dovršetka svih navedenih koraka, dobivenu RDF datoteku moguće je vizualizirati i ujedno validirati pomoću RDF Validator.⁴³ Ukoliko ne postoje sintaktičke pogreške u datoteci, RDF Validator prikazat će poruku da je datoteka ispravna. Ako je provjera valjanosti uspješna, RDF tripleti prikazat će se u obliku grafa (Slika 7.) i u tablici.

⁴³RDF Validator. // W3C Consortium. URL: <https://www.w3.org/RDF/Validator/> (2024-08-05)



Slika 7 – Dio RDF grafa RDF dokumenta generiran RDF validatorom.

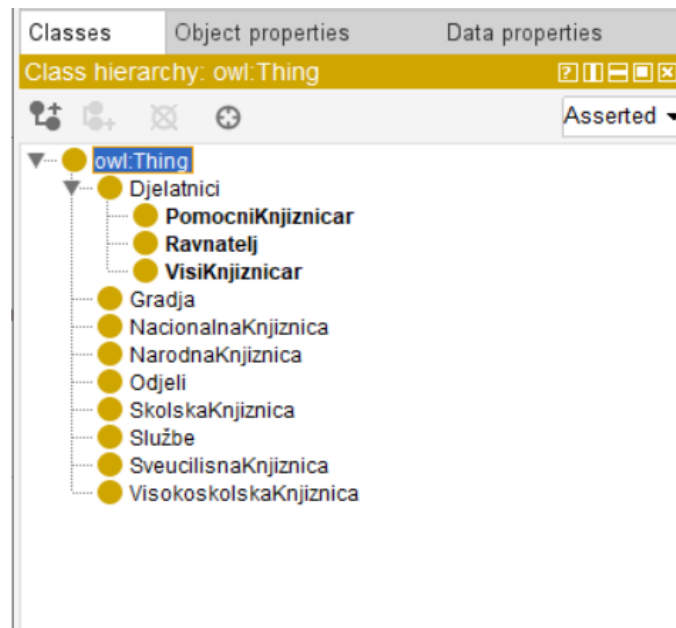
Na Slici 7. prikazan je samo jedan dio RDF grafa odabranog skupa podataka (prvih pet setova podataka), zbog preglednosti rezultata. Također, RDF graf prikazuje i kako je svaki podatkovni element predstavljen RDF tripletom. Ova vrsta prezentacije podataka omogućuje bolje razumijevanje međusobnih odnosa entiteta u određenoj domeni znanja.

5.3. Izrada ontologije

Ontologija je izrađena u Protégé besplatnome programu za izradu i razvoj ontologija. Protégé je osmislio Mark Musen 1987. godine, i otad ga razvija tim na Sveučilištu Stanford. Protégé pruža grafičko korisničko sučelje za kreiranje ontologije. Također uključuje deduktivne klasifikatore kako bi se potvrdila dosljednost modela i izvele nove informacije na temelju analize ontologije. Ovaj program napisan je u programskome jeziku Java. Protégé podupire W3C standarde, kao što su OWL i RDF te omogućava pohranu ontologije u raznolikim formatima pohrane, Turtle, RDF/XML, OWL/XML i dr. Postoje dvije inačice ovoga programa,

a to su Web Protégé te Protégé Desktop aplikacija.⁴⁴ Potonja je primijenjena u realizaciji ove ontologije, a sama instalacija programa je vrlo jednostavna putem mrežne stranice programa. Ogljedna ontologija knjižnica korištena u radu za potrebe kreiranja grafa znanja napisana je u OWL ontologijskom jeziku i pohranjena u Turtle formatu.

U ontologiji je kreirano 12 klasa i potklasa koje su prikazane na Slici 8.



Slika 8 – Prikaz klasa i potklasa ontologije.

Nadalje, ontologija uključuje pet objektnih svojstava. Svojstva :imaOdjel, :imaSluzbu, :imaVrstuGradje deklarirana su kao objektna svojstva. Navedena svojstva imaju istu domenu, koja uključuju klase :NarodnaKnjiznica, :SveucilisnaKnjiznica, VisokoskolskaKnjiznica, :SkolskaKnjiznica te :NacionalnaKnjiznica. Opseg svojstva :imaOdjel je klasa :Odjel, svojstva :imaSluzbu je klasa :Sluzba, a svojstva :imaVrstuGradje je klasa :Gradja.

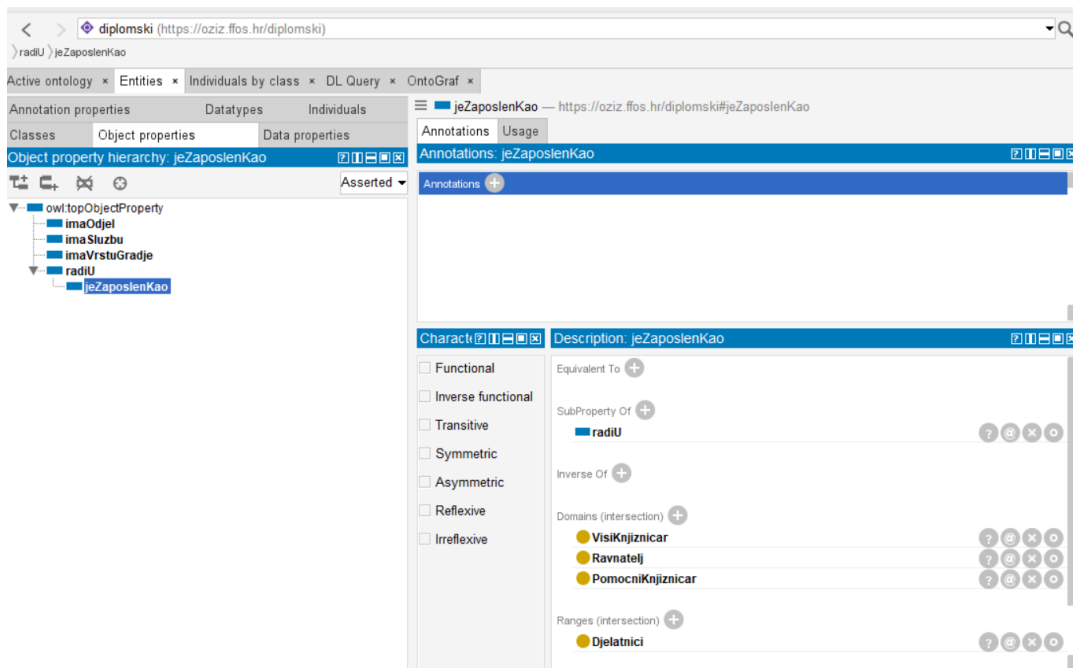
Navedenim svojstvom :imaVrstuGradje izražava se koje sve vrste građe posjeduje određena knjižnica, kao u primjeru Gradske knjižnice i čitaonice Vinkovci, kao instance klase :NarodnaKnjiznica:

```
:Gradska_knjižnica_i_čitaonica_Vinkovci rdf:type owl:NamedIndividual ,
                                           :NarodnaKnjiznica ;
                                           :imaVrstuGradje :Clanak ,
```

⁴⁴ Protégé. URL: <https://protegeproject.github.io/protege/> (2024-08-07)

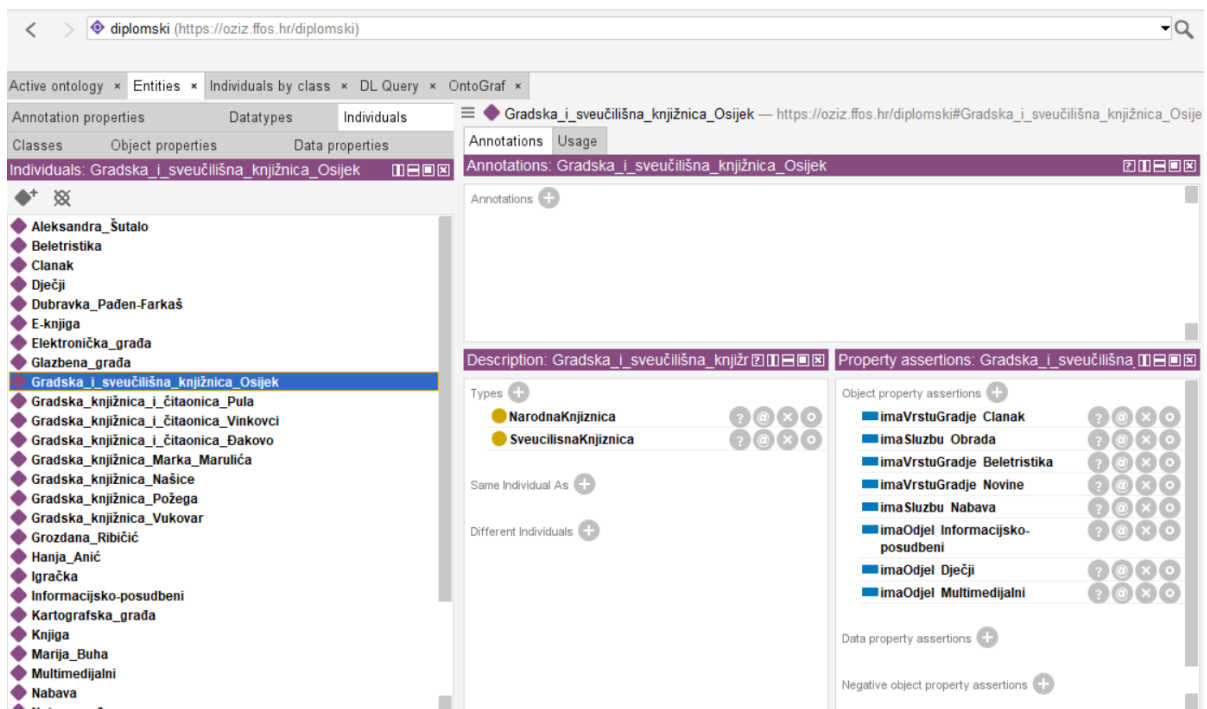
:E-knjiga ,
:Elektronička_građa ,
:Knjiga ,
:Notna_građa ,
:Novine .

Naposljetku, objektno svojstvo :radiU posjeduje i objektno podsojstvo :jeZaposlenKao.



Slika 9 – Prikaz objektnih svojstava u ontologiji knjižnica.

U ontologiju knjižnica uključeno je i 32 instance. Jednostavnim primjerom jedne od instanci ontologije, najlakše je razumjeti kako se instance deklariraju, određuje njihova pripadnost određenoj klasi ili klasama te kako se povezuju s objektnim svojstvima.



Slika 10 – Prikaz instanci klase.

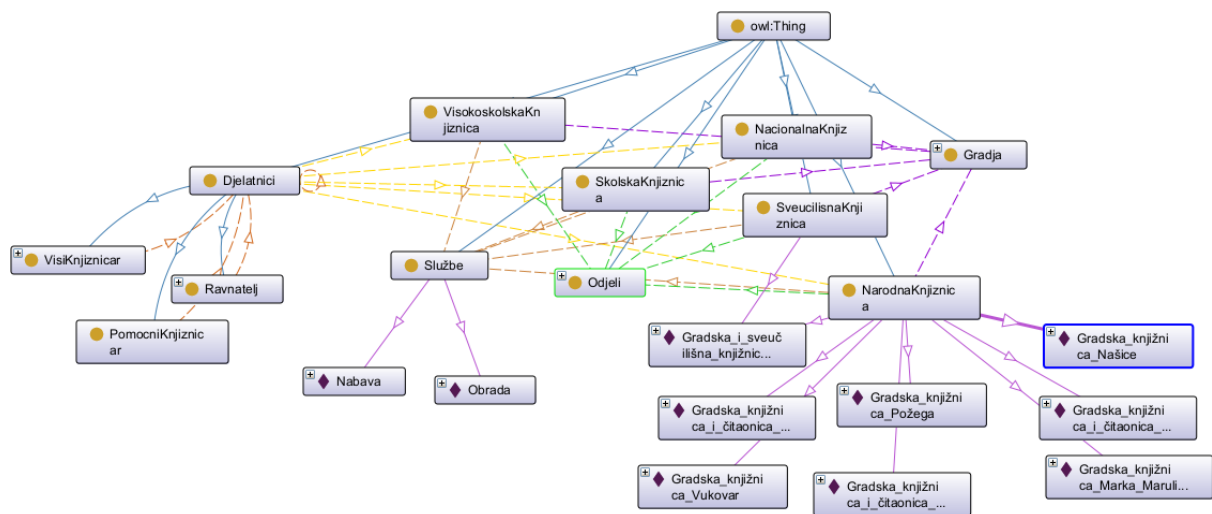
Kao što je vidljivo na Slici 10. instanca `:Gradska_i_sveučilišna_knjiznica_Osijek` pripada dvjema klasama `:NarodnaKnjiznica` te `:SveucilisnaKnjiznica` s obzirom da je jedina knjižnica u Hrvatskoj koja ima dvojaku funkciju. Primjer deklariranja navedene instance donosi se u nastavku:

```
:Gradska_i_sveučilišna_knjiznica_Osijek rdf:type owl:NamedIndividual ,
                                           :NarodnaKnjiznica ,
                                           :SveucilisnaKnjiznica ;
:imaOdjel :Dječji ,
           :Informacijsko-posudbeni ,
           :Multimedijalni ;
:imaSluzbu :Nabava ,
           :Obrada ;
:imaVrstuGradje :Beletristika ,
                :Clanak ,
                :Novine .
```

Owl:NamedIndividual odnosi se na izraz kojim se u označava instanca klase u OWL-u.

Naposlijetku, kreiran je i graf u Ontograf dodatku (engl. *plug-in*) programu Protégé koji pruža podršku interaktivnom kretanju između odnosa OWL ontologije. Podržani su

različiti prikazi svojstava, klasa i potklasa, instanci i sl. Odnosi i vrste čvorova mogu se filtrirati kako bi se stvorio željeni prikaz.



Slika 11 – Ontograf ontologije knjižnica.

Na Slici 11. prikazane su sve klase navedene ontologije, uključujući i potklase. Nadalje, moguće je uočiti i instance klase :NarodnaKnjiznica. Naime, klikom na određenu klasu moguće je prikazati sva svojstva, potklase te sve instance navedene klase.

5.4. Izrada grafa znanja

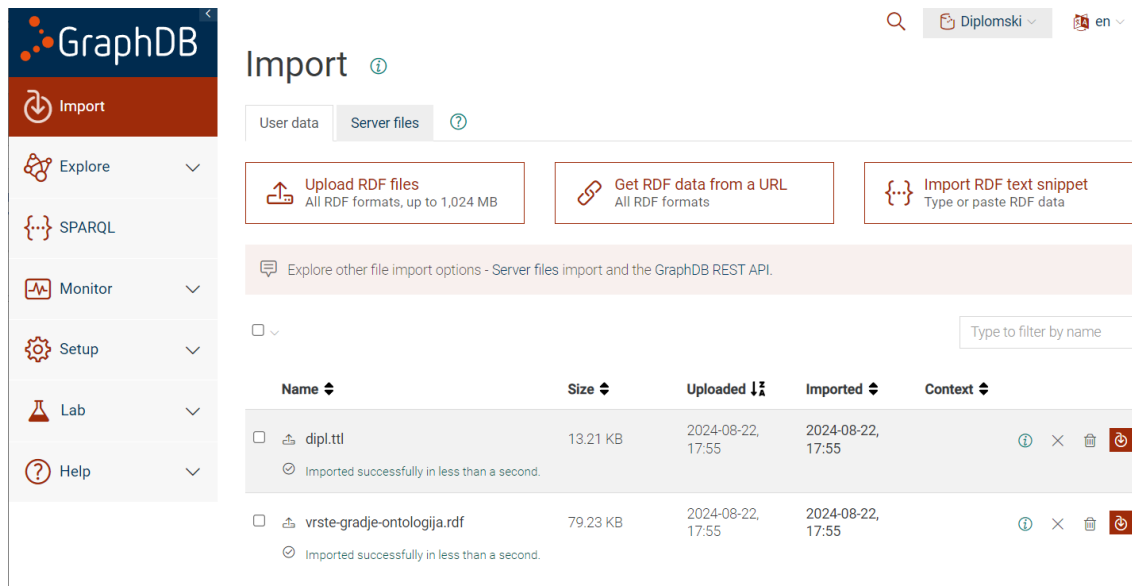
Konačno, za izradu grafa znanja korišten je besplatan program GraphDB (dostupan u *open-source* i komercijalnim verzijama) koji se koristi kao semantička grafička baza podataka ili RDF triplet-spremište (engl. *RDF Triplestore*). GraphDB upravlja RDF podacima, upitima i zaključcima u stvarnom vremenu. Obrazloženje i procjena upita provode se preko trajnog sloja pohrane, dok se učitavanje i rasuđivanje odvija izuzetno brzo, čak i u slučaju ontologija i skupova povezanih podataka koji pohranjuju velike količine podataka. Osim toga, nakon uvoza datoteka moguće je postavljati SPARQL upite putem kojih se mogu kreirati grafove znanja.⁴⁵

Za kreiranje grafa znanja, neophodno je preuzeti već navedeni program sa službene mrežne stranice programa.⁴⁶ Na mrežnim stranicama GraphDB programa nalaze se upute za instalaciju te sva potrebna dokumentacija. Nakon uspješne instalacije, program se otvara u mrežnome pregledniku te se na početnom zaslonu prikaže vodič s kratkim uputama o korištenju

⁴⁵ Usp. GraphDB: Documentation. URL: <https://graphdb.ontotext.com/documentation/10.7/index.html> (2024-08-12)

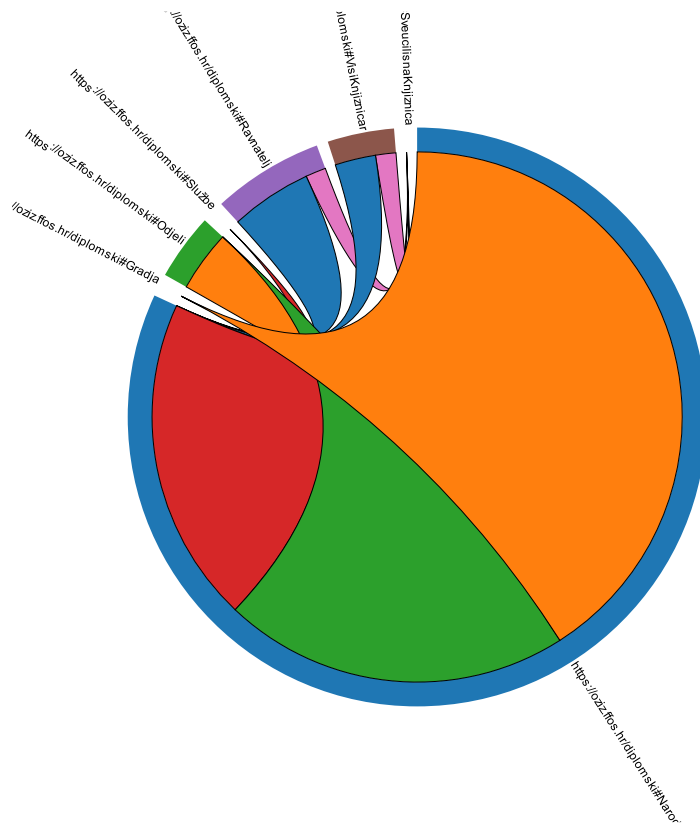
⁴⁶ GraphDB. URL: <https://graphdb.ontotext.com/> (2024-08-12)

programa. Prvi korak uključuje kreiranje repozitorija te učitavanje datoteka koje sadrže skup povezanih podataka te ontologiju, kao što je prikazano na Slici 12. GraphDB podržava učitavanje lokalnih datoteka na poslužitelj te učitavanje podataka s web adresa, oba načina uključuju formate poput .rdf, .ttl, .xml, .json i drugih. Također, postoji mogućnost i ručnog unošenja teksta RDF datoteka.



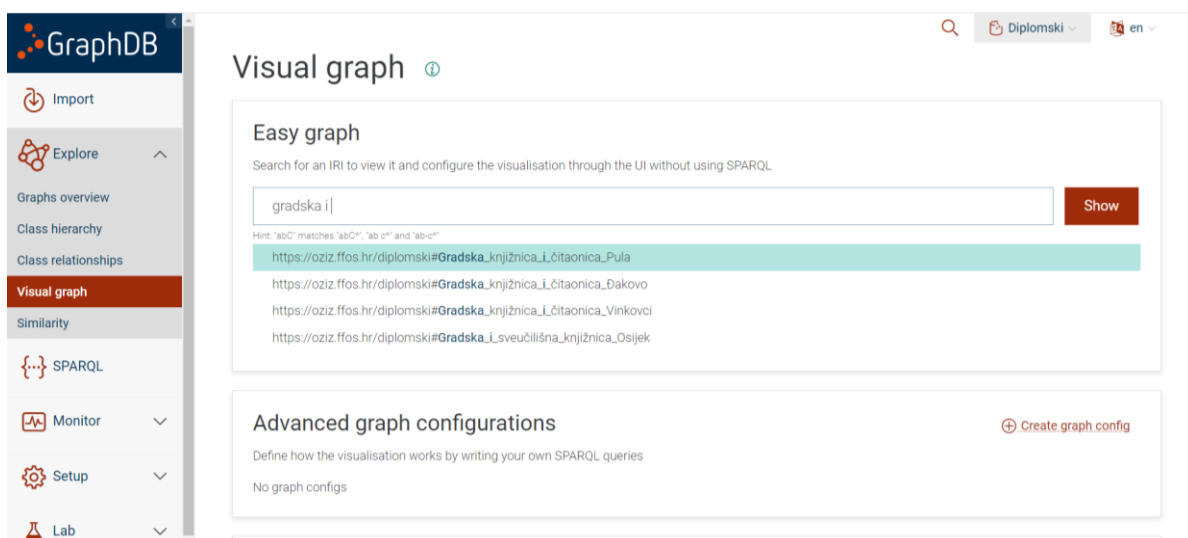
Slika 12 – Učitavanje datoteka ontologije i skupa povezanih podataka na poslužitelj.

Nakon uspješnog učitavanja datoteka ontologije i skupa povezanih podataka na poslužitelj, u lijevom padajućem izborniku Istraživanja (engl. *Explore*), postoje brojne opcije prikaza samih podataka. Tako, primjerice, opcija pod nazivom *Class Relationships* prikazuje odnose između RDF klasa, gdje je svaki odnos prikazan vezama između pojedinačnih instanci dviju klasa. Svaki odnos je RDF izjava u kojoj je subjekt instanca jedne klase, objekt instanca druge klase, a odnos predikat, kao što je prikazano na Slici 13.



Slika 13 – Prikaz odnosa između klasa u ontologiji knjižnica.

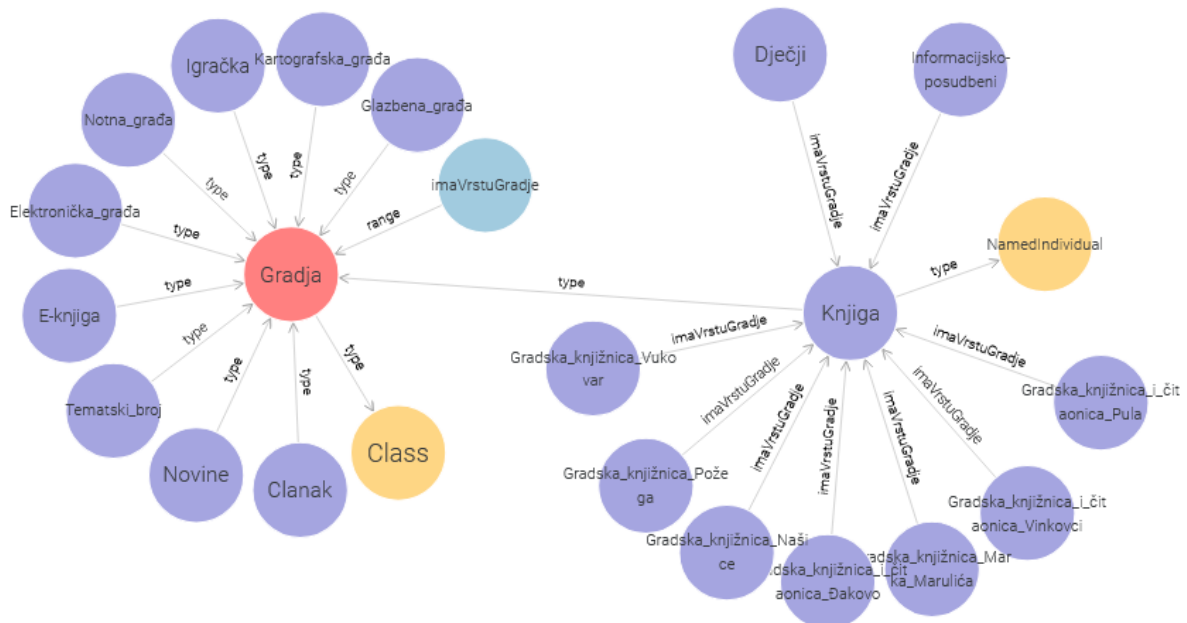
Nadalje, stavka u izborniku *Explore* pod nazivom *Visual Graph* pruža vizualni prikaz dijelova RDF grafa u obliku grafa znanja. Sama vizualizacija započinje od jednog resursa upisom njegovog URI/IRI-ja te se može proširiti na resurse koji su s njim povezani, kao što je prikazano na Slici 14.



Slika 14 – Konfiguracija grafa znanja putem URI/IRI-ja.

U primjeru odabran je URI/IRI: https://oziz.ffos.hr/diplomski#Gradska_knjiznica_i_čitaonica_Vinkovci putem kojeg je kreiran graf znanja, prikazan na Slici 15. Dvostrukim klikom na klasu :NarodnaKnjiznica graf znanja se proširuje dodatnim svojstvenim odnosima te objektnim svojstvima. Konačno, klikom na pojedinu vrstu građe, u ovom primjeru na Knjige, otvara se desna bočna traka u kojoj se prikazuju podaci iz skupa podataka. Iz grafa znanja moguće je očitati da Gradska knjižnica i čitaonica Vinkovci posjeduje 76.664 primjerka Knjiga.

Sličnim postupkom kreiran je i graf znanja o vrstama građe u narodnim knjižnicama, koji je prikazan na Slici 16. Naime, dvostrukim klikom na instancu :Knjiga unutar klase :Gradja pružen je uvid u posjedovanje iste vrste građe u određenim knjižnicama. Važno je napomenuti da se dvostrukim klikom na klasu, instancu klase ili objektno svojstvo graf znanja vrlo lako i brzo proširuje.



Slika 16 – Graf znanja o vrsti građe u knjižnicama.

S druge strane, SPARQL upitima nad ontologijom i skupom povezanih podataka također je moguće kreirati grafove znanja. SPARQL, kao upitni jezik za RDF, nudi nekoliko tipova upita, od kojih su, kao što je spomenuto u teorijskom dijelu rada, najvažniji SELECT, CONSTRUCT, DESCRIBE, i ASK. Za potrebe kreiranja i manipulacije grafovima znanja rabe se CONSTRUCT i DESCRIBE vrste upita, dok SELECT i ASK upiti imaju drugačiju svrhu. SELECT upiti odgovaraju na upite u tabličnom formatu, što je korisno za analizu i izradu različitih dijagrama, poput stupičastog grafikona (*Bar Chart*), 'Torta' dijagrama (*Pie Chart*), linijskih dijagrama (*Line Chart*), da spomenemo samo neke, ali koji se ne mogu ubrojiti u grafove znanja, dok ASK upiti vraćaju jednostavan odgovor o postojanju određenih obrazaca u grafu, potvrđujući uvjete definirane upitom. Iako ovi upiti ne generiraju nove grafove znanja, služe kao važne alatke za pregled i provjeru podataka unutar postojećih grafova.

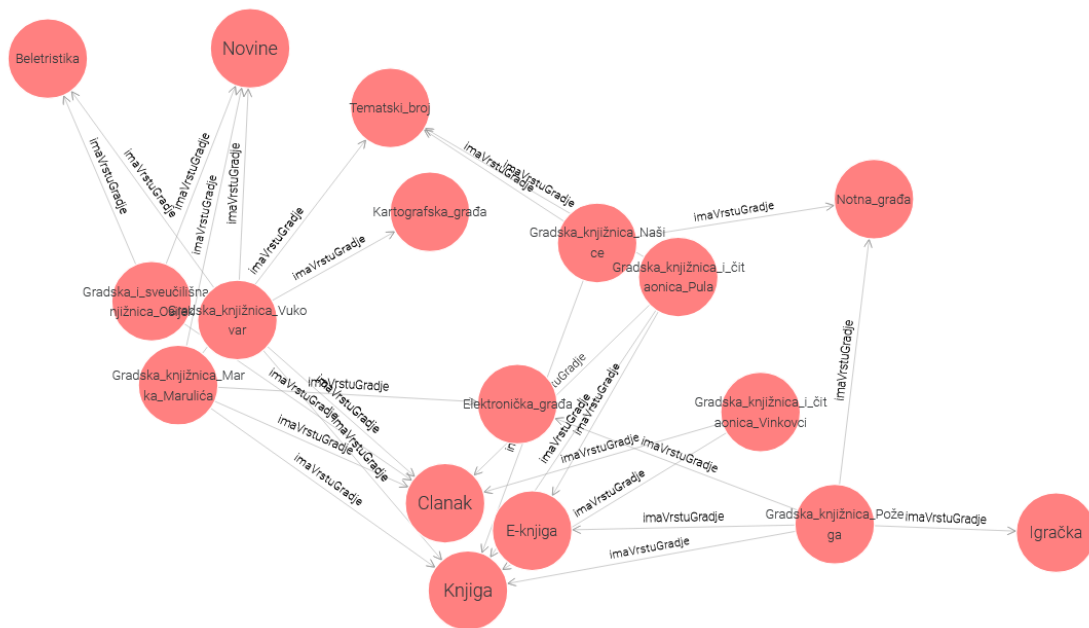
U nastavku rada fokusirat ćemo se na upotrebu CONSTRUCT i DESCRIBE upita, s obzirom na njihov ključni značaj u procesu izgradnje i prikaza grafova znanja. CONSTRUCT upiti

omogućuju generiranje novih RDF grafova na temelju definiranih obrazaca, čime se aktivno proširuje graf znanja. S druge strane, DESCRIBE upiti pružaju cjelovit uvid u podatke povezane s određenim resursom, vraćajući sve relevantne RDF triplete u obliku grafa.

U nastavku se donosi primjer CONSTRUCT upita koji povezuje knjižnice s vrstama građe.

```
PREFIX : <https://oziz.ffos.hr/diplomski#>
CONSTRUCT {
?knjiznica :imaVrstuGradje ?vrstaGradje .
}
WHERE {
?knjiznica a :NarodnaKnjiznica . ?knjiznica :imaVrstuGradje ?vrstaGradje .
}
```

Graf znanja koji ovaj upit generira prikazan je na Slici 17.



Slika 17 – Primjer grafa znanja temeljenog na SPARQL CONSTRUCT upitu.

Graf znanja u programu GraphDB moguće je generirati i putem SPARQL DESCRIBE upita. Putem DESCRIBE upita, primjerice, mogu se dobiti sve informacije o specifičnoj knjižnici koja je identificirana određenim URI/IRI-jem. U nastavku se donosi primjer DESCRIBE upita:

```
PREFIX : <https://oziz.ffos.hr/diplomski#>
DESCRIBE <https://oziz.ffos.hr/diplomski#Gradska_knjiznica_i_čit_aonica_Vinkovci>
```


6. Zaključak

U ovome radu istražena je važnost semantičkih grafičkih baza podataka u kontekstu suvremenog digitalnog okruženja, u kojem se količina i složenost informacija kontinuirano povećava. Prikazano je kako se tradicionalni sustavi za upravljanje podacima suočavaju s izazovima kada je riječ o organizaciji i povezivanju podataka na smislen način. Semantički web i grafovi znanja predstavljaju odgovor na izazove suvremenog digitalnog okruženja, nudeći okvir koji omogućava bolje razumijevanje i korištenje podataka. U drugom poglavlju predstavljen je pregled semantičkog weba, s objašnjenjem ključnih komponenti, poput slojeva i tehnologija koje omogućuju njegovu implementaciju. Treće poglavlje detaljno se bavilo konceptom ontologije, koja je ključna za strukturu i organizaciju podataka unutar semantičkih sustava. Četvrto poglavlje usporedilo je relacijske i grafičke baze podataka, s posebnim naglaskom na prednosti grafičkih baza podataka u radu sa složenim i povezanim podacima. U istom poglavlju pružen je i teorijski okvir za razumijevanje i primjenu grafova znanja kao temeljne tehnologije u razvoju inteligentnih sustava sposobnih za napredno pretraživanje i analizu podataka. Peto poglavlje prikazalo je praktičnu primjenu svih prethodnih teorijskih koncepata kroz izradu grafa znanja. Detaljno su opisani koraci u izradi skupa povezanih podataka, ontologije i konačne implementacije grafa znanja, čime je prikazana snaga i fleksibilnost semantičkih grafičkih baza podataka u stvarnim primjenama.

Na koncu, ovaj rad zaključuje da semantičke grafičke baze podataka predstavljaju značajan korak naprijed u organizaciji i analizi podataka u digitalnom okruženju. Njihova sposobnost povezivanja i smislenog organiziranja podataka ima potencijal transformirati način na koji pristupamo informacijama, omogućujući stvaranje inteligentnih sustava koji se mogu prilagoditi potrebama korisnika. Nadalje, rezultati rada ukazuju na veliki potencijal daljnjeg razvoja i primjene semantičkih grafičkih baza podataka u raznim područjima, uključujući poslovne sustave, znanstvena istraživanja i razvoj umjetne inteligencije. Ovaj rad doprinosi širem razumijevanju važnosti semantičkih tehnologija i njihove uloge u budućem razvoju informacijskih sustava. Preporučuje se daljnje istraživanje u ovom području, posebno u kontekstu integracije ovih tehnologija u postojeće informacijske sustave i njihove primjene u različitim industrijama. Također se predlaže daljnje istraživanje u smjeru optimizacije i skalabilnosti grafova znanja te njihove primjene u realnim okruženjima, što bi moglo dodatno unaprijediti učinkovitost upravljanja informacijama i donošenja odluka temeljenih na podacima.

Popis literature

1. Batra, Shalini; Tyagi, Cahru. Comparative Analysis of Relational And Graph Databases. // International Journal of Soft Computing and Engineering (IJSCE) 2, 2(2012), str. 509-512. URL: <https://www.semanticscholar.org/paper/Comparative-Analysis-of-Relational-Databases-and-Tyagi-Batra/f6bbca68da24d4e4708b58f01ad844bba00dfb17> (2024-07-23)
2. Batra, Shalini; Tyagi, Charu. Comparative Analysis of Relational And Graph Databases. // International Journal of Soft Computing and Engineering (IJSCE) 2, 2(2012), str. 509-512. URL: https://www.academia.edu/2891354/Comparative_Analysis_of_Relational_And_Graph_Databases (2024-07-14)
3. Berners-Lee, T., Hendler, J., Lassila, O., et al. The semantic web. // Scientific American 284, 5(2001), str. 28–37. URL: 10.1038/scientificamerican052001-yL7Vw7HIOZ4iSjlnEeVsJ (2024-07-03)
4. Berners-Lee, Tim. Linked data, 2006. URL: <https://www.w3.org/DesignIssues/LinkedData.html> (2024-07-04)
5. Berners-Lee, Tim; Bizer, Christian; Heath, Tom. Linked Data: The Story so Far. // International journal on Semantic Web and information systems, 5(2009), str. 1-22. URL: <https://ideas.repec.org/a/igg/jswis0/v5y2009i3p1-22.html> (2024-06-30)
6. Database. // Encyclopedia Britannica. URL: <https://www.britannica.com/technology/database> (2024-08-01)
7. DBpedija. URL: <https://www.dbpedia.org/> (2024-08-27)
8. Decker, Stefan et al. The Semantic Web: the roles of XML and RDF. // IEEE Internet Computing 4, 5(2000), str. 63-73. URL: 10.1109/4236.877487 (2024-07-28)
9. Ding, L. et al. Using Ontologies in the Semantic Web: A Survey. // Springer, 14(2007), str. 79–113. URL: https://doi.org/10.1007/978-0-387-37022-4_4 (2024-07-10)
10. Domingue, John; Fensel, Dieter; Hendler, James A. Introduction to the Semantic Web Technologies. // Handbook of Semantic Web Technologies / Domingue, John; Fensel, Dieter; Hendler, James A. New York: Springer Publishing Company, 2011. Str. 1-41. URL: <https://link.springer.com/referencework/10.1007/978-3-540-92913-0> (2024-07-21)

11. Gopal, Pandey. „The Semantic Web: An Introduction and Issues“ // International Journal of Engineering Research and Applications, 1(2012), str. 780 – 786. URL: https://www.academia.edu/1959581/IJERA_www_ijera_com_ (2024-07-04)
12. GraphDB. URL: <https://graphdb.ontotext.com/> (2024-08-12)
13. GraphDB: Documentation. URL: <https://graphdb.ontotext.com/documentation/10.7/index.html> (2024-08-12)
14. Guia, José; Gonçalves Soares, Valéria; Bernardino, Jorge. Graph Databases: Neo4j Analysis. // International Conference on Enterprise Information Systems, 2017, str. 351-356. URL: <https://www.semanticscholar.org/paper/Graph-Databases%3A-Neo4j-Analysis-Guia-Soares/c867c372ed78e27c53feaf788e74aa7506302bbe#citing-papers> (2024-07-24)
15. Hadi Hor, A.E. et al. A Semantic Graph Database for Bim-Gis Integrated Information Model for An Intelligent Urban Mobility Web Application. // ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 4, 4(2018), str. 89-96. URL: <https://ui.adsabs.harvard.edu/abs/2018ISPRAn.4.4...89H/abstract> (2024-07-15)
16. Ham, Kelli. OpenRefine: Free, open-source tool for cleaning and transforming data. // Journal of the Medical Library Association: JMLA, 101, 3(2013), str. 233–234. URL: 10.3163/1536-5050.101.3.020 (2024-08-05)
17. Matthews, Brian. Semantic Web Technologies. // E-learning and Digital Media 6, 6(2005), str. 1-21. URL: https://www.researchgate.net/publication/30408878_Semantic_Web_Technologies (2024-07-04)
18. Miller, Eric; Schloss, Bob. Resource Description Framework (RDF): Model and Syntax, 1997. URL: <https://www.w3.org/TR/WD-rdf-syntax-971002/> (2024-07-14)
19. Ontotext. What is SPARQL. URL: <https://www.ontotext.com/knowledgehub/fundamentals/what-is-sparql/> (2024-08-01)
20. OpenRefine. Documentation. URL: <https://openrefine.org/documentation.html> (2022-08-05)
21. OWL Web Ontology Language Guide. // W3C. URL: <https://www.w3.org/TR/owl-guide/> (2024-08-27)

22. Protégé. URL: <https://protegeproject.github.io/protege/> (2024-08-07)
23. RDF 1.1 Concepts and Abstract Syntax. // W3C. URL: <https://www.w3.org/TR/rdf-concepts/#section-rdf-graph> (2024-08-28)
24. *The Linked Open Data Cloud*. URL: <https://lod-cloud.net/> (2024-07-24)
25. Valle, Emanuele Della; Ceri, Stefano. Querying the Semantic Web: SPARQL . // Handbook of Semantic Web Technologies / Domingue, John; Fensel, Dieter; Hendler, James A. New York: Springer Publishing Company, 2011. Str. 299-363. URL: <https://link.springer.com/referencework/10.1007/978-3-540-92913-0> (2024-07-21)
26. Wilson. M.; Matthews, B. The semantic Web: prospects and challenges. // 7th International Baltic Conference on Databases and Information Systems, 2006, str. 26-29. URL: 10.1109/DBIS.2006.1678469. (2024-07-21)
27. Zou, Xiaohan. A Survey on Application of Knowledge Graph. // Journal of Physics: Conference Series 1487, 4(2020), str. 1-12. URL: <https://iopscience.iop.org/article/10.1088/1742-6596/1487/1/012016> (2024-07-10)