Ontologije u arhivistici: pregled, dizajn, implementacija i održavanje

Arijan Rajh

Filozofski fakultet Sveučilišta u Zagrebu
Ivana Lučića 3
HR – 10 000 Zagreb
arajh@m.ffzg.hr

Izvorni znanstveni rad
Primljeno: 12. 4. 2024.
Prihvaćeno: 12. 7. 2024.
004.822
111
DOI: 10.58565/vda.5.1.1

Sažetak

U radu se analiziraju ontologije arhivističke domene, odnosno formalne konceptualizacije znanja o arhivskim entitetima, od raznih prijašnjih autorskih i projektnih ontologija do ontologije međunarodne norme Zapisi u kontekstima (RIC-O). Autor nadalje predstavlja metodologije razvoja ontologija i prikazuje odabrani način razvoja na primjeru ontologije arhivističke etike. Naposljetku, autor diskutira o načinima korištenja ontologija u arhivistici te donosi zaključak i preporuke vezane uz mogući smjer razvoja i održavanja ontologija.

Ključne riječi: arhivski opis, deskriptivna logika, modeliranje ontologija, povezani podaci, semantički web, zapisi u kontekstima

1. Uvod

Ovu temu treba smjestiti u širi kontekst semantičkog weba. Semantički web čini skup koncepata i decentraliziranih kooperativnih[1] praksa kojima se omogućuje softversko dohvaćanje i obrada informacija o entitetima, bez potrebe za izravnim ljudskim posredovanjem. Ti koncepti uključuju označavanje informacija o entitetima i pridruživanje jasno određenih značenja.[2] Povezani podaci vrsta su implementacije koncepata semantičkog weba[3] koja koristi standardizirane načine identifikacije imenovanih entiteta (engl. *Uniform Resource Identifiers*, URI)[4] i RDF model za

1 Cleborne C. Maddux, „The Semantic Web and educational technology", *Educational technology* 48, 1 (2008), 5, pristupljeno 18. 6. 2024. https://www.jstor.org/stable/44429538.

2 Tim Berners-Lee, James Hendler, Ora Lassila. „The Semantic Web: a new form of web content that is meaningful for computers will unleash a revolution of new possibilities", *Scientific American* 5 (2001), pristupljeno 29. 3. 2024. https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf.

3 Tom Heath, Christian Bizer, „Principles of Linked Data", u: *Linked Data: evolving the Web into a global data space,* ur. Tom Heath, Christian Bizer (Cham.: Springer, 2011), 7.

4 Berners-Lee, Tim. Linked Data. (2006), pristupljeno 18. 3. 2024. https://www.w3.org/DesignIssues/LinkedData.html.

razmjenu podataka o entitetima putem interneta (engl. *Resource Description Framework*, RDF).[5] Da bi se informacije o entitetima mogle dohvaćati i koristiti, RDF model propisuje njihovo strukturiranje u oblik rečenica ili iskaza. Iskazi, nazvani tripletima, imaju subjekte, predikate i objekte. Objekt jednog tripleta može postati subjektom drugog i tako se informacije o entitetima povezuju (primjer posljednja dva retka Tablice 1).

Tablica 1. Razlike između klasičnog arhivističkog opisa i opisa koji trebaju obraditi softveri

Jezik koji razumiju arhivisti				**Jezik koji razumiju njihova računala**		
Subjekt	Predikat	Objekt		Subjekt	Predikat	Objekt
Ferdo Šišić	zanimanje	povjesničar		<https://dbpedia.org/page/Ferdo_%C5%A0i%C5%A1i%C4%87>	rico:OccupationType	”povjesničar”
arhivska kutija 1 u ostavštini Jaroslav Šidak	sadržaj	Kutija s rukopisima radova i tiskanim sadržajem.		<file:/uploaded/generated/Kutija01>	rico:scopeAndContent	”Kutija s rukopisima i tiskanim sadržajem.”
prvi skup dokumenata u kutiji	ima količinu	40 listova	→	<file:/uploaded/generated/SkupDokumenata01-14>	rico:hasExtent	a rico:InstantiationExtent ; rico:quantity ”40” ; rico:unitOfMeasurement ”List”
Stockholm	je bio mjesto	10. međunar. kongresa povijesnih znanosti		<https://dbpedia.org/page/Stockholm>	rico:isOrWasLocationOf	”10. međunarodni kongres povijesnih znanosti”
10. međunar. kongres pov. zn.	se događao	21. – 28. 8. 1960.		<file:/uploaded/generated/XIthInternationalCongressOfHistoricalSciences>	rico:occuredAtDate	”21. – 28. 8. 1960.”

Predikati mogu biti svojstva subjekata ili odnosi kojima su subjekti i objekti povezani. Subjekti su entiteti označeni URI-jima, a objekti su točno određene vrijednosti svojstava subjekata koje nazivamo literalima (brojevi, nizovi znakova, datumi) ili drugi entiteti identificirani URI-jima.

Povezivanje podataka različite provenijencije o nekom segmentu svijeta uvelike je olakšano postojanjem ontologije o tom segmentu svijeta. Ontologije su „repre-

5 W3C. Resource Description Framework (RDF). (2014a).

zentativni vokabulari [odnosno] formalne reprezentacije znanja neke domene",[6] „zbirke informacija"[7] koje su pažljivo definirane i koje mogu referirati različiti informacijski sustavi. „[U] informacijskim znanostima ontologija je artefakt inženjerstva, obično konceptualni model (nekog aspekta) svijeta, tipično formaliziran kao logička teorija."[8] Thomas Grubner ontologije je inicijalno odredio kao formalne reprezentacije konceptualizacije, a Rudi Studer kao računalno čitljive „formalne, eksplicitne specifikacije [određene] dijeljene konceptualizacije".[9] Konceptualizacija je misaona operacija nužna za poimanje nekog segmenta svijeta, pojednostavljenje tog dijela svijeta, odnosno vrsta tumačenja sa svrhom razumijevanja. Dijeljena konceptualizacija je ona koja je dana nekoj zajednici na korištenje. Entiteti svijeta u ontologijama oblikuju se u koncepte i njihove relacije. Entitete općenite razine koji dijele ista svojstva i postoje unutar hijerarhijskih struktura zovemo klasama. Klase imaju svoje nad- i podklase, primjerice u ontologiji norme Zapisi u kontekstima (engl. *Records in Contexts Ontology,* dalje u tekstu: RIC-O), klasa Agent ima podklase Osoba, Grupa, Položaj i Mehanizam, a podklasa Grupa ima podklase Obitelj i Korporativno tijelo. Prema klasama se određuju individuali, konkretni entiteti koji prema načelima semantičkog weba mogu stajati na poziciji subjekata i objekata tripleta. Ontologije ne propisuju samo klase i njihove individuale nego i svojstva. Svojstva mogu biti podatkovna, koja opisuju neki entitet, ili objektna, koja povezuju dva entiteta. Hijerarhijski odnosi klasa nisu jedini odnosi koji se mogu definirati ontologijama, postoje i razni drugi aksiomi. Na primjer, za objektno svojstvo isAgentAssociatedWithAgent možemo izraziti aksiom da je ono simetrično jer polazi od agenta i završava na drugom agentu. Svojstva se koriste kao predikati tripleta; njihove domene su subjekti, a dosezi objekti. Primjerice, za RIC-O objektno svojstvo rico:isContentTypeOf, koje se u opisu koristi kao predikat, domena je klasa VrstaSadržaja (rico:ContentType), a doseg je Dokument (rico:Record). Ovo znači da će u opisu koji se oslanja na RIC-O određena vrsta sadržaja biti subjektom, a određeni dokument koji je te vrste sadržaja objektom tripleta. Ovo vrijedi za

6 Thomas R. Grubner, „A translation approach to portable ontology specifications", *Knowledge acqusition* 5 (1993), 199. Ontološke se „definicije imena entiteta u univerzumu diskursa (klase, odnosi, funkcije ili drugi objekti) pridružuju ljudski čitljivom tekstu koji opisuje što ta imena znače i formalnim aksiomima koji ograničavaju interpretaciju i formiraju korištenje ovih termina"., Thomas R. Grubner, „Toward principles for the design of ontologies used for knowledge sharing", *International journal of human-computer studies* 43, 5–6 (1995), 908. U tom radu Grubner o ontologijama piše kao o „jezicima za komunikaciju različitih alata u distribuiranim, heteronomnim okolinama". *Isto*, 912.

7 Berners-Lee, Hendler, Lassila, „The Semantic Web: a new form of web content.

8 Ian Horrocks, „What are ontologies good for?", u: *Evolution of semantic systems*, ur. Bernd Olaf Küppers, Udo Hahn, Stefan Artmann. Berlin, Heidelberg: Springer, 2013., 186.

9 Rudi Studer, Richard Benjamins, Dieter Fensel, „Knowledge engineering: principles and methods", u: *Data & knowledge engineering* 25, 1-2 (1998), 25.

objektna svojstva, dok je kod podatkovnih svojstava tipičan objekt slobodan unos ili odabrana vrijednost iz liste propisanih vrijednosti (primjerice iz liste https://id.loc.gov/vocabulary/contentTypes.html). Ontologije, dakle, definiraju značenja klasa entiteta te propisuju njihova svojstva, odnose i ograničenja. Ta su značenja istodobno raspoloživa za opis ili upravljanje entitetima određenih domena i neovisna o tome upotrebljava li ih konkretan informacijski sustav ili aplikacija.[10]

2. Ontologije u arhivistici

Arhivski opis je proces sažimanja sadržaja, interpretacije i izrade informacija o gradivu te proizvod tog procesa u obliku klasičnoga obavijesnog pomagala, računalno čitljivog obavijesnog pomagala ili samostalnog opisa priređenoga za računalnu obradu. Arhivski opis mora se osloniti na međunarodne norme da bi bio razumljiv članovima globalne arhivističke zajednice. Međunarodne norme opisa u zadnjih gotovo tridesetak godina, od 1994. do studenoga 2023. godine, bile su ISAD(G) norma i povezane norme ISAAR(CPF), ISDF i ISDIAH. ISAD(G) normirao je opis arhivskoga gradiva, ISAAR(CPF) stvaratelja, ISDF poslovnih funkcija, a ISDIAH opis ustanova koje čuvaju arhivsko gradivo. Ove su se norme za računalnu obradu i korištenje obavijesnih pomagala i opisa u različitim sustavima i aplikacijama oslanjale na druge, dodatne, norme: EAD (engl. *Encoded Archival Description*), EAC (engl. *Encoded Archival Context*) i EAG (engl. *Encoded Archival Guide*). EAD i takvi formati oslanjali su se na XML normu serijaliziranja metapodataka za razmjenu i obradu u računalnim sustavima. EAD je ISAD(G)-kompatibilan opis gradiva formatiran u XML datoteku koja služi za razmjenu tog opisa između više aplikacija. EAC se oslanja na ISAAR(CPF) i priprema za razmjenu opis stvaratelja, a EAG čini isto za opis ustanova s arhivskim gradivom, pripremljenom prema ISDIAH normi.

2.1. RIC-O

Od 2012. Međunarodno arhivsko vijeće i Stručna grupa za arhivski opis na čelu s Danielom Pittijem (engl. *Expert Grup On Archival Description*, EGAD) radili su na novoj jedinstvenoj normi opisa koja bi objedinila mogućnosti opisivanja svih arhivskih entiteta i osuvremenila struku povezujući arhivski opis s mehanizmima semantičkoga weba i pružajući arhivistima mogućnosti korištenja novih tehnologija. Nova norma, nazvana Zapisi u kontekstima (engl. *Records in Contexts,* RIC),

10 Grubner, „Towards principles for the design of ontologies", 913.

sastoji se od uvodnog dijela (*RIC Foundations of Archival Description*, RIC-FAD), konceptualnog modela entiteta (RIC-CM), ontologije (RIC-O) i smjernica za implementaciju. Prvi nacrt bio je predstavljen 2016., a prva konačna inačica bila je objavljena u studenome 2023. Smjernice za implementaciju trebale bi biti objavljene u bliskoj budućnosti. Konceptualni model opisuje entitete, njihova svojstva i relacije. Ontologija te entitete razrađuje većinom u klase, a atribute entiteta bilo u klase, bilo u podatkovna svojstva. Primjer razrade atributa konceptualnog modela u klasu ontologije je transformacija atributa *documentary form type* u klasu *DocumentaryFormType*, a primjer kada je atribut konceptualnog modela ostao (slično) podatkovno svojstvo je *creationDate*. Podatkovna svojstva imaju vrijednost literala slobodnog unosa ili vrijednosti s lista termina. Relacije iz konceptualnog modela uobličene su u ontologiji u objektna svojstva (npr. relacija *has or had location* iz RIC-CM-a u objektno svojstvo *hasOrHadLocation* RIC-O-a). Ontologija pritom definira i pravila odnosa između različitih klasa i osigurava dosljednost onoga što je modelirano. Ontologije služe da bi najrazličitije implementacije, u našem slučaju opisi arhivskih entiteta u raznim sustavima, kao i opisi formatirani ili izvezeni u samostalne datoteke, mogle koristiti iste pojmove i na temelju toga se povezivati. RIC se ne oslanja na EAD, EAC i EAG formate, već se izravno serijalizira u XML ili RDF datoteke koje mogu obrađivati informacijski sustavi. Dok ISAD(G) obitelj normi računa na izradu klasičnih obavijesnih pomagala, tiskanih ili objavljenih na internetu, a EAD, EAC i EAG na izradu računalno čitljivih obavijesnih pomagala, RIC pretpostavlja izradu samostalnih strojno čitljivih opisa. RIC-O uključuje više od stotinu klasa i podklasa, šezdesetak podatkovnih svojstava i četiristotinjak objektnih svojstava. Neka su se od podatkovnih svojstava možda mogla izuzeti (npr. *measure* i njegova podsvojstva) i umjesto toga se u četvrtom dijelu norme moglo pozvati na postojeće ontologije iz drugih domena koje su već obradile ta svojstva, no taj četvrti dio norme nije još objavljen i oba su pristupa dozvoljena. Uz RIC-O, kao službenu ontologiju međunarodne stručne opisne norme, javile su se i druge ontologije povezane s arhivskim gradivom. One su za ovaj rad bile analizirane pregledom literature i njihovih OWL ili Turtle datoteka, ako su te datoteke bile dostupne. Turtle je jedan od oblika (serijalizacija) RDF modela podataka, tekstni prikaz SPO tripleta koji podržava sintaksu njihova skraćenog zapisivanja, a OWL je jezik koji proširuje mogućnosti RDF-a logičkim mehanizmima.

2.2. Prijašnje i alternativne ontologije u arhivistici

Jedna od početnih ontologija koja je bila povezana s arhivskim gradivom ona je autora Daphne Kyriaki-Manessi i Markosa Dendrinosa iz 2014. Autori su radili na

rješavanju problema organizacije digitalnoga gradiva Atenskoga tehnološko-edukacijskog instituta i na dizajniranju ontologije koja bi im u tome mogla pomoći. Ova ontologija zapravo i nije u svojoj osnovi arhivistička jer joj je primarna domena viša edukacija u području tehnologija, pa se najviše oslanja na VIVO i HERO ontologije iz domene obrazovanja. Sljedeće ontologije mogu se smatrati arhivističkim: ARKIVO, ArDo i ArCO, a KAM ontologija podjednako je arhivistička, knjižničarska i muzeološka. Ontologija ARKIVO dosta se oslanja na druge ontologije: *Dublin Core*, *Bibliographic Ontology* (BIBO) i *Schema.org* entitete te *Frend of a Frend* ontologiju za entitete osoba, *Linking open description of events* ontologiju za događaje i Geonames za lokacije.[11] ARKIVO se koristila za opis dokumentacije poljske provenijencije koja je uoči II. svjetskog rata bila prenesena u SAD (Institut Jozef Pilsudski). Archive Dynamics Ontology (ArDO) također se u većoj mjeri oslanjala na druge ontologije, a autori su je koristili za opis gradiva njemačke Weimarske republike (1919. – 1933.) nakon što su ocijenili da je to gradivo previše specifično za opis uz korištenje tada dostupne inačice RIC-O ontologije.[12] ArCo ontologija (tal. *Architettura della Conoscenza*) ontologija je talijanskoga Ministarstva kulturne baštine i ima pet zasebnih arhivskih klasa (ArhivskaZbirka, ArhivskaJedinica, RazinaJedinice, ObavijesnoPomagalo, KriterijSređivanja), dva arhivska objektna svojstva (imaRazinuJedinice, imaObavijesnoPomagalo) i tri podatkovna svojstva (sustavSređivanja, količina, dužniProstorPolica), ali nadograđuje se na klase i svojstva drugih povezanih vokabulara iz šire domene kulturne i prirodne baštine.[13] Nama bliži primjer ontologije koja se može koristiti za arhivsko gradivo, uz gradivo drugih AKM domena, jest ontologija Pravilnika za opis i pristup građi u knjižnicama, arhivima i muzejima (https://pravilnik.kam.hr//). Ontologija Pravilnika KAM, koja je trenutno u završnoj fazi testiranja, spaja knjižničarske entitete IFLA LRM modela, s arhivskim entitetima RIC-a i muzejskim entitetima modela CIDOC-CRM.[14] U Tablici 2 prikazane su razlike u opisima jedne vrste gradiva

---

11 Laura Pandolfo i dr., "Towards an ontology for describing archival resources.", u: *Proceedings of the second workshop on humanities in the Semantic Web (WHiSe II) co-located with the 16th International Semantic Web conference.* Beč, WHiSe@ISWC, 2017, 112-113. Ontologija je popraćena literaturom iz 2017., no još se nalazi u nacrtu inačice 0.3, što vjerojatno znači da se njezin razvoj prekinuo. Informacije o njoj i sama datoteka s ontologijom mogu se pronaći na GitHub stranicama https://github.com/ArkivoTeam/ARKIVO.

12 Oleksandra Vsesviatska i dr. „ArDO: an ontology to describe the dynamics of multimedia archival records", u: *Proceedings of the 36th annual ACM symposium on applied computing (SAC '21).* New York: Association for computing machinery, 2021.

13 GitHub stranica ArCo ontologije, podstranica s archive.owl datotekom, pristupljeno 2.4.2024., https://github.com/ICCD-MiBACT/ArCo/blob/master/ArCo-release/ontologie/archive/archive.owl.

14 Mirna Willer i dr., „Ontologija Pravilnika za opis i pristup građi u knjižnicama, arhivima i muzejima", *Vjesnik bibliotekara Hrvatske* 66, 2 (2023), 46.

koje proizlaze iz činjenice da ontologije mogu na različite načine konceptualizirati neki zadani segment svijeta.

Tablica 2. Prikaz razlika ontologija na primjeru opisa filmskog gradiva

RICO	ARKIVO	ArDO	ArCO	KAM
filmsko gradivo – nema zasebnu klasu, već klasu Content Type (rico:ContentType) ili se može koristiti objektno svojstvo rico: hasContentOfType	filmsko gradivo – nema zasebnu klasu, koristi se BIBO klasa AV dokumenta (bibo:AudioVisualDocument)	filmsko gradivo – zasebna klasa (podklasa BIBO AV dokumenta, ardo:Film)	filmsko gradivo – zasebna klasa FilmDocumentaton (podklasa klase Documentation, onda klase Object pa klase Entity)	filmsko gradivo – nema zasebne klase, no ima specifičnih objektnih svojstava povezanih s filmom kao klasom entiteta djelo (npr. imaOsobuFilmskogRedatelja)

3. Korištenje ontologija i njihova izrada

Iako RIC-O, dakle, nije jedina arhivistička ontologija, ona se danas može smatrati najrelevantnijom za arhiviste jer iza njezine izrade stoji obimniji grupni rad i jer je podržava Međunarodno arhivsko vijeće.

3.1. Korištenje ontologija

Dobar primjer korištenja ontologija u arhivistici, konkretno korištenja RIC-O ontologije, jest vježba osmišljena za studente druge godine prijediplomskog studija Informacijskih i komunikacijskih znanosti, na kolegiju Sređivanje i opis arhivskoga gradiva u ljetnom semestru ak. god. 2023./2024.[15] Vježba se izvodila u prvih pet tjedana nastave na kolegiju, od kraja veljače. Svrha vježbe bila je podučiti studente o normi opisa RIC, njezinoj ontologiji i normama semantičkog weba te propitati dvije hipoteze – prvu o primjerenosti RIC norme za suradnju pri opisivanju gradiva i drugu o mogućnostima dobivanja novih informacija o arhivskim entitetima iz tako priređenih opisa. Studenti su opisivali ostavštinu prof. Jaroslava Šidaka. Ona se sastoji od 22 arhivske kutije, od čega su studenti obradili 14 arhivskih kutija. Nastavnik im je na temelju RIC-CM i RIC-O dijelova norme te Specifikacije

15 Arian Rajh, „Archival description turns truly collaborative: an excercise in Records in Contexts standard", Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja, Radenci 16. 5. 2024., prezentacija na konferenciji, pristupljeno 18. 6. 2024. http://www.pokarh-mb.si/storage/app/media/Radenci%202024/Predstavitve_2024/Rajh_Adian_Radenci_2024.pdf. Rad je prihvaćen i za objavu u narednom broju časopisa Moderna arhivistika za 2024. godinu.

metapodataka za opis dokumentarnog i arhivskog gradiva, koju je objavio Hrvatski državni arhiv u siječnju 2022., izradio datoteku-predložak opisa kutije u Turtle formatu te dvije datoteke s opisom cijele ostavštine i stvaratelja. Uz to su studenti od nastavnika dobili i konvenciju imenovanja entiteta i bilježenja opisa kako bi njihovi opisi bili usklađeni. Analizirali su sadržaj svojih kutija i opisali ga u obliku RDF Turtle datoteke. [16] U njoj su naveli metapodatke o logičkom i fizičkom aspektu kutije, skupova dokumenata i dokumenata u kutiji te o agentima, lokacijama i aktivnostima koje se spominju u dokumentaciji. Osmero studenata uspjelo je u roku samostalno izraditi i predati validne datoteke s opisom arhivske kutije. Nastavnik je te datoteke učitao u graf bazu podataka[17] i povezao sve predane opise kutija s opisima stvaratelja i ostavštine kao dokumentarne cjeline. Time se dokazala prva hipoteza o podršci RIC-a za kolaborativan rad pri opisivanju gradiva. Nakon toga nastavnik je izradio SPARQL upite. SPARQL je jezik za postavljanje upita na graf baze podataka. Jedan je upit poslužio za propitivanje o povezanosti agenata, aktivnosti u kojima su osobe sudjelovale, datuma aktivnosti i identifikatora dokumenata ili skupova dokumenata u kojima se te aktivnosti spominju. Drugi SPARQL upit bio je konstruiran kako bi se dobila posljednja godina nastanka obrađenih dokumenata jer je naslov ostavštine u pogledu te informacije neodređen (navodi se raspon godina od 1933. do 1980-ih). Vježba se planira nastaviti i sljedeće ak. god. kako bi se obradila cijela ostavština Jaroslava Šidaka. Rezultati ovih i sličnih SPARQL upita idu u prilog dokazivanja druge hipoteze, formulirane za navedenu vježbu, jer su na relativno jednostavan način bile dobivene informacije koje nisu bile evidentne iz parcijalnih opisa kutija: informacije o kompleksnim odnosima raznih entiteta dokumentacije i onih koji se u dokumentaciji spominju, vremenski entiteti itd. Oslanjanje na RIC-O uvelike je olakšalo i ubrzalo kolaborativan rad i konstrukciju novih, složenih, informacija o gradivu. Kombinacija korištenja povezanih podataka za opis i SPARQL upita omogućuje raznoliko korištenje arhivskog gradiva, bez obzira na to jesu li arhivisti koji su izradili opis takvo korištenje predvidjeli.

3.2. Modeliranje ontologija

Ontologije se mogu dizajnirati ili modelirati različitim metodologijama, a sam pristup modeliranju ovisi o tome imamo li ili nemamo jasno definirane koncepte i entitete u domeni. Prvi pristup naziva se pristupom utemeljenim na znanju, a drugi

16 RDF Turtle datoteke sa svim opisima, onima koje je izradio nastavnik i onima koje su izradili studenti, kao i SPARQL upite moguće je pregledati na GitHub stranici: https://github.com/rarianSp/soag/tree/main/Jarosav%20Sidak.

17 Korišten je sustav Ontotext GraphDB 10.6.2. (inačica koja se ne plaća).

evaluacijskim pristupom.[18] U arhivističkoj domeni postoje bliskoznačni koncepti, što se može uočiti kod, primjerice, pojmova dokumenta i zapisa ili pojma izlučivanja, koji je u širem značenju blizak pojam vrednovanju, a u užem uništavanju. Zbog toga je za modeliranje arhivističkih ontologija pogodniji evaluacijski pristup. Ontologija ovdje, na neki način, rafinira diskurs.

U literaturi se mogu pronaći brojne metodologije modeliranja ontologija. Oyelade navodi Simplified Agile Methodology for Ontology Development (SAMOD), Pattern-based Ontology Development (PBOD), Menthodology, KACTUS, Horrocksovu metodu, DOGMA metodologiju te Fuzzy, Expert2OWL i TOVE. Iqbal i dr. uz većinu ovih metodologija navode i Enterprize ontology metodologiju, IDEF5, CYC, Common KAD, On-To_knowledge, ONIONS, MENELAS, UPON i 101 metodologiju te na kraju zaključuju da „nijedna metodologija nije potpuno zrela".[19] Banu razlikuje proces izrade ontologije iz početka i proces izrade ontologije preuzimanjem dijelova postojećih ontologija. Prvi proces uključuje korake (1) određivanja opsega domene, (2) organizacije odnosa koncepata, (3) određivanja objektnih svojstava, (4) testiranja ontologije, (5) određivanja podatkovnih svojstava i anotacija, (6) kreiranja individuala i (7) određivanja ograničenja svojstava.[20] U drugom procesu treba (1) identificirati domenu i osnovne entitete, (2) pretražiti relevantne ontologije, (3) analizirati ih, (4) sjediniti ih s entitetima iz prvog koraka, (5) dovršiti i (6) testirati ontologiju.[21] Alati koji se za modeliranje mogu koristiti jesu Ontolingua, Protégé, Swoop, OntoEdit, Apollo, OntoViz, WebVowl, SENSUSQ i drugi uređivači i alati za ontologije. Uz alate za uređivanje koriste se zaključivački alati (engl. *reasoner tools*) kao što su HermiT, Pellet i drugi. Od formata u koje se modelirane ontologije pohranjuju najčešće se koristi OWL format. Ontologiju je moguće, uz OWL format (engl. *web ontology language*), pohraniti i u RDF Turtle datoteke.

Grubner navodi pet odlika koje modelirana ontologija mora zadovoljavati: jasnoću i kompletnost definicija, koherentnost i usklađenost zaključaka s definicijama, proširivost novim terminima s obzirom na moguće upotrebe, nepristranost u kodiranju, koliko god moguća bila, kao i minimalno ontološko obvezivanje korisnika, odnosno definiranje samo nužnih klasa i ostavljanje korisnicima da

18 Ayesha Banu, „Ontologies for knowledge representation", u: *Semantic Web technologies: research and applications,* ur. Archana Patel, Narayan C. Debnath, Bharat Bhushan. Boca Raton: CRC Press, 2022., 229.

19 Rizwan Iqbal, Masrah Azirfah Azmi Murad, Aida Mustapha, Nurfadhlina Mohd Sharef, „An analysis of ontology engineering methodologies: a literature review", u: *Research journal of applied sciences, engineering and technology* 6, 16 (2013), 2999.

20 Banu, „Ontologies for Knowledge Representation", 229.

21 *Isto*, 233.

sami definiraju individuale.[22] Validacija se može provoditi korištenjem posebnih dodataka uređivačkih softvera za ontologije kao što je Protégé ili uz pomoć samostalnih alata za zaključivanje i skripata koje prolaze po datotekama ontologija. U koracima testiranja ontologije koriste se alati poput OntoCleana, HermiT-a i Pelleta. Primjerice, kod modeliranja ArDo metodologije, prema literaturi, korišten je OntoClean. U razvoju RIC-O ontologije EGAD-ov tim koji je na tome radio nije slijedio niti jednu formalnu metodologiju, no analizirao je ontologije koje su u odnosu na RIC-O općenitije (https://www.loa.istc.cnr.it/dolce/overview.html), ontologije srodnih AKM disciplina (LRM, FRBRoo) i povezanih domena (PREMIS, PROV-O). Modeliranje je bilo odrađeno u alatu Protégé.[23] U nastavku rada prikazat će se metoda modeliranja na primjeru jedne manje ontologije koju možemo svrstati u poddomenu arhivističke etike.

3.3. Modeliranje ontologije arhivističke etike

U ovom će se dijelu rada prikazati modeliranje arhivističke i etičke ontologije koristeći i kombinirajući naputke metodologije za izradu ontologije uz preuzimanje dijelova postojećih ontologija, koju opisuje Banu, s naputcima metodologije iz priručnika „Ontology Development 101". Prva je metodologija izabrana samo zato da se izdvoje klase koje već postoje u drugim ontologijama kako se ne bi ponavljale u ontologiji koja se modelira. Druga je metodologija bila izabrana zbog svojega naglaska na koracima formiranja klasa i svojstava – Ontology 101 priručnik donosi metodologiju koja ima korake: (1) određivanja domene i opsega uz pomoć kompetencijskih pitanja ili drugim sredstvima, (2) razmatranje upotrebe postojećih ontologija ili dijelova, (3) definiranje pojmova, (4) oblikovanje hijerarhije klasa, (5) određivanje svojstava klasa, (6) detaljno određivanje svojstava, njihovih domena i dosega te (7) oblikovanje instanci.[24] Obje metodologije koje su se razmatrale za modeliranje bile su objedinjene u jedinstveni proces od pet koraka (koraci A – E u daljnjem tekstu). Ontologija koja se ovdje modelira odnosi se na jednu moguću domenu arhivističke etike, a odabrana je jer ima razvijen konceptualni model opisan u radu prihvaćenom za objavu u Arhivskom vjesniku za 2024. godinu.[25] U tom bi

22 Grubner, „Toward principles for the design of ontologies", 909-910.

23 Prema informacijama dobivenima od Florence Clavaud, voditeljice RIC-O tima EGAD grupe, vidi https://groups.google.com/g/records_in_contexts_users/c/mhAkB6k-E9c/m/p1r0QZgVAgAJ.

24 Natalya Noy, Deborah Mcguinness, „Ontology Development 101: A guide to creating your first ontology", u: *Knowledge systems laboratory* 32 (2001), 4-11.

25 Arian Rajh, „Prilog arhivističkoj etici: osobine arhivista, perspektiva pravednosti i vrednovanje", rukopis prihvaćen za objavu u broju Arhivskog vjesnika za 2024.

smislu ontologija dizajnirana u ovome radu predstavljala nadogradnju konceptualnog modela iz rada u Arhivskom vjesniku, a konceptualni model predstavljao bi, obratno, stepenicu prema ontologiji. Upravo zbog ovoga ne bismo, kao u jednoj od definicija iz uvodnog dijela ovog rada, izjednačavali konceptualne modele i ontologije. I RIC i predložena domena arhivističke etike pokazuju manja, ali dovoljna, odstupanja između svojih konceptualnih modela i ontologija, koja mogu opravdati ovakvo mišljenje.

(A) Određivanje domene i opsega – ontologiju koja se ovdje modelira uvrštavamo, dakle, u poddomenu arhivističke etike. Središnji je koncept ove poddomene pravedno ponašanje. Konceptualni model deontološko-aretaičke etike, odnosno etike koja se oslanja na postupke i osobine, koji ovdje služi kao polazište modeliranja ontologije, oslanja se na Rawlsovu teoriju pravednosti, no zbog praktične izvedivosti njegovi se pojmovi „izvorne pozicije" i „vela neznanja"[26] zamjenjuju dijametralno suprotnim ishodištem, arhivističkom samorefleksivnošću, a principi pravednosti koji bi prema Rawlsu trebali osigurati maksimalnu korist za one s najslabijim društvenim pozicijama, empatijom i inkluzivnošću. Svemu navedenom dodaje se osobina transparentnosti koja dokumentira i priopćava ishode pravednog ponašanja. Osobine su djelomično formulirane prema teoriji arhivističkih vrlina Scotta Clinea. Nazivi entiteta ontologije (rdfs:label) pripremljeni su na hrvatskom jeziku, a novodizajnirana ontologija može se pronaći na GitHub stranici.[27]

(B) Razmatranje upotrebe i analiza postojećih ontologija – analizirana je RIC-O ontologija, a zbog definicija vezanih uz ponašanje bila je razmatrana i Behaviour Change Intervention ontologija (BCIO). Međutim, donesena je odluka da se u ovoj iteraciji razvoja ontologije modeliraju samo entiteti te arhivističko-etičke ontologije. Kao budući rad može se promisliti njezino eventualno uklapanje u RIC-O ontologiju, u vidu etičkog modula koji bi bilo jedna od mogućih eksperimentalnih nadogradnji RIC-a.

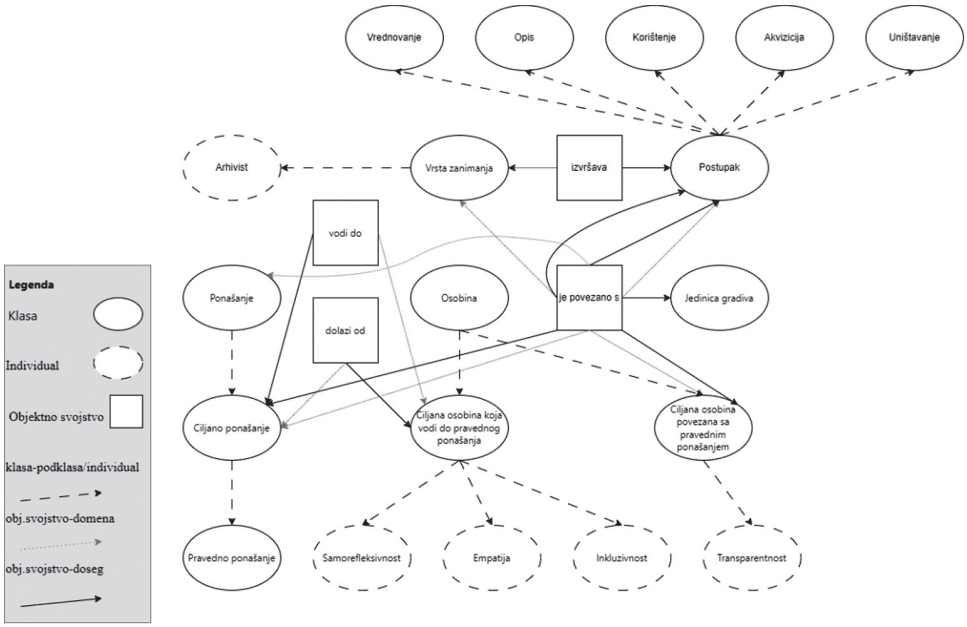
(C) Definiranje klasa, indivuduala, svojstava i aksioma te njihovo postavljanje u odnose.[28] Za početne klase ontologije u nastanku određene su klase Ponasanje s podklasom CiljanoPonasanje i podklasom PravednoPonasanje te Osobina s kompleksnom podklasom CiljanaOsobina. CiljanaOsobina je unija podklasa CiljanaOsobina-

26 John Rawls, „A theory of justice". Cambridge: Harvard Uni. Press, The Belknap Press, 1999., 120-121. Člankom Rajh, „Prilog arhivističkoj etici" u Arhivskom vjesniku opširno se objašnjavaju teorijska polazišta ove arhivističko-etičke ontologije. Ovaj rad nastavlja razvoj iz konceptualnog modela u ontologiju, a pritom služi i za objašnjenje moguće metode razvoja ontologija.

27 Arian Rajh, „Daetika", *GitHub* (2024), pristupljeno 18. 6. 2024. https://github.com/rarianSp/DAEtika.

28 U ovom koraku koristio se alat draw.io (https://app.diagrams.net/) zbog njegove jednostavnosti i mogućnosti grafičkog modeliranja entiteta uz prikaz svih klasa i individuala te deskriptivna logika o kojoj će biti riječi u tekstu.

KojaVodiDoPravednogPonasanja (s individualima Samorefleksivnost, Empatičnost, Inkluzivnost) i CiljanaOsobinaPovezanaSPravednimPonasanjem (s individualom Transparentnost). Dodana je i klasa Postupak s podklasama Vrednovanje, Opis, Korištenje, Akvizicija, Unistavanje. Za tu se klasu pri razvoju ontologije moralo naročito pripaziti da se modeliraju podklase jasno međusobno razdvojenih značenja. Ostale klase iz dijagrama prikazanog na Slici 1. uključuju JedinicuGradiva i Arhivista (prema RIC-O bio bi uvršten pod klasu OccupationType), [29] no te se klase, prema odluci iz prethodnog koraka, neće uvrstiti u ontologiju koja se modelira.



Slika 1. Konceptualizacija entiteta za ontologiju izrađena alatom draw.io

Tablica 3. Objektna svojstva, klase koje su njihove domene i klase koje su njihov doseg

Domena	Objektno svojstvo	Doseg
Ciljana osobina koja vodi do pravednog ponašanja	vodi do	Pravedno ponašanje
Ciljana osobina povezana s pravednim ponašanjem	je povezan s	Pravedno ponašanje
Ciljano ponašanje	dolazi od	Ciljana osobina
Ponašanje	je povezan s	Postupak
Postupak	je povezan s	Ponašanje

29 Kako je zaključeno u diskusiji Google grupe RIC https://groups.google.com/g/records_in_contexts_users/c/qc4PqaZVF24/m/lcIVuvJHAgAJ.

U ovom su koraku također analizirana i modelirana objektna svojstva, predikati, a svakomu od njih pridružena je domena i doseg – na ovaj se način s njima povezuju klase koje služe kao subjekti i objekti. Rezultati analize prikazani su u Tablici 3. Bilo bi poželjno koncepte prije definiranja u ontologiji modelirati primjenom deskriptivne logike. Deskriptivna logika (engl. *description logic*, DL) jest „obitelj […] formalizama koji reprezentiraju znanje aplikacijske domene ('svijeta'), prvo, definiranjem njezinih relevantnih koncepata (terminologije) te, nakon toga, upotrebom tih koncepata za određivanje svojstava objekata i individuala koji se u domeni pojavljuju (opis svijeta)".[30] Primjerice, definicijski aksiom odnosno definicija za koncept ili, kasnije, klasu ontologije PravednoPonasanje glasio bi:[31]

PravednoPonasanje

$\equiv$ *Ponasanje*

$\sqcap \forall$ *dolaziOd.CiljanaOsobinaKojaVodiDoPravednogPonasanja*

$\sqcap \forall$ *jePovezanoS.CiljanaOsobinaPovezanaSPravednimPonašanjem.*

Pravedno ponašanje ekvivalentno je ponašanju, i to ponašanju kod kojega se uvijek iskazuju osobine samorefleksivnosti, empatičnosti i inkluzivnosti i kod kojega su sve situacije povezane s transparentnošću kao ciljanom osobinom.

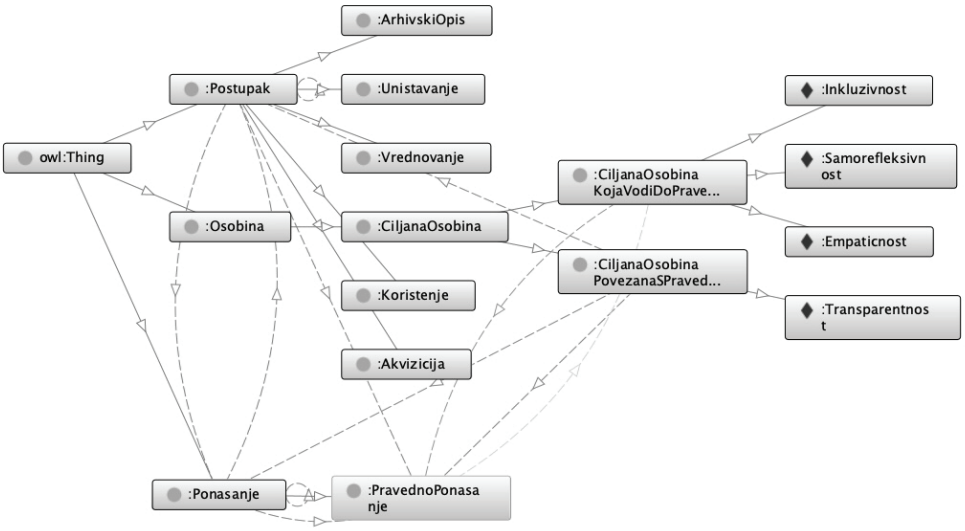
(D) Izrada ontologije – u modeliranju ontologije koja je sada nazvana „Daetika", prema „deontološko-aretaička etika", korišten je alat za uređivanje Protégé.[32] Za ontologiju su odabrane samo one klase koje ne postoje u RIC ontologiji, dakle u nju nije uvrštena ni podklasa Arhivist ni objektna svojstva koja postoje u RIC normi. Za svaku klasu ontologije Daetika stavljene su anotacije (rdfs:label, rdfs:comment). Uputno je napisati anotacije u skladu s definicijama koncepata deskriptivne logike, no to se u ontologijama uvijek može i naknadno doraditi. Dodani su individuali Samorefleskivnost, Empatičnost, Inkluzivnost i Transparentnost za klase CiljanaOsobinaKojaVodiDoPravednogPonasanja i CiljanaOsobinaPovezanaSPravednimPonasanjem. Individuali osobina morali su se razdvojiti u dvije podklase entiteta CiljnaOsobina jer je Transparentnost osobina koja, za razliku od Samorefleksnosti,

30 Franz Baader i dr., *The Description Logic handbook: theory, implementation, and applications*. (Cambridge: Cambridge University Press, 2003.), 43.

31 Péter Szeredi, Gergely Lukácsy i Tamás Benkő, *The Semantic Web explained: the technology and mathematics behind web 3.0.* (Cambridge: Cambridge University Press, 2014.), 206.

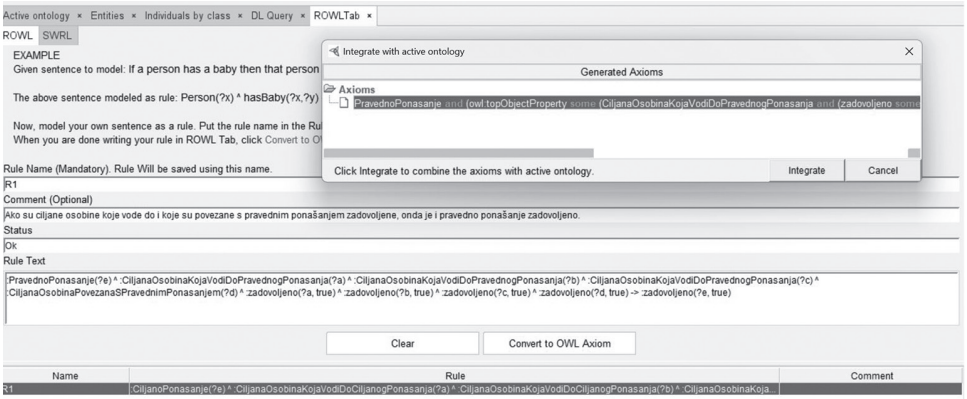
32 Alat Protégé može pohranjivati ontologije u raznim sintaksama pa je, zbog raširenosti upotrebe i autorova poznavanja sintakse, za pohranu datoteke ontologije odabrana sintaksa Turtle. Datoteka je kreirana pomoću OWL API (https://github.com/owlcs/owlapi), JAVA aplikacijskog programskog sučelja ugrađenoga u Protégé.

Empatičnosti i Inkluzivnosti, ne vodi do Pravednog ponašanja, ali je s njime nužno povezana. Klase Daetike koje ne mogu imati iste individuale bile su dizajnirane kao nepovezane klase. Dodana su objektna svojstva prikazana u Tablici 3. Prema konceptualnom modelu ontologije, Rawlova izvorna pozicija jest zamijenjena samorefleksivnošću arhivista, no ta samorefleksivnost može uslijediti i nakon iskazane empatičnosti i inkluzivnosti, stoga nije dizajnirano posebno objektno svojstvo koje bi ciljalo na samorefleskivnost (inicira, započinje ili sl.).



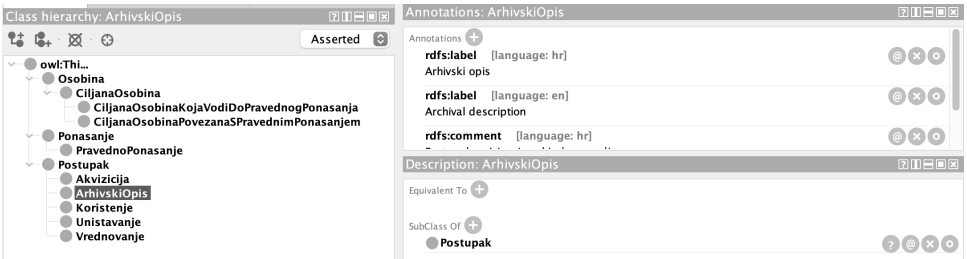
Slika 2. Graf ontologije Daetika izrađen alatom Protégé

U ontologiju je dodano podatkovno svojstvo „zadovoljeno" vrste Boolean (moguće vrijednosti *true/false*) za domenu PravednoPonašanje i CiljaneOsobine, a zatim se pomoću ROWL potprograma u alatu Protégé kreirao aksiom koji nalaže da vrijednost svojstva „zadovoljeno" za individual PravednoPonašanje može biti „točno" – ako su vrijednosti istoga svojstva točne za individuale klasa CiljanaOsobinaKojaVodiDoPravednogPonasanja i CiljanaOsobinaPovezanaSPravednimPonasanjem. Dakle, ponašanje može biti pravedno uz posjedovanje osobina samorefleksivnosti, empatičnosti, inkluzivnosti i transparentnosti. Pravilo je izrađeno u SWRL jeziku za izradu pravila (engl. *Semantic Web Rule Language*) i prikazano na Slici 3.

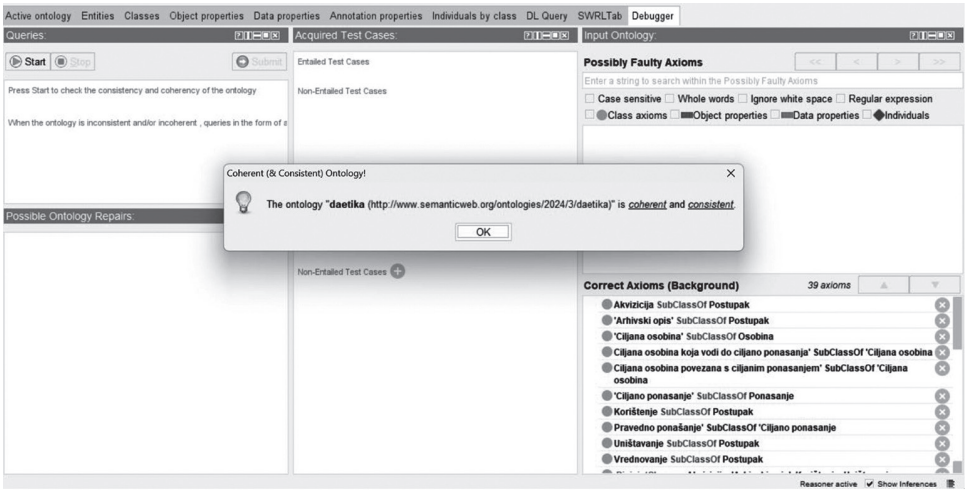


Slika 3. Ugradnja pravila na instancu PravednoPonasanje u alatu Protégé

(E) Redaktura i testiranje ontologije – redakturom su svi pojmovi usklađeni i dobiven je konačan izgled ontologije, a tijekom modeliranja i nakon njega za testiranje ontologije korišteni su dostupni zaključivači i softver za debugiranje u aplikaciji Protégé, kao što je prikazano na Slici 5. Ontologija, koju je moguće naći na GitHub stranici, kako je navedeno u bilješci 27, sadrži 12 klasa i četiri individuala, jedno podatkovno i četiri objektna svojstva. Dakle, ovo je ontologija manjeg obima koja ilustrira kako izraditi ontologije u ovoj domeni i koja se može, prema potrebi, koristiti kao modul većih ontologija.



Slika 4. Klase nakon čišćenja ontologije



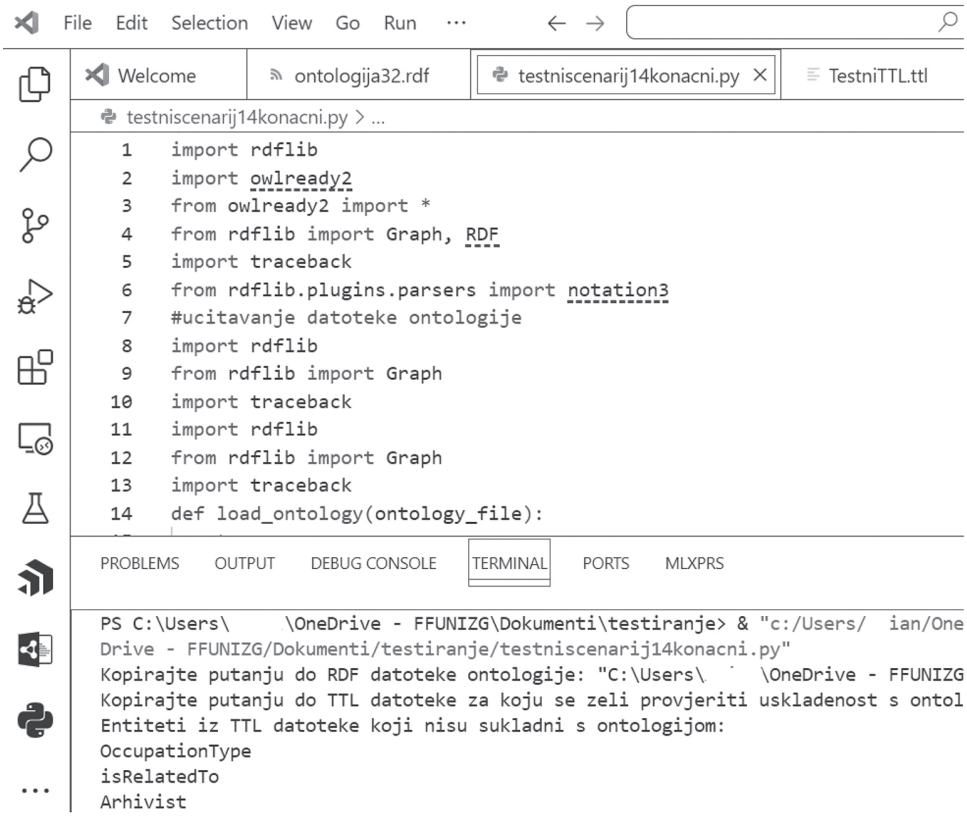
Slika 5. Primjena funkcije za debugiranje u alatu Protégé

Izrađena ontologija može se koristiti i za provjeru sadržaja datoteka s opisima entiteta pa je tako izrađeno nekoliko datoteka Turtle s mješovitim entitetima ontologije Daetika i drugih ontologija. Pomoću virtualne okoline Python za izrađene je datoteke opisa testirana njihova usklađenost s ontologijom Daetika. Testne datoteke su uz Daetika entitete sadržavale i RIC-O entitete. Skripta je prikazana na GitHub stranici,[33] a rezultati su prikazali one entitete iz testnih datoteka koji nisu pripadali ontologiji Daetika. Primjer rezultata može se vidjeti na Slici 6. Skripte poput ove mogu se koristiti za provjeru usklađenosti opisa arhivskih entiteta s ontologijama na koje bi se ti opisi trebali oslanjati, kao što je RIC-O.

Iako je Daetika izrađena na hrvatskom jeziku, oznake svih entiteta prevedene su i na engleski kako bi se u nekom trenutku, pomoću neke od odabranih tehnika interoperabilnosti,[34] ova etika mogla uklopiti u RIC-O kao njegov etički modul. Ovo trenutno nije napravljeno zbog opsega rada i uže svrhe ontologije Daetika koja se ostvaruje u objašnjavanju procesa modeliranja arhivističkih ontologija.

33 Ovo je rađeno u aplikaciji Visual Studio Code s ekstenzijama Pyton Extension Pack, knjižnicom rdflib Pyton i paketom za rad s ontologijama Owlready2. Pokazalo se da skripta bolje prolazi kroz veliki broj tripleta u RDF datotekama od datoteka Turtle pa je ontologija za ovaj test bila reserijalizirana. V. skriptu na adresi: https://github.com/rarianSp/DAEtika/blob/main/testniscenarij14konacni.py.

34 Tomislav Galba, „Algoritam transformacije ontologije u strukturu taksonomije za evidencijsko zaključivanje" (doktorska disertacija, Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija, 2018.), 77-80.



Slika 6. Izvršavanje skripte Python izrađene radi traženja entiteta koji nisu prema ontologiji dozvoljeni

4. Rasprava

Arhivistička ontologija osigurava prikaz i organizaciju znanja o svijetu arhivskoga gradiva i povezanih entiteta, interoperabilnost njihovih opisa koje kolaborativno rade razni arhivisti koristeći različite tehnologije te interpretabilnosti tih opisa drugih arhivista, korisnika i softverskih agenata. Već je spomenuto da ontologije mogu poslužiti uređivanju diskursa neke domene. Korištenje deskriptivne logike može na sličan način osigurati primjerenost koncepata njezina diskursa. Iako je poželjno koristiti deskriptivnu logiku prije strukturiranja ontologije, kao što je prikazano u koraku C metodologije izrade ontologija, deskriptivna se logika može koristiti i retroaktivno, za provjeru klasa postojeće ontologije:

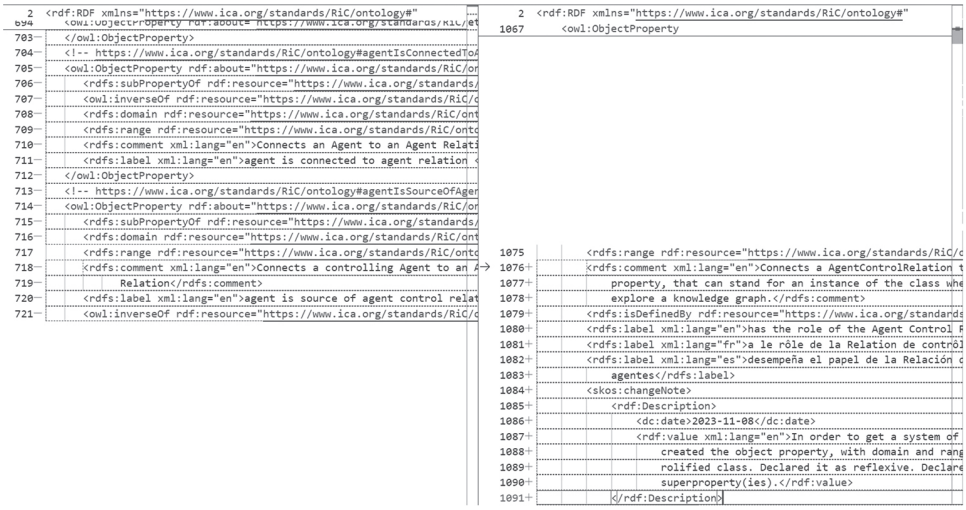
$$Family \equiv \geq 2 hasFamilyAssociationWithAgent.Person$$

U ovom primjeru definiran je koncept Obitelj, koji je klase Agent (stvaratelj) iz RIC norme i ontologije. Koncepti deskriptivne logike (C, Obitelj) korespondiraju klasama RIC-O ontologije (Obitelj ili Family, Osoba ili Person), a uloge (R, hasFamilyAssociationWithAgent) objektnim svojstvima. Obitelj je u RIC-O ontologiji definirana kao „dvije ili više osoba koje su u srodstvu rođenjem, bračnom vezom, posvojenjem, civilnim partnerstvom ili drugom društvenom konvencijom koja ih povezuje u društveno prepoznatu obiteljsku grupu".[35] Ovaj koncept je, dakle, definiran u okviru deskriptivne logike koristeći postojeću definiciju iz RIC ontologije. Poželjno je, dakle, da su ontološke definicije i definicije koncepata međusobno povezive i uskladive kako bi se lakše mogla provjeravati konzistentnost ontologije. Zbog toga bi bilo dobro prvo definirati koncepte pomoću deskriptivne logike i onda iz toga izvući definicije koje se objavljuju na webu u sklopu ontologija. Kako kod RIC ontologije to nije napravljeno, bilo bi korisno pojedine koncepte provjeriti uporabom alata deskriptivne logike te tako revidiranu ontologiju u nekom trenutku objaviti kao novu inačicu ontologije. Ontologije ionako treba redovito održavati pa se ovo može smatrati dijelom njihova uobičajenog životnog ciklusa.

Kako bi se RIC ontologija približila domaćoj stručnoj javnosti, bilo bi je isto tako korisno prevesti na hrvatski jezik, kao što su bila prevedene i ISAD(G) i ISAAR(CPF) norme. U datoteci RIC ontologije postoje engleski, francuski i španjolski termini. Hrvatski termini bili bi dobrodošli kako bi naši arhivisti razumjeli i prihvatili RIC-O normu. U siječnju 2022. Hrvatski državni arhiv (HDA) objavio je Specifikaciju metapodataka dokumentarnog i arhivskog gradiva. Specifikacija je napravljena prema tada važećoj 0.2. inačici nacrta norme RIC. U odnosu na RIC ontologiju Specifikacija detaljnije razrađuje neke od entiteta u složene semantičke sastavnice, no takve nacionalne implementacije mogu se pretpostavljati kod pripreme međunarodnih normi. Razlike između konačne inačice RIC norme i navedenog nacrta na koji se oslanjala Specifikacija metapodataka nisu velike, stoga se Specifikacija bez većih problema može koristiti u prijevodu ontologije RIC norme.[36] Na Slici 7 mogu se vidjeti neke od razlika ovih inačica RIC-O ontologije; konačna inačica ima jednu klasu, podatkovno svojstvo i dvadesetak objektnih svojstava manje.

35 Prijevod određenja klase Obitelj, v. Family, https://github.com/ICA-EGAD/RiC-O/blob/master/ontology/current-version/CSV_lists_of_components/RiC-O_1-0_list-of-classes.csv

36 Više informacija na https://groups.google.com/g/Records_in_Contexts_users/c/scWHTlhHIvg.



Slika 7. Izvadak iz detaljne usporedbe inačica ontologije 0.2 i 1.0 u programu Visual Studio Code 1.87.2

Pri izradi prijevoda RIC ontologije na hrvatski jezik trebalo bi koristiti CSV liste objavljene na EGAD-ovoj GitHub stranici.[37] Za svaku klasu i podatkovno i objektno svojstvo trebalo bi dodati rdfs:label naslov na hrvatskom jeziku i opis. Prijevod bi u CSV obliku onda bilo preporučljivo poslati službenim kanalom EGAD grupi. Pretpostavka je da bi EGAD grupa htjela dobiti što više nacionalnih prijevoda RIC entiteta te ih uvrstiti u datoteku ontologije. Konačni rezultat bio bi da tim odgovoran za ažuriranje datoteke i objave ažurirane datoteke na GitHub stranici EGAD-a uvrsti hrvatske pojmove u RIC-O datoteku. Ažurirana datoteka izgledala bi kao u primjeru prikazanom na Slici 8: za svaki prevedeni entitet pojavila bi se linija s rdfs:label svojstvom i natpis u skos:changeNote čvoru. U prijevodu bi se izgubili detalji razrade nekih od entiteta na složene semantičke sastavnice, osim ako se ne modelira svojevrsno proširenje kao nacionalni dodatak na normu. Prijevod norme ili prošireni prijevod bilo bi korisno imati jer je Specifikacija namijenjena javnim ustanovama Republike Hrvatske pod nadzorom ustanova Hrvatskog državnog arhiva, a RIC normu mogu koristiti i svi drugi imatelji i stvaratelji dokumentarnoga i arhivskoga gradiva.

37 https://github.com/ICA-EGAD/RiC-O/tree/master/ontology/current-version/CSV_lists_of_components

```
<owl:Class rdf:about="https://www.ica.org/standards/RiC/ontology#RecordSet">
    <rdfs:subClassOf rdf:resource="https://www.ica.org/standards/RiC/ontology#RecordResource"/>
    <rdfs:comment xml:lang="en">One or more records that are grouped together by an Agent based
        on the records sharing one or more attributes or relations.</rdfs:comment>
    <rdfs:isDefinedBy rdf:resource="https://www.ica.org/standards/RiC/ontology"/>
    <rdfs:label xml:lang="en">Record Set</rdfs:label>
    <rdfs:label xml:lang="fr">Ensemble d'objets informationnels</rdfs:label>
    <rdfs:label xml:lang="es">Agrupación documental</rdfs:label>
    <rdfs:label xml:lang="hr">Skup dokumenata</rdfs:label>
    <skos:changeNote>
        <rdf:Description>
            <dc:date>2023-11-12</dc:date>
            <rdf:value xml:lang="en">Aligned with RiC-CM 1.0: updated the description
                (rdfs:comment), skos:scopeNote.</rdf:value>
        </rdf:Description>
    </skos:changeNote>
    <skos:changeNote>
        <rdf:Description>
            <dc:date>2023-11-02</dc:date>
            <rdf:value xml:lang="en">Added a new rdfs:label in Spanish.</rdf:value>
        </rdf:Description>
    </skos:changeNote>
    <skos:changeNote>
        <rdf:Description>
            <dc:date>2023-08-25</dc:date>
            <rdf:value xml:lang="en">Added a new rdfs:label in French.</rdf:value>
        </rdf:Description>
    </skos:changeNote>
    <skos:changeNote>
        <rdf:Description>
            <dc:date>2024-04-04</dc:date>
            <rdf:value xml:lang="en">Added a new rdfs:label in Croatian.</rdf:value>
```

Slika 8. Kako bi izgledala RIC-O datoteka s hrvatskim prijevodom

5. Zaključak i preporuke

Pomoću ontologija u arhivistici možemo raditi opis gradiva i taj opis može zaista biti kolaborativan zbog veće razine zajedničkog razumijevanja entiteta. Primjer uspješne suradnje na opisivanju gradiva na temelju korištenja novih normi i tehnologija vježba je na kolegiju Sređivanje i opis arhivskoga gradiva u ljetnom semestru ak. god. 2023./2024., opisana u dijelu 3.1. ovoga rada i prezentirana međunarodnoj stručnoj javnosti na ovogodišnjoj arhivističkoj konferenciji u Radencima u Sloveniji. Opisivanje baštine, od fondova i zbirki do opisa na nižim razinama, potpomognuto novim tehnologijama pružalo bi, zbog vizualizacije opisa i povezivanja entiteta, daleko bogatije korisničko iskustvo od prijašnjega. RIC i norme i tehnologije semantičkog weba mogu nam pomoći u dobivanju odgovora na složenija pitanja od pretraživanja relacijskih baza podataka o gradivu,[38] u konstruiranju novih in-

38 Horrocks, What are ontologies good for?, 181.

formacija i u dijeljenju tog novonastalog znanja među arhivistima i korisnicima. U primjene ove vrste mogu se iskoristiti koncepti i tehnologije umjetne inteligencije za programiranje softverskih agenata za pretraživanje, odnosno sloja za izradu SPARQL upita iz prirodnog jezika korisnika.[39] Ovo je dobrodošla mogućnost jer arhivisti koji rade opis ne mogu, a kako je ustvrdio Hans Booms,[40] niti bi trebali, predvidjeti sve korisničke upite i načine korištenja gradiva. Možemo li „semantičke" i povezane arhive, upite koji bi se postavljali tako da umjetna inteligencija prevodi korisnička pitanja i novostečene informacije smatrati protutežom i rješenjem za situacije u kojima bismo mi trebali biti, kako se Booms našalio, vidovnjaci? Ontologije se mogu koristiti u razne svrhe i nisu pomodna tema, a s vremenom će i stručna arhivistička zajednica uvidjeti koje sve koristi od njih može imati.

Na kraju rada preporučilo bi se prevođenje RIC-O ontologije na hrvatski jezik kako bi se domaća stručna javnost u najvećoj mjeri mogla s tom normom familijarizirati; zatim proširenje tako prevedene RIC-O ontologije sa semantičkim jedinicama iz Specifikacije metapodataka za opis dokumentarnog i arhivskoga gradiva (HDA) i nadogradnja nacionalnog digitalnog arhivskog sustava u sljedećem ciklusu kako bi sustav mogao izravno koristiti glavnu arhivističku stručnu ontologiju, odnosno nacionalni prošireni prijevod te ontologije. Ova preporuka vrijedi ako se inzistira na korištenju semantičkih jedinica Specifikacije HDA, odnosno na opisu koji ide razinu dublje od razine koja je normirana RIC-om, što može biti poslovno opravdano u nacionalnim implementacijama međunarodnih normi. Naposljetku, preporuča se provjera postojećih ontologija alatima deskriptivne logike i zaključivačkim alatima[41] te, kod uočenih nekonzistentnosti ontologija, korištenje deskriptivne logike za razradu koncepata – sve radi konsolidacije, optimizacije i održavanja ontologija u njihovim budućim inačicama.

---

39 Radovi o ovoj temi mogu se pratiti još od rada Abraham Bernstein, Esther Kaufmann i Christian Kaiser, „Querying the semantic web with ginseng: a guided input natural language search engine", u: 15th Workshop on information technologies and systems, Las Vegas, 11-12.12.2005 (2005). Povezanost umjetne inteligencije i semantičkog weba na općenitoj razini može se vidjeti i u radu Jim Hendler, Tim Berners-Lee, „From the Semantic Web to social machines: a research challenge for AI on the World Wide Web", *Artificial Intelligence* 174, 2 (2010), 158-159.

40 Hans Booms, „Uberlieferungsbildung: keeping archives as a social and political activity", u: Archivaria 33 (Winter 1991-1992), 26.

41 Prilikom pisanja glavnog dijela ovoga rada na snazi je bila RIC ontologija u inačici 1.0.1. Ova inačica pokazivala je greške prilikom rada s alatima za automatsko zaključivanje (primjerice u okolini Protégé, u kojoj je autor testirao ontologiju RIC). Za inačicu RIC-a 1. 0. 2., koja se pojavila u rujnu 2024., radna grupa koja se bavi tom ontologijom najavila je da su navedene greške otklonjene, v. diskusiju na tematskoj grupi Google https://groups.google.com/g/Records_in_Contexts_users/c/J_7dPXKgOeI/m/WhCyEQ6KAgAJ, 6. 9. 2024.

IZVORI I LITERATURA

Arhivski izvori:

Filozofski fakultet Sveučilišta u Zagrebu (Hrvatska)
34. Ostavština Jaroslava Šidaka (1933-1980-ih)

Mrežni izvori:

GitHub stranica ArCO ontologije. Pristupljeno 2. 4. 2024. https://github.com/ICCD-MiBACT/ArCo.

GitHub stranica ArDO ontologije. Pristupljeno 2. 4. 2024. https://github.com/ISE-FIZKarlsruhe/ArDO.

GitHub stranica ARKIVO ontologije. Pristupljeno 2. 4. 2024. https://github.com/ArkivoTeam/ARKIVO.

GitHub stranica RIC-O ontologije. Pristupljeno 2. 4. 2024. https://github.com/ICA-EGAD/RiC-O/tree/master/ontology/current-version.

Rajh, Arian. „Daetika", *GitHub* (2024). Pristupljeno 18. 6. 2024. https://github.com/rarianSp/DAEtika

Github stranica sa Turtle datotekama iz primjera Šidak. Pristupljeno 18. 6. 2024. https://github.com/rarianSp/soag/tree/main/Jarosav%20Sidak

Literatura:

Archive Ontology (ArCo network). Pristupljeno 2. 4. 2024. http://wit.istc.cnr.it/arco/lode/extract?lang=en&url=https://raw.githubusercontent.com/ICCD-MiBACT/ArCo/master/ArCo-release/ontologie/archive/archive.owl.

Avello, Daniel Gayo; Gutiérrez, Dario Álvarez. „The cooperative web: a complement to the Semantic Web". U: *Proceedings 26th annual international computer software and applications*, ur. Danielle C. Martin, 179-183. Oxford, UK, 2002., 179-183, doi: 10.1109/CMPSAC.2002.1044550.

Baader, Franz; Calvanese, Diego; McGuinness, Deborah L.; Nardi, Daniele; Patel-Schneider, Peter F., ur. *The Description Logic handbook: theory, implementation, and applications.* Cambridge: Cambridge University Press, 2003.

Banu, Ayesha. „Ontologies for knowledge representation". U: *Semantic web technologies: research and applications*, ur. Archana Patel, Narayan C. Debnath, Bharat Bhushan. Boca Raton: CRC Press, 2022., 223-243. https://doi-org.ezproxy.nsk.hr/10.1201/9781003309420.

Berners-Lee, Tim; J. Hendler, James; Lassila, Ora. „The semantic web: a new form of web content that is meaningful for computers will unleash a revolution of new possibilities." *Scientific American* 5 (2001). Pristupljeno 29. 3. 2024. https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf.

Berners-Lee, Tim. „Linked data." (2006). Pristupljeno 18. 3. 2024. https://www.w3.org/DesignIssues/LinkedData.html.

Bernstein, Abraham; Kaufmann, Esther; Kaiser, Christian. „Querying the Semantic Web with ginseng: a guided input natural language search engine". U: *15th Workshop on information technologies and systems, Las Vegas*, 11-12.12.2005 (2005). Pristupljeno 22. 3. 2024. https://www.merlin.uzh.ch/contributionDocument/download/2283.

Booms, Hans. „Uberlieferungsbildung: keeping archives as a social and political activity". *Archivaria* 33 (Winter 1991-1992), 25-33. Pristupljeno 2. 1. 2024. https://archivaria.ca/index.php/archivaria/article/view/11796/12747.

Cline, Scott. *Archival virtue: relationship, obligation, and the just archives*. Chicago: SAA, 2021.

De Coulon, Baptiste. „Déploiement de la norme records in contexts pour la gestion des collections de la fondation SAPA", *Revue électronique Suisse de science de l'information* (RESSI), 24 (2024). Pristupljeno 1. 3. 2024. https://doi.org/10.55790/journals/ressi.2024.e1511

Furner, Jonathan. „'Records in Context' in context: a brief history of data modelling for archival description". U: *Engaging with records and archives: histories and theories*, ur. Fiorella Foscarini, Heather MacNeil, Mak Bonnie, Oliver Gillian, 41–62. London: Facet, 2016.

Galba, Tomislav. „Algoritam transformacije ontologije u strukturu taksonomije za evidencijsko zaključivanje". Doktorska disertacija, Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija, 2018. Pristupljeno 8. 4. 2024. https://urn.nsk.hr/urn:nbn:hr:200:731662.

Gracy, Karen. F. „Archival description and linked data: a preliminary study of opportunities and implementation challenges", *Archival Science* 15 (2015), 239–294 Pristupljeno 1. 3. 2024. https://doi.org/10.1007/s10502-014-9216-2

Grubner, Thomas R. „A translation approach to portable ontology specifications", *Knowledge acqusition* 5 (1993), 199-220.

Grubner, Thomas R. „Toward principles for the design of ontologies used for knowledge sharing". *International Journal of Human-Computer Studies,* vol. 43, br. 5–6 (1995), 907-928, https://doi.org/10.1006/ijhc.1995.1081. Pristupljeno 1. 4. 2024., https://www.sciencedirect.com/science/article/pii/S1071581985710816.

Grüninger, Mihael; Chow, Amanda; Wong, Janette. „Semiautomatic design of ontologies". U: *The practice of enterprise modeling. PoEM 2023. Lecture Notes in Business Information Processing, vol 497,* ur. Joao Paulo A. Almeida, Monika Kaczmarek-Heß, Agnes Koschmider, Henderik A. Proper. Cham.: Springer, 2024. Pristupljeno 1. 3. 2024. https://doi.org/10.1007/978-3-031-48583-1_9

Guarino, Nicola; Welty, Christopher A. „An overview of Ontoclean". U: *Handbook on ontologies. international handbooks on information systems,* ur. Steffen Staab, Rudi Studer. Berlin, Heidelberg: Springer, 2009., 201-220. https://doi.org/10.1007/978-3-540-92673-3_9.

Guarino, Nicola; Oberle, Daniel; Staab, Steffan. „What is an ontology?". U: *Handbook on ontologies. international handbooks on information systems,* ur. Steffen Staab, Rudi Studer. Berlin, Heidelberg: Springer, 2009., 1-17. Pristupljeno 3. 4. 2024. https://doi.org/10.1007/978-3-540-92673-3_9.

Heath, Tom; Bizer, Christian. „Principles of Linked Data". U: *Linked Data: evolving the Web into a global data space,* ur. Tom Heath, Christian Bizer. Cham.: Springer, 2011., 7-27. Pristupljeno 3. 4. 2024. https://doi.org/10.1007/978-3-031-79432-2_2.

Hendler, Jim; Berners-Lee, Tim. „From the Semantic Web to social machines: a research challenge for AI on the World Wide Web". *Artificial Intelligence* 174, 2 (2010), 156-161. https://doi.org/10.1016/j.artint.2009.11.010.

Horrocks, Ian. „What are ontologies good for?". U: *Evolution of semantic systems*, ur. Bernd Olaf Küppers, Udo Hahn, Stefan Artmann. Berlin, Heidelberg: Springer, 2013. 175-188. Pristupljeno 10. 4. 2024. https://doi.org/10.1007/978-3-642-34997-3_9.

Hrvatski državni arhiv. „Specifikacija metapodataka za opis dokumentarnog i arhivskog gradiva". Zagreb: Hrvatski državni arhiv, 2022. Pristupljeno 1. 4. 2024. http://www.arhiv.hr/Portals/0/Specifikacija.pdf.

Human Behaviour Change Project. Behaviour Change Intervention Ontology (BCIO). Pristupljeno 29. 11. 2023. https://www.ebi.ac.uk/ols4/ontologies/bcio.

Iqbal, Rizwan; Azirfah Azmi Murad, Masrah; Mustapha, Aida; Mohd Sharef, Nurfadhlina. „An analysis of ontology engineering methodologies: a literature review". *Research journal of applied sciences, engineering and technology* 6, 16 (2013), 2993-3000. Pristupljeno 3. 4. 2024. https://www.maxwellsci.com/msproof.php?-doi=rjaset.6.3684.

Kyriaki-Manessi, Daphne; Dendrinos, Markos. „Developing ontology for the university archives: the domain of technological education". *Procedia - Social and Behavioral Sciences* 147, (2014), 349-359. Pristupljeno 1. 4. 2024. https://doi.org/10.1016/j.sbspro.2014.07.111.

Lamy, Jean-Baptiste. „Owlready: ontology-oriented programming in Python with automatic classification and high level constructs for biomedical ontologies". *Artificial intelligence in medicine* 80 (2017). 11-28. Pristupljeno 10. 4. 2024. https://www.sciencedirect.com/science/article/abs/pii/S0933365717300271.

Lohmann, Steffen; Link, Vincent; Marbach, Eduard; Negru, Stafan. „WebVOWL: Web-based visualization of ontologies". U: *Knowledge engineering and knowledge management*. EKAW 2014. Lecture notes in computer science, vol 8982, ur. Patrick Lambrix, Eero Hyvonen, Eva Blomquist, Valentina Presutti, Guilin Qi, Uli Sattler, Ying Ding, Chiara Ghidini. Cham: Springer, 2015. Pristupljeno 1. 3. 2024. https://doi.org/10.1007/978-3-319-17966-7_21

Maddux, Cleborne C. „The Semantic Web and educational technology". *Educational technology* 48, 1 (2008), 3-7. Pristupljeno 18. 6. 2024. https://www.jstor.org/stable/44429538.

Niu, Jinfang. „Linked Data for Archives". *Archivaria* 82 (2016), 83-110. Pristupljeno 1.3.2024. https://archivaria.ca/index.php/archivaria/article/view/13582

Noy, Natalya F.; Mcguinness, Deborah L. „Ontology development 101: a guide to creating your first ontology". *Knowledge systems laboratory* 32. (2001). Pristupljeno 1. 3. 2024. https://protege.stanford.edu/publications/ontology_development/ontology101.pdf

Oyelade, Olaide N. „Ontology Modelling", U: *Semantic Web technologies: research and applications*, ur. Archana Patel, Narayan C. Debnath, Bharat Bhushan. Boca Raton: CRC Press, 2022., 223-243. https://doi-org.ezproxy.nsk.hr/10.1201/9781003309420.

Pandolfo, Laura; Pulina, Luca; Zielinski, Marek. "Towards an ontology for describing archival resources". U: *Proceedings of the second workshop on humanities in the Semantic Web (WHiSe II) co-located with the 16th International Semantic Web conference, WHiSe@ISWC*, ur. Alessandro Adamou, Enrico Daga, Leif Isaksen, 111-116.

Beč, Austrija, 22. 10. 2017., (2017). Pristupljeno 1. 3. 2024. https://ceur-ws.org/Vol-2014/paper-12.pdf

Rajh, Arian. „Archival description turns truly collaborative: an excercise in Records in Contexts standard", Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja, Radenci 16. 5. 2024., prezentacija na konferenciji. Pristupljeno 18. 6. 2024., http://www.pokarh-mb.si/storage/app/media/Radenci%202024/Predstavitve_2024/Rajh_Adian_Radenci_2024.pdf

Rajh, Arian. „Prilog arhivističkoj etici: osobine arhivista, perspektiva pravednosti i vrednovanje." *Arhivski vjesnik* 67 (2024)

Rawls, John. „A theory of justice". Cambridge: Harvard Uni. Press, The Belknap Press, 1999.

Russey Roke, Elisabeth; Kitchin Tillman, Ruth. „Pragmatic principles for archival linked data", *The American Archivist* 85, 1 (2022), 173–201. Pristupljeno 1. 3. 2024. https://doi.org/10.17723/2327-9702-85.1.173

Sirin, Evren; Parsia, Bijan; Cuenca Grau, Bernardo; Kalyanpur, Aditya; Katz, Yarden. „Pellet: a practical OWL-DL reasoner". *Journal of Web Semantics* 5, 2 (2007), 51-53, https://doi.org/10.1016/j.websem.2007.03.004.

Studer, Rudi; Benjamins, Richard V.; Fensel, Dieter. „Knowledge engineering: principles and methods". *Data & knowledge engineering*, 25, 1-2, (1998), 161-197. https://doi.org/10.1016/S0169-023X(97)00056-6.

Szeredi, Péter; Lukácsy, Gergely; Benkő, Tamás. *The Semantic Web explained: the technology and mathematics behind Web 3.0.* Cambridge: Cambridge University Press, 2014.

Vsesviatska, Oleksandra; Tietz, Tabea; Hoppe, Fabian; Sprau, Mirjam; Meyer, Nils; Dessì, Danilo; Sack, Harald. „ArDO: an ontology to describe the dynamics of multimedia archival records". U: *Proceedings of the 36th annual ACM symposium on applied computing (SAC '21),* ur. Chih-Cheng Hung, Jiman Hong, *1855-1863.* New York: Association for computing machinery, 2021. 1855–1863. https://doi.org/10.1145/3412841.3442057

W3C. OWL 2 Web ontology language primer (II. izd), W3C (2012). Pristupljeno 1. 3. 2024. http://www.w3.org/TR/owl2-primer/

W3C. OWL 2 Web ontology language Manchester syntax (2. izd.). Pristupljeno 1. 4. 2024. https://www.w3.org/TR/owl2-manchester-syntax/.

W3C. Resource Description Framework (RDF). (2014a). Pristupljeno 18. 3. 2024. https://www.w3.org/RDF/.

W3C, RDF 1.1 Turtle W3C recommendation (2014b). Pristupljeno, 1. 3. 2024. https://www.w3.org/TR/turtle/

W3C. SPARQL 1.1 query language. (2011). Pristupljeno, 1. 3. 2024. http://www.w3.org/TR/sparql11-query/.

W3C. SWRL A Semantic Web rule language combining OWL and ruleML (2004). Pristupljeno 8. 4. 2024. https://www.w3.org/submissions/SWRL/.

Willer, Mirna; Vukadin, Ana; Bosančić, Boris; Rajh, Arian; Zlodi, Goran. „Ontologija Pravilnika za opis i pristup građi u knjižnicama, arhivima i muzejima". *Vjesnik bibliotekara Hrvatske* 66 (2023), 2, 25-52.

Archival ontologies: a review, design, and implementation

Summary

The author familiarizes readers with Semantic Web standards and technologies such as RDF, Turtle, SPARQL, and ontologies. Ontologies and their elements are defined according to literature, and the author discusses ontologies concerning the archival domain, like ontology made by Kyriaki-Manessi and Dendrinos, ARKIVO, ArDO, ArCO, and KAM ontology. The author analyzes the Records in Contexts archival standard and its ontology RIC-O. After this analysis, the author describes various ontology modeling methodologies and uses a particular method to construct his example.

The constructed ontology called 'Daetika' is based on the conceptual model of archival ethics with elements of just behavior and characteristics of self-reflexivity, empathy, inclusivity, and transparency. The conceptual model was created similarly to the theory of justice by John Rawls and the theory of virtues by Scott Cline. The author defines the 'world' of exemplar ontology by using visual conceptualization and Description Logic. He designs Daetika primarily in the Protégé system. After modeling classes, data, object properties, and axioms in the ontology, the author tests the ontology's coherence and consistency with reasoners. He also creates several RDF Turtle files with descriptions – comprising Daetika classes and properties and classes, individuals, and properties of other ontologies. The author tests their compliance with Daetika ontology with a specially designed script.

The paper also concludes on possible archival implementations of ontologies - for more powerful descriptions, better collaboration while describing entities, construction of new knowledge about entities, better dealing with user inquiries that are impossible to anticipate, and for cleansing archival discourse. Can AI-powered

'semantic' and linked archives solve situations where archivists need to have 'clairvoyant' abilities (imaginatively put by Hans Booms)? At the very end, the author recommends translating RIC-O into Croatian to familiarize the domestic archival community with RIC. This may possibly involve expanding the translated ontology to encompass all semantic units from national metadata specification. The last recommendation is the most general and considers enhancing ontology consistency by utilizing Description Logic.

Keywords: archival description, Description Logic, Linked Data, ontology modeling, Records in Contexts, Semantic Web