Prema strukturiranim podacima na internetu

Milošević, Marija

Undergraduate thesis / Završni rad

2021

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:142:874782

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-04-19**

FILOZOFSKI FAKULTET
SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU

Repository / Repozitorij:

FFOS-repository - Repository of the Faculty of Humanities and Social Sciences Osijek

zir.nsk.hr

SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU

dabar
DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

Sveučilište Josipa Jurja Strossmayera u Osijeku Filozofski
fakultet Osijek
Preddiplomski studij informatologije

Marija Milošević

Prema strukturiranim podacima na internetu

Završni rad

Mentor: izv.prof.dr.sc. Boris Bosančić

Osijek, 2021.

Sveučilište Josipa Jurja Strossmayera u Osijeku Filozofski
fakultet Osijek
Odsjek za informacijske znanosti
Preddiplomski studij informatologije

Marija Milošević

Prema strukturiranim podacima na internetu

Završni rad

Društvene znanosti, informacijske i komunikacijske znanosti, informacijski sustavi i informatologija

Mentor: izv.prof.dr.sc. Boris Bosančić

Osijek, 2021.

IZJAVA

Izjavljujem s punom materijalnom i moralnom odgovornošću da sam ovaj rad samostalno napravila te da u njemu nema kopiranih ili prepisanih dijelova teksta tuđih radova, a da nisu označeni kao citati s napisanim izvorom odakle su preneseni.
Svojim vlastoručnim potpisom potvrđujem da sam suglasna da Filozofski fakultet Osijek trajno pohrani i javno objavi ovaj moj rad u internetskoj bazi završnih i diplomskih radova knjižnice Filozofskog fakulteta Osijek, knjižnice Sveučilišta Josipa Jurja Strossmayera u Osijeku i Nacionalne i sveučilišne knjižnice u Zagrebu.

U Osijeku, 07. rujna 2021.

Marija Miloševič, 0122228714
ime i prezime studenta, JMBAG

Sažetak

Svrha ovoga rada je prikazati Schema.org projekt koji je nastao kako bi se olakšalo pretraživanje i pronalaženje relevantnog sadržaja na mrežnim stranicama pomoću označavanja strukturiranim podacima. Schema.org projekt obuhvaća razvoj istoimenog rječnika sastavljenog od predmeta odnosno vrsta ili tipova sadržaja u hijerarhijskom odnosu te svojstava koji se rabe njihov opis. U prvom dijelu rada govori se o problemima koji su doveli do pokretanja projekta, zatim, o razvoju mikropodataka i načina njihova označavanja, te ulozi povezanih i strukturiranih podataka na semantičkom webu. Poseban naglasak u radu stavljen je upravo na postupak označavanja sadržaja putem mikropodataka odgovarajućim atributima iz Schema.org rječnika, te je u središnjem dijelu rada prikazan primjer takvog označavanja. Dalje se u radu opisuju karakteristike samog projekta i njegov razvoj kroz proširenja koja su omogućila širu upotrebu rječnika za potrebe raznih zajednica te se navodi i opisuje rad alata koji provjeravaju kreirane Schema.org strukturirane podatke. U posljednja dva poglavlja pozornost je usmjerena na područja primjene Schema.org rječnika te se pretpostavljaju buduće promjene s obzirom na dosadašnje aktivnosti ovoga projekta.

Ključne riječi: Schema.org, strukturirani podaci, semantički web, metapodaci, tražilice

Sadržaj

1. Uvod.....6

2. Motivacija za pokretanje Schema.org projekta 7

3. Schema.org.....9

3.1. Općenito o Schema.org 9

3.2. Korištenje mikropodataka unutar Schema.org 10

3.3. Schema.org proširenja.....12

3.4. Kreiranje i provjera Schema.org strukturiranih podataka 15

3.5. Primjena Schema.org 17

3.6. Budućnost Schema.org.....19

4. Zaključak 22

Literatura.....23

1. Uvod

Web se pretražuje svakodnevno pri čemu korisnici teže jednostavnijem pretraživanju koje bi rezultiralo preciznijim rezultatima s obzirom na traženi upit. Kako bi se omogućilo pronalaženje relevantnih i što preciznijih rezultata pomoću tražilica, pojedine organizacije započele su s izvođenjem različitih projekata čiji je cilj razumijevanje sadržaja pohranjenog na mrežnim stranicama. U ovome radu opisuje se jedan takav projekt koji je nastao udruživanjem nekoliko organizacija 2011. godine. Rezultat te suradnje je Schema.org projekt koji je detaljnije opisan u sljedećim poglavljima ovoga rada.[1] Rad je podijeljen na tri dijela, u prvom dijelu se opisuju koncepti na kojima se Schema.org temelji, poput HTML5, zatim standardi kao što su mikropodaci te, na kraju, i koncepti koji se odnose na povezane i strukturirane podatke. U prvom dijelu rada razmatraju se i problemi koji su doveli do pokretanja Schema.org projekta, razvoj samih mikropodataka i način njihova označavanja te uloga povezanih i strukturiranih podataka na semantičkom webu koji je uopće i doveo do nastanka spomenutog projekta. U drugom se dijelu rada opisuje što je to Schema.org te njezini sastavni dijelovi i kako funkcionira, odnosno način na koji se Schema.org rječnik rabi za označavanje sadržaja mrežnih stranica. Dalje je u radu detaljnije opisano označavanje sadržaja odnosno podataka mikropodacima koji se najčešće i koriste, jer osiguravaju podatke u ljudski čitljivom obliku pomoću unaprijed definiranih atributa. Nadalje, opisuju se Schema.org proširenja koja su nastala iz potrebe za što jednostavnijim modelom, odnosno okvirom označavanja vanjskog sadržaja i prikazivanja njegovih međusobnih odnosa. Poglavlje u kojem se govori o Schema.org proširenjima strukturirano je tako da se prvo opisuju najkorištenija proširenja, vanjska i unutarnja, a zatim su definirana i takozvana 'neriješena proširenja', kao i 'meta proširenja'. Uz to, navedena su posebna proširenja za određene organizacije i mrežne stranice u obliku novih rječnika koji se najčešće koriste za opisivanje sadržaja vezanih uz područja medicine i zdravstva, banaka i financijskih proizvoda i usluga, hotela i ostalih smještaja te automobila i ostalih vozila, a na kraju navedenog poglavlja opisuje se način na koji se mogu kreirati vlastita proširenja. Sljedeća se dva poglavlja odnose na načine na koje se može strojno kreirati i provjeriti kreirane Schema.org strukturirane podatke te ih primijeniti pomoću različitih alata kreiranih upravo za tu svrhu. Konačno, u zadnjem se poglavlju opisuje u kojem će se smjeru odvijati budući razvoj Schema.org projekta zajedno s proširenjima istoimenog rječnika i

[1] Usp. Schema.org. URL: https://schema.org/ (2021-04-12)
Usp.

primjenom te što organizacije koje su ga pokrenule i ostali istraživači očekuju od njegovog daljnjeg razvoja.

2. Motivacija za pokretanje Schema.org projekta

Razvojem informacijske tehnologije i web se razvijao te se i dalje razvija. Danas postoji semantički web koji se temelji na više slojeva koji daju podršku izgradnji strojno čitljivih aplikacija i koji predstavlja novu fazu u razvoju weba. Jedan od tih slojeva je i XML (engl. *eXstensible Markup Language*) koji predstavlja zamjenu za SGML (engl. *Standard Generalized Markup Language*). XML se javio kao pogodan format za strukturiranje i prijenos podataka između strojeva koji omogućuje dodavanje ograničenja na određene elemente i njihove vrijednosti.[2] Njegova strojna čitljivost očituje se u dodavanju početne deklaracije unutar XML dokumenta te definiranju strukture dokumenta pomoću DTD-ja (engl. *Document Type Definition*) ili u novije vrijeme sve korištenije XML Scheme koja će u budućnosti u potpunosti zamijeniti dosadašnji DTD.[3] Najpoznatija SGML aplikacija je HTML (engl. *HyperText Markup Language*) koja se razvila za potrebe kreiranja mrežnih stranica. HTML se usporedno s razvojem tehnologije sve više razvijao te je razvijao i svoje semantički utemeljene elemente koji pružaju informacije u ljudski i strojno čitljivom obliku.[4] Za razliku od XML-a, HTML ima točno definirane elemente koji su se tijekom vremena, a posebice u verziji HTML5, razvili kako bi što bolje odgovarali sadržaju i razumijevanju sadržaja. HTML5 pruža elemente koji, za razliku od elemenata u prethodnim verzijama HTML-a, stvaraju HTML dokument koji je lakši za razumijevanje čovjeku, ali i stroju. Međutim, navedeni HTML5 elementi ne pridonose potpunom razumijevanju sadržaja dokumenta, odnosno ne mogu odgovoriti na pitanje o čemu sadržaj na mrežnoj stranici govori. Iz tog je razloga u najnovijoj verziji HTML-a uključena sintaksa mikropodataka kako bi se omogućila povezanost strukturiranih podataka i njihovog prikaza na mrežnoj stranici.[4] Mikropodaci su se pojavili 2009. godine te upravo ističu značenje sadržaja na webu, a njihova jednostavna sintaksa koja rabi HTML atribute namijenjena je

---

[2] Usp. Antoniou, Grigoris; van Harmelen, Frank. A semantic Web primer. Massachusetts: Institute of Technology, 2004. Str. 25. URL: https://books.google.hr/books?hl=hr&lr=&id=cIYAiS1wbRsC&oi=fnd&pg=PR13&dq=semantic+web&ots=icwOBtRfBB&sig=k_2zVZ5nqtUkbfaz0eB45yQ3fw8&redir_esc=y#v=onepage&q&f=true (2021-04-02)

[3] Usp. Isto, str. 26. [4]

Usp. Isto, str. 23.

[4] Usp. Ronallo, Jason. HTML5 Microdata and Schema.org. // Code4Lib Journal 16 (2012). URL: https://journal.code4lib.org/articles/6400 (2021-04-02)

Usp.

autorima koji kreiraju mrežne stranice. Ono što se opisuje mikropodacima naziva se predmet, a svaki se predmet sastoji od para svojstva i vrijednosti. Model takvih podataka čine atributi *@itemscope* koji identificira novi predmet, *@itemtype* koji specificira vrstu predmeta te

@itemprop koji povezuje predmet s odgovarajućim svojstvom i vrijednosti.[5] Danas se sve više pažnje posvećuje mikropodacima, jer ih koristi sve više web tražilica u obliku pružanja kratkih isječaka pojedinih poveznica koji obuhvaćaju informacije o sadržaju mrežnih stranica.

Primjerice Google prikazuje kontekst u kojemu se nalaze ključne riječi pretraživanja na pojedinoj mrežnoj stranici. Na takav način korisniku se olakšava pregledavanje weba koje postaje znatno brže. Iako Google prikazuje isječke na takav način, brojne druge tražilice omogućuju autorima da sami odluče što će se prikazati u pojedinom isječku s obzirom na dodijeljene metapodatke, a sami metapodaci mogu se koristiti unutar koncepta povezanih podataka označeni ili 'enkodirani' u RDF-u (engl. *Resource Description Framework*). Važno je napomenuti kako metapodaci nisu povezani podaci, već podaci koji pružaju opis sadržaja koji se nalazi na mrežnoj stranici.[6] Međutim, mikropodaci mogu kreirati povezane podatke, odnosno mogu se koristiti za povezivanje izvora s odgovarajućim podacima koji ga opisuju. Za to se koristi *@itemid* atribut koji povezuje izvor s globalnim URL identifikatorom, a koristi se i u RDF formatu.[7] Takvo povezivanje, kao i svi povezani podaci, omogućuju strojevima stvaranje veza između sadržaja na webu olakšavajući korisnicima pretraživanje pomoću poveznica. Povezani podaci su dio strukturiranih podataka koji su povezani s ostalim podacima kako bi postali primjenjivi na semantičkom webu. Za kreiranje takvih strojno čitljivih podataka potrebne su različite web tehnologije od kojih su najznačajnije HTTP (engl. *HyperText Transfer Protocol*), RDF i URI (engl. *Uniform Resource Identifier*).[8]

Strukturirani podaci pokušavaju što bolje opisati mrežnu stranicu i njezin sadržaj kako bi ih razumjele web tražilice. U tu je svrhu najčešće potreban rječnik.[9] Jedan od najpoznatijih rječnika za opis sadržaja mrežnih stranica za potrebe učinkovitijeg pretraživanja je Schema.org rječnik koji se započeo razvijati u okviru istoimenog projekta. Rječnik mogu koristiti osobe

[5] Usp. Isto.
[6] Usp. Isto.
[7] Usp. Isto.
[8] Usp. What is linked data? URL: https://jsld.org/ (2021-04-02)
[9] Usp. Van de Rakt, Marieke. SEO basics: What is structured dana?, 2020. URL: https://yoast.com/what-isstructured-data/ (2021-04-06)

koje objavljuju sadržaj, ali i osobe kojima je sadržaj potreban, odnosno korisnici.[10] Djeluje na način da tražilice čitaju kôd te ga prikazuju u razumljivijem i preglednijem formatu za korisnika, najčešće u obliku isječaka s pridodanom vrijednošću (engl. *Rich Snippets*). Danas se sve više koriste različiti rječnici za opis podataka, pa tako postoje i rječnici povezanih podataka. Schema.org rječnik koristi isti model podataka, sintaksu i jezik kao i rječnik povezanih podataka, ali se razlikuje u smjernicama.[11] S obzirom da HTML format propisuje sam prikaz sadržaja, a ne i njegovo značenje, javila se potreba za drugačijim načinom opisa podataka bliža konceptu semantičkog weba. Također, iako se Schema.org pojavila kako bi omogućila učinkovitiji rad web tražilica, pokazala se učinkovitom i u rješavanju problema postojanja više srodnih rječnika te različitih korištenja tih rječnika. Upravo iz tog razloga često se zanemarivalo označavanje kako bi se izbjegle složenosti formata označavanja i kako bi se osigurao jedinstveni, univerzalni rječnik za širok raspon predmeta kao što su ljudi, mjesta, proizvodi i slično.[12]

3. Schema.org

3.1. Općenito o Schema.org

Schema.org projekt nastao je kao zajednički projekt četiri organizacije s ciljem kreiranja, održavanja i promoviranja rječnika koji su prilagođeni strukturiranim podacima na internetu. Prvotno i najvećim dijelom u kreiranju Schema.org sudjelovale su kompanije Google, Microsoft i Yahoo! kojima se vrlo brzo priključio Yandex, najpopularnija ruska web tražilica, ali su i brojne druge internetske zajednice na svoj način doprinijele razvoju te inicijative.[13] Schema.org predstavlja rječnik koji se koristi za označavanje različitih sadržaja koje preglednici prikazuju pomoću HTML-a. Rječnik, koji se još naziva i ontologijom, propisuje naziv vrste i svojstva izvora te njihove odnose i ograničenja.[14] Kao što je već navedeno, HTML ne utječe na način razumijevanja sadržaja stranice te preglednik ne zna koje je značenje prikazanog sadržaja. Upravo iz tog razloga pojavio se 2011. godine Schema.org projekt kojemu je cilj pružiti

[10] Usp. Guha, R.V.; Brickley, Dan; Macbeth, Steve. Schema.org: Evolution of Structured Dana on the Web. // ACM Queue 13, 9(2015), str. 1-28. URL: https://dl.acm.org/doi/pdf/10.1145/2857274.2857276 (2021-04-07)
[11] Usp. Isto, str. 22.
[12] Usp. Isto, str. 5.
[13] Usp. Schema.org. About. URL: https://schema.org/docs/about.html (2021-04-12)
[14] Usp. Barker, Phil; Campbell, Lorna M. What is schema.org?: A Cetis Briefing paper form LRMI, 2014. URL: http://publications.cetis.org.uk/2014/960 (2021-04-12) [16]
Usp. Schema.org. Nav.dj.
Usp.

korisniku što relevantnije rezultate pri pretraživanju sadržaja u odnosu na korisnički upit.[16] Relevantniji rezultati prikazuju se nakon što se sadržaj mrežnih stranica okvalificira odgovarajućim vrstama ili tipovima sadržaja (engl. *types*) i označi svojstvima (engl. *properties*) iz rječnika koje zatim prepoznaju tražilice Google, Microsoft, Yahoo! i Yandex. Uz rječnik, Schema.org uključuje i način označavanja sadržaja u obliku strojno čitljivih mikropodataka (engl. *Microdata*), RDF-a ili JSON-LD (engl. *JavaScript Object Notation for Linked Data*) formata putem kojih se svojstva rječnika i njihove vrijednosti ugrađuju u sadržaje mrežnih stranica.[15] Schema.org trenutno sadrži 779 vrsta sadržaja i 1390 svojstva koji su raspoređeni u dvije hijerarhije, jedna za tekstualne vrijednosti svojstava, a druga za stvari koje opisuju. Neke od najčešće korištenih vrsta sadržaja su Mrežna stranica (engl. *WebSite*), Osoba (engl. *Person*), Proizvod (engl. *Product*), Slikovna datoteka (engl. *ImageObject*), Članak (engl. *Article*), Blog (engl. *Blog*) i slično, dok su neka od najčešće korištenih svojstava naziv (engl.*name*), URL (engl. *Uniform Resource Locator*), opis (engl. *description*), slika (engl. *image*), cilj (engl. *target*), autor (engl. *author*) i druga.[16] Iako se Schema.org može koristiti i bez mikropodataka, kao što se i mikropodaci mogu koristiti neovisno o Schema.org, tražilice preferiraju mikropodatake, pa su vrste sadržaja i svojstva Schema.org najčešće opisani sintaksom mikropodataka.[17]

3.2. Korištenje mikropodataka unutar Schema.org

Mikropodaci omogućuju dodatne oznake koje su uključene u HTML dokument mrežne stranice kako bi ih tražilice i druge aplikacije razumjele, a time osigurale i razumijevanje sadržaja koji se nalazi na stranici.[18] Sintaksa mikropodataka nastala je kako bi se osigurala ravnoteža između pragmatičnosti i jednostavnosti, a preporučile su je osobe koje se bave testiranjem iskoristivosti mrežnih stranica.[21] Sintaksa mikropodataka već je spomenuta u poglavlju „Motivacija za pokretanje Schema.org projekta“, a u ovom će se poglavlju sada detaljnije opisati.

Označavanje mikropodacima temelji se na dodavanju atributa *@itemscope* i *@itemtype* unutar odjeljka HTML dokumenta koji se odnosi na informacije o sadržaju koji je označen. Navedeni atributi i njihove vrijednosti omogućuju tražilicama razumijevanje označenog

[15] Usp. Isto.
[16] Usp. Guha, R.V.; Brickley, Dan; Macbeth, Steve. Nav.dj., str. 12.
[17] Usp. Ronallo, Jason. Nav.dj.
[18] Usp. Schema.org. Getting Started. URL: https://schema.org/docs/gs.html (2021-04-12)
[21] Usp. Guha, R.V.; Brickley, Dan; Macbeth, Steve. Nav.dj., str. 13. [22] Usp. Schema.org. Getting Started. Nav.dj.

sadržaja unutar HTML dokumenta. Atribut *@itemtype* sadrži URL svojstava koja se koriste unutar svakog elementa odjeljka HTML dokumenta koji želimo dodatno pojasniti tražilici, a pojedina svojstva su zatim navedena u atributu *@itemprop*.[22] Postupak označavanja određenog sadržaja mikropodacima, primjerice knjige, prikazan je u nastavku.

Odjeljak HTML dokumenta koji pruža općenite informacije o knjizi, primjerice, o njezinu autoru, naslovu, godini izdanja i sl. identificira se u okviru elementa koji objedinjuje takve podatke, poput elemenata , <section> i sl. U primjeru se rabi element u kojemu se navodi atribut *@itemscope* koji specificira da se određeni odjeljak HTML dokumenta odnosi upravo na određeni sadržaj, u primjeru knjigu (*Book*) te se uz atribut *@itemscope* dodaje atribut *@itemtype* čija vrijednost predstavlja odgovarajući URL koji upućuje na tip, odnosno vrstu sadržaja u Schema.org hijerarhiji. Nakon što je određena vrsta sadržaja navode se pojedina Schema.org svojstva u okviru atributa *@itemprop* uz svaki element sadržaja (npr. autor) dodjeljujući im značenje. Označeni odjeljak HTML dokumenta bi izgledao ovako:[19]

```
<div itemscope itemtype= "http://schema.org/Book">

<h1 itemprop="name">Ubiti pticu rugalicu</h1>
<span>Autor: <span itemprop="author">Harper Lee</span></span>
<a href= https://www.youtube.com/watch?v=NQJbEMoimVI
itemprop="url">Audioknjiga</a>

</div>
```

U primjeru označen je sadržaj koji se odnosi na knjigu Ubiti pticu rugalicu pomoću mikropodataka koji koriste Schema.org svojstva kako bi tražilice razumjele sadržaj koji se nalazi na određenoj stranici. Tražilica nakon ovoga postupka označavanja - moglo bi se tako reći - "razumije" da se na HTML stranici nalazi audioknjiga Ubiti pticu rugalicu autorice Harper Lee. Sadržaj je označen pomoću Schema.org svojstava na sljedeći način: u okviru najopćenitije vrste sadržaja Stvari (engl. *Thing*) izabrana je specifičnija vrsta sadržaja Kreativno djelo (engl. *CreativeWork*) i unutar njega Knjiga (engl. *Book*), a zatim Audioknjiga (engl. *Audiobook*). Osim knjiga, pomoću Schema.org svojstava, mogu se opisati i brojni drugi sadržaji kao što su filmovi koristeći vrstu sadržaja Film (engl. *Movie*) ili mrežne stranice koristeći vrstu sadržaja Mrežna stranica (engl. *WebPage*). Sadržaj na mrežnim stranicama mogao se označiti i koristeći

[19] Označavanje je provedeno na temelju primjera prikazanog na stranici https://schema.org/docs/gs.html. Usp.

samo svojstva za Stvari, no preciznijim definiranjem odgovarajućih svojstva, hijerarhija pruža specifičnija, prikladnija svojstva za označavanje sadržaja. Također, može se dogoditi da za pojedini opis HTML sadržaja odgovara više od jednog svojstva, primjerice umjesto svojstva *author* može se koristiti i svojstvo *creator* itd.

Slično ugnježđivanju HTML elemenata, propisana svojstva i tipovi ili vrste sadržaja Schema.org hijerarhije mogu zajedno sa vrijednošću koja im je pridodana postati predmet s vlastitim svojstvima. Prilikom takvog označavanja atribut @itemscope se nalazi odmah nakon atributa @itemprop zajedno s odgovarajućim vrijednostima. Schema.org pruža opsežan popis predmeta (engl. *item*) odnosno vrsta sadržaja, a unutar predmeta se nalazi veliki broj svojstava (engl. *properties*) koji se koriste kako bi se opisali predmeti. Najobuhvatnija i najopćenitija vrsta sadržaja ili predmet Schema.org rječnika je Stvar (engl. *Thing*), koja se sastoji od četiri svojstva: ime ili naziv, opis, URL i slika.[20] Stvar tako sadrži u osnovi samo ta četiri svojstava, ali također sadrži i specifičnije predmete koji, uz četiri osnovna svojstva, sadrže više drugih specifičnih svojstava. To znači da ukoliko jedan predmet sadrži više drugih predmeta i svaki taj predmet sadrži više drugih predmeta, utoliko svi navedeni predmeti sadrže, uz vlastita svojstva, svojstva roditeljskog predmeta.[21] Tako je u gornjem primjeru unutar predmeta Stvari pronađeno Kreativno Djelo unutar kojega je pronađena Knjiga te Audioknjiga. Upravo iz tog razloga moguće je pronaći više svojstava za isti opis.

3.3. Schema.org proširenja

Uz propisane klase, odnosno vrste sadržaja ili predmete i svojstva, Schema.org pruža prilagodljivost u obliku proširivog mehanizma (engl. *Extension Mechanism*) koji korisnicima omogućuje kreiranje podvrsta i podsvojstava u okviru postojećih vrsta sadržaja i svojstava.[22] Proširenja su se tijekom godina mijenjala kako bi postala učinkovitija za korisnike te su prvo kreirana unutarnja proširenja (engl. *hosted extensions*) 2015. godine, a zatim i vanjska proširenja (engl. *external extensions*) 2019. godine. Unutarnja proširenja su dijelovi Schema.org koji su označeni unutar poddomene i objavljeni kao dio Schema.org projekta, dok su vanjska proširenja objavljenja bilo gdje na webu i o njima skrbe pojedine mrežne

[20] Usp. Schema.org. Getting Started. Nav.dj.
[21] Usp. Isto.
[22] Usp. Paulheim, Heiko. What the Adoption of schema.org Tells About Linked Open Data, 2015. URL: http://ceur-ws.org/Vol-1362/PROFILES2015_paper6.pdf (2021-04-14)

organizacije.[23] Primjer poddomene su bib.schema.org[24] ili auto.schema.org[25], dok je primjer vanjskih proširenja GS1 web rječnik[26] koji ujedno predstavlja i prvo objavljeno vanjsko proširenje Schema.org koje ima vlastitu strukturu i tijek rada, procese i infrastrukturu. GS1 se razvio iz B2B (engl. *Business-to-Business*) standarda, oblikovan na temeljnim preporukama W3 konzorcija (engl. *World Wide Web Consortium*) kao što je JSON-LD te je primjenjiv uz Schema.org rječnik.[27] Iako se GS1 može koristiti uz Schema.org rječnik, razvio je i vlastiti web rječnik koji se smatra jednim od vanjskih proširenja. Svako unutarnje proširenje, za razliku od vanjskih proširenja, ima vlastitu prilagođenu početnu stranicu koja opisuje jezik i status pojmova te poveznice na dodatne informacije.[28] Schema.org verzija 3.0 uvela je nova proširenja, primjerice, proširenja rječnika koja su u tijeku, odnosno rječnika koja još nisu usklađena (engl. *„pending schemas" extension*) sa Schema.org rječnikom. Neriješeno proširenje obuhvaća pojmove koji još nisu prihvaćeni u glavnom rječniku te se oni još uvijek mogu mijenjati u skladu s recenzijama, koje sadrže prijedloge za dodavanje novih pojmova, odnosno vrsta sadržaja i svojstava.[29] Schema.org organizacija predlaže izbjegavanje tih proširenja, no u nuždi mogu se koristiti uz povećani oprez. Uz navedena tri proširenja, Schema.org sadrži i manji broj pojmova koji se mogu koristiti za definiranje rječnika, ali oni nisu namijenjeni općenitoj uporabi na webu i oni se nazivaju meta proširenjima (engl. *„meta" extension*).[30]

Kao dodatne vrste proširenja pojavili su se rječnici koji omogućuju kreiranje strukturiranih podataka i označavanje koje je prilagođeno određenom sadržaju pojedinih područja kao što su zdravlje i medicina, automobili, hoteli te banke i financijske institucije. Takvi rječnici olakšavaju označavanje osobama koje objavljuju sadržaj na webu koji se odnosi na navedena područja. Dokumentacija se sastoji od tabličnog popisa entiteta, odnosno vrsta sadržaja i mogućih svojstava uz navođenje specifičnijih vrsta sadržaja. Glavna namjena rječnika je olakšavanje pronalaska traženog sadržaja na webu pružajući strukturirane informacije na

[23] Usp. Schema.org. Extensions. URL: https://schema.org/docs/extension.html (2021-04-14)
[24] Usp. Schema.org Section: bib, 2021. URL: https://schema.org/docs/bib.home.html (2021-04-14)
[25] Usp. Schema.org Section: auto, 2021. URL: https://schema.org/docs/auto.home.html (2021-04-14)
[26] Usp. GS1 Web Vocabulary, 2021. URL: https://www.gs1.org/voc/ (2021-04-14)
[27] Usp. Schema blog: GS1 Web vocabulary: welcoming the first schema.org external extension, 2016. URL: http://blog.schema.org/2016/02/gs1-milestone-first-schemaorg-external.html (2021-04-14)
[28] Usp. Schema.org. Extensions. Nav.dj.
[29] Usp. Isto.
[30] Usp. Schema.org. How We Work. URL: https://schema.org/docs/howwework.html (2021-04-14)
Usp.

mrežnim stranicama. Rječnik za definiranje zdravstvenih i medicinskih vrsta sadržaja nije namijenjen definiranju ili kodiranju novog medicinskog rječnika, već služi kao dopuna postojećih rječnika i ontologija kao što su MeSH (engl. *Medical Subject Headings*) ili SNOMED CT (engl. *Systemized Nomenclature of Medicine – Clinical Terms*).[31] Navedeni rječnik pruža popis vrsta sadržaja koje obuhvaćaju informacije o zdravstvenom stanju, lijekovima, smjernice za dijagnosticiranje i liječenje određenog zdravstvenog stanja, prikaz opisa, sažetka i punog teksta objavljenih medicinskih publikacija te ostalih informacija vezanih uz medicinu i zdravstvo kao i njihovih međusobnih odnosa, a označavanje se može provesti na dva načina: označavanjem vrste sadržaja neovisno o odnosima među različitim vrstama ili označavanjem odgovarajućim vrstama sadržaja zajedno s njihovim vezama. Rječnik koji pruža informacije i primjere za vrste sadržaja i svojstva prilikom označavanja hotela i ostalih smještaja obuhvaća tri skupine: hotele, hostele, odmarališta i ostale slične smještaje poput hotelskih soba i apartmana te ponuda za najam spomenutih smještaja.[32] Te tri skupine razlikuju se u označavanju s obzirom na korištenje različitih vrsta sadržaja i svojstava. Tako se za opis ponude koriste vrste sadržaja koje označavaju bilo koji ponudu bilo kojeg tipa, a za opis smještaja koriste se vrste sadržaja za smještaj (engl. *Accommodation*) ili za proizvod (engl. *Product*). Također, rječnik osigurava razna svojstva za takve vrste sadržaja kao što su informacije o iznajmljivanju, broju zvjezdica određenog smještaja i tko ih je dodijelio, radno vrijeme, recenzije drugih korisnika, cijene i slično.[37] Rječnik za označavanje automobila i ostalih vozila razvio se u okviru projekta GAO, odnosno Generičke automobilske ontologije, i sličnih ontologija. Služi za opis najpopularnijih vrsta automobila i njihovih konfiguracija, karakteristika motora te značajki koji su povezani uz komercijalne elemente. Taj rječnik pruža četiri vrste sadržaja te četrdeset četiri svojstava, kao i primjere označavanja mikropodacima u RDF i JSON-LD formatu.[33] Zadnji rječnik opisan u ovome radu, koristi se za opis banaka, financijskih ustanova i njihovih proizvoda kao što su krediti te ponude za klijente. Započeo je kao manja skupina pojmova koji su se razvili u samostalni rječnik. Hijerarhija se sastoji od različitih vrsta sadržaja i svojstava koje se isprepliću s vrstama sadržaja glavne Schema.org

[31] Usp. Schema.org. Documentation for health/medical types. URL: https://schema.org/docs/meddocs.html (2021-04-16)
[32] Usp. Schema.org. Markup for Hotels. URL: https://schema.org/docs/hotels.html (2021-04-16)
[37] Usp. Isto.
[33] Usp. Schema.org. Markup for Autos. URL: https://schema.org/docs/automotive.html (2021-04-16)

hijerarhije s njih ukupno osamnaest podvrsta i trideset jednim svojstvom.[34] Bitno je naglasiti da vrste sadržaja i svojstva u svakom pojedinom rječniku od prethodno navedenih su namijenjeni laicima, istraživačima i stručnjacima u određenom području te su kreirani primarno za web, a ne za označavanje poslovnih i inih podataka u automobilskoj industriji ili hotelijerstvu ili pak za razmjenjivanje podataka među organizacijama.

Poboljšanjem Schema.org, odnosno novijim verzijama, unapređuju se i proširenja te je u jednoj od verzija omogućeno da se opisi i definicije pojmova mogu premjestiti iz osnovne Schema.org jezgre u njezina unutarnja proširenja ili obrnuto. Svaki termin se označava kao dio jezgre ili proširenja te se definira kao širi pojam ili proširenje.[35] Uz navedeni način proširenja strukturiranih podataka postoje brojni drugi načini na koji je to moguće učiniti. Neki od načina su objavljivanje podataka u različitim formatima strukturiranih podataka kao što su mikroformati, RDF, JSON-LD formati ili mikropodaci, zatim opisivanje podataka pomoću para svojstvo-vrijednost, anotiranje mehanizmom za dodavanje dodatnih informacija te korištenje sintakse koja omogućuje korištenje više rječnika za opis.[36] Na opisani način osiguravaju se kvalitetniji strukturirani podaci koji mogu biti neovisni o Schema.org rječniku.

3.4. Kreiranje i provjera Schema.org strukturiranih podataka

Kako je već spomenuto, za kreiranje strukturiranih podataka rabe se mikropodatci, RDF ili JSON-LD formati. Tako kreirani strukturirani podaci trebaju biti što kvalitetniji. Upravo iz tog razloga razvili su se alati koji olakšavaju i pružaju podršku prilikom kreiranja Schema.org strukturiranih podataka. Navedeni alati korisni su prilikom izrade kratkih sažetaka ili upravljanja automatski generiranim označenim mrežnim stranicama.[37] Neki od alata su Googleov *Structured Data Markup Helper, Microdata Generator, Schema-creator by Raven, WordPress plugin i Drupam Schema.org modul*. Googleov pomagač za označavanje strukturiranih podataka pomaže u označavanju elemenata na mrežnoj stranici kako bi sadržaj, odnosno podaci, koji se nalazi na njoj postao razumljiv Google tražilici i vidljiviji prilikom pretraživanja. Tim alatom označava se događaj ili email kako bi se omogućilo prikazivanje kratkih isječaka, odnosno prikaz osnovnih informacija određenog emaila u Gmailu.

[34] Usp. Schema.org. Markup for Banks and Financial Institutions. URL: https://schema.org/docs/financial.html (2021-04-16)
[35] Schema.org. How We Work. Nav.dj..
[36] Usp. Isto.
[37] Usp. Barker, Phil; Campbell, Lorna M. Nav.dj.
Usp.

Označavanje se provodi tako da se odredi stranica koju je potrebno označiti, zatim odredi vrsta podataka koja se nalazi na mrežnoj stranici te se na kraju unese i odgovarajući URL ili HTML iste stranice. Zatim započinje označavanje tako što se odrede najvažniji i posebni dijelovi iz kojih se generira novi HTML koji se može preuzeti i testirati prema valjanosti. Na sličan se način označava email.[38] Generator mikropodataka koristi se, kako i samo ime kaže, za kreiranje mikropodataka, ali i JSON-LD oznaka. Za razliku od prethodno objašnjenog alata, generator se temelji na unaprijed definiranim poljima kao što su vrsta sadržaja i naziv posla, adresa, grad, telefon, koji se zatim ispunjavaju te se generira JSON-LD ili HTML kôd.[39] Ravenov kreator rječnika trenutno se nalazi na stranici seoscout.com te omogućuje provjeru već kreiranih strukturiranih podataka ili kreiranje novih na temelju unesenih podataka.[40] Wordpress dodatak kreira kôd mikropodataka koji se zatim dodaje YouTube videima kako bi se sažetci obogatili dodatnim opisima[41], dok Drupal modul kreira Google kratke sažetke za ljude, događaje i recepte pomoću softvera koji je automatski povezan sa Schema.org vrstama sadržaja i svojstvima.[42] Alati koji su prethodno opisani olakšavaju kreiranje strukturiranih podataka.

Sljedeći alati koji su navedeni provjeravaju valjanost kreiranih podataka. Jedan od takvih alata je Googleov Alat za testiranje strukturiranih podataka (engl. *Google Structured Data Testing Tool*) koji s obzirom na upisani URL ili isječak kôda testira i ukazuje na pogreške i upozorenja u kôdu.[48] Taj alat je zamijenila nova verzija koja se naziva Test pridodanih vrijednosti rezultatima pretraživanja (engl. *Rich Result Test*) koji provodi testiranje na sličan način, a osim toga, prikazuje koji se rezultati s pridodanim vrijednostima mogu kreirati iz postojećih strukturiranih podataka.[43] Yandex validator strukturiranih podataka (engl. *Yandex Structured data validator*) provjerava skoro sve formate za mikrooznačavanje kao što su mikropodaci, mikroformati, Schema.org, RDF i OpenGraph. Također se može unijeti URL ili

[38] Usp. Structured Data Markup Helper. URL: https://support.google.com/webmasters/answer/3069489?hl=en (2021-04-20)
[39] Usp. Microdata Generator: Local Business Schema Generator. URL: https://microdatagenerator.org/localbusiness-microdata-generator/ (2021-04-20)
[40] Usp. Schema Generator & Structured Data Testing Tool. URL: https://seoscout.com/tools/schema-generator (2021-04-20)
[41] Usp. Wordpress.org: plugins. URL: https://wordpress.org/plugins/tags/schemaorg/ (2021-04-20)
[42] Usp. Drupal: Schema.org. URL: https://www.drupal.org/project/schemaorg (2021-04-20) [48] Usp. Google: Alat za tesitranje strukturiranih podataka. URL: https://search.google.com/structureddata/testing-tool (2021-04-20)
[43] Usp. Google: Test obogaćenih rezultata. URL: https://search.google.com/test/rich-results (2021-04-20)

HTML kôd na temelju kojih se provjerava semantičko označavanje.[44] Kako bi se Yandexov validator mogao koristiti potrebna je prethodna registracija, a Yandex uz validator za strukturirane podatke pruža provjeru kreiranog XML-a i druge mogućnosti. Zadnji ovdje spomenuti alat za provjeru kreiranih strukturiranih podataka je Bingov validator označavanja (engl. *Bing's Markup Validator*) koji pruža slične značajke kao i prethodna dva spomenuta, ali se pomoću njega može provjeriti i samo URL određene mrežne stranice.[45] Postoje i drugi alati za kreiranje i provjeru kreiranih strukturiranih podataka koji ovdje nisu opisani, a također postoje i razni drugi alati koji ih na neki način organiziraju ili interpretiraju. *WebMoves microdata parser* samo je jedan od takvih sličnih alata koji parsiraju, odnosno interpretiraju, komponente podataka kako bi one postale razumljive nekom drugom jeziku, a u ovom slučaju parsiraju HTML5 mikropodatke putem URL-a određene mrežne stranice i prikazuju ih u hijerarhijama.[46] Daljnjim razvojem strukturiranih podataka razvijaju se i brojni drugi alati s različitim funkcijama, ali neki od njih nisu u mogućnosti pratiti razvoj strukturiranih podataka te više nisu u funkciji.

3.5. Primjena Schema.org

S obzirom da se Schema.org koristi za opis sadržaja mrežnih stranica, njezina primjena je vrlo široka. Prvu veću primjenu ostvarila je nedugo nakon nastanka, 2011. godine u vidu Googleovih takozvanih isječaka s pridodanom vrijednošću (engl. *Rich Snippets*), a daljnjim razvojem integrirana je u sve veći broj aplikacija.[47] Unutar aplikacija ima mnoge prednosti poput pružanja dodatnih informacija koje se koriste prilikom korisničkog postavljanja upita ili u obliku podsjetnika pojedinih aplikacija. Neke od aplikacija koje koriste Schema.org su virtualni asistent Microsoft Cortana, tražilica Yandex, društvena mreža Pinterest te Appleov iOS 9.[48] Email aplikacije, kao na primjer Google Gmail koriste Schema.org rječnik kako bi dohvatile strukturirane podatke i pružale obavijesti o pojedinoj transakciji putem pametnih telefona. Na sličan način djeluje i Cortana koja prima takve obavijesti te djeluje kao pomagač i podsjetnik pojedinom korisniku, dok je Yandexu i Pinterestu potreban takav rječnik kako bi korisnicima

[44] Usp. Yandex Webmaster: Structured data validator. URL: https://webmaster.yandex.com/tools/microtest/ (2021-04-20)
[45] Usp. Microsoft Bing: URL Inspection. URL: https://www.bing.com/webmasters/help/url-inspection-55a30305 (2021-04-20)
[46] WebMoves: Internet Strategists. URL: https://www.webmoves.net/tools/microdata (2021-04-21)
[47] Usp. Guha, R.V.; Brickley, Dan; Macbeth, Steve. Nav.dj., str. 6.
[48] Usp. Isto, str. 6.
Usp.

prikazao što relevantnije rezultate, čak i pomoću 'obogaćenih' kratkih isječaka. Slično svemu nabrojanom, Schema.org se pojavljuje i u Appleovom devetom izdanju mobilnog operativnog sustava za potrebe virtualnog asistenta Siri, za potrebe pretraživanja.[49] Veliku popularnost među mrežnim stranicama Schema.org može zahvaliti svojoj sposobnosti iskazivanja odnosa i vrsta pojedinih sadržaja koristeći atribute kao što su *@type* i *@context*. Zbog toga se primjenjuje u različitim područjima vijesti (BBC), filmova (IMDB), glazbe (SoundCloud) i videa (YouTube), poslovanja i trgovine (eBay, Alibaba), događaja (Meetup) i slično.[50] S obzirom da su Schema.org jezik kreirali Google, Yandex, Yahoo! i Microsoft Bing ne čude podaci da najveću prednost od primjene Schema.org rječnika i označavanja koriste upravo navedene tražilice. Brojne provedene ankete i slična istraživanja 2013. godine pokazala su kako se samo na mrežnim stranicama koje se prikazuju prilikom Google pretraživanja nalazi više od pet milijuna Schema.org strukturiranih podataka koji su pridodani kako bi pretraživanje bilo učinkovitije.[51] U proteklih osam godina zasigurno se taj broj povećao. Također je potrebno stalno usavršavanje i razvoj Schema.org rječnika kako bi ga sve veći broj tražilica koristio u pretraživanju, a tražilice će ga koristiti ako ga mrežne stranice implementiraju u svoj sadržaj.[52] To je ciklus koji je potrebno redovito provoditi i razvijati. Uz rječnik potrebno je neprestano razvijati Schema.org mrežnu stranicu na kojoj se nalazi popis svih vrsta sadržaja i svojstava u obliku hijerarhije te dokumentacija koja opisuje način označavanja sadržaja putem istih koja je namijenjena svim zainteresiranim osobama u ovom području. Već spomenute vrste sadržaja i svojstva definirani su u hijerarhiji unutar modela podataka koji je izveden iz RDF modela podataka, a takav rječnik može biti proširen ako njegova primjena to bude zahtijevala.[53] Schema.org rječnik ne predstavlja jedinstvenu i originalnu ontologiju, već prilagođenu okruženju kako bi zadovoljio potrebe organizacija koje su ga i kreirale. Tako rječnik propisuje uporabu već postojećih svojstava iz drugih rječnika umjesto kreiranja mnoštva sličnih svojstava za određeni sadržaj.[60] Zbog mogućnosti proširenja i prilagodljivosti dodavanja novih svojstava Schema.org ima vrlo široku primjenu u mnogim područjima. Tako je pronašla mjesto i u turizmu. Naime, dodavanjem vrste sadržaja Hotel, Hostel ili Motel na mrežne stranice pojedinih hotela, hostela ili motela, oni se javljaju češće u rezultatima upita odnosno povećava se njihov

[49] Usp. Isto, str. 6-8.
[50] Usp. Isto, str. 9-11.
[51] Usp. Barker, Phil; Campbell, Lorna M. Nav.dj.
[52] Isto.
[53] Usp. Schema.org. Data Model. URL: https://schema.org/docs/datamodel.html (2021-04-08)
[60] Usp. Isto.

odziv. Sličnu primjenu Schema.org pronašla je za potrebe geografskih informacija koje omogućavaju i lociranje pojedinih hotela i ostalih smještaja što je od iznimne važnosti za turizam.[54] Jedno od istraživanja provedeno na temu Schema.org označavanja i geografskih informacija provedeno je 2018. godine na Sveučilištu u Innsbrucku u Austriji. Postupak označavanja temeljio se na mapiranju postojećih svojstava geografske domene s odgovarajućim vrstama sadržaja i svojstvima iz Schema.org rječnika. U tom je istraživanju umjesto mikropodataka za označavanje korišten JSON. Nakon strukturiranja, zapis je sadržavao jedinstveni identifikator, naziv, opis, polje za unos geografskih koordinata, id kategorije i metapodatke različitih sadržaja.[62] Sljedeći koraci obuhvaćaju mapiranje kategorija i svojstva koji se prikazuju kao skup JSON-LD objekata koristeći pritom Schema.org rječnik. Na kraju, provedeno je čak i vrednovanje kako bi se otkrili nedostatci i prednosti označavanja strukturiranim podacima Schema.org rječnika.[63] Slična su istraživanja provedena i u svrhu povezivanja Schema.org rječnika i mode te gastronomije. Problemi kod takvih istraživanja najčešće su bili povezani s činjenicom da Schema.org rječnik postoji samo na engleskom jeziku iz čega proizlaze propusti kod prevođenja, čije otklanjanje predstavlja prioritet u budućem razvoju rječnika. Uspješna implementacija Schema.org rječnika potvrđuje se u velikom broju mrežnih stranica koje koriste Schema.org rječnik za označavanje sadržaja te u količini usluga koje promoviraju označeni sadržaj.[55] Schema.org rječnik može se primijeniti i u knjižnicama, arhivima i muzejima za potrebe lakšeg pretraživanja i pronalaženja tražene građe, odnosno gradiva. Iako se primarno primjenjuje na webu, posebice u okviru semantičkog weba, rječnik je moguće uključiti u razne organizacije kojima su potrebni strukturirani i povezani podaci kako bi što brže otkrili informacije koje su im potrebne.

3.6. Budućnost Schema.org

Kao što je već navedeno, Schema.org primjenjiva je na veliki broj aplikacija i mrežnih stranica na različite načine. Svoj dosadašnji uspješan razvoj može zahvaliti organizaciji podijeljenoj u dvije grupe: zajedničkoj grupi ili grupi zajednice (engl. *Community Group*) koja upravlja razvojem, raspravom i integracijom rječnika i upravljačkoj grupi (engl. *Steering Group*) koja

[54] Usp. Panasiuk, Oleksandra… [et al.]. Representing GeoData for Tourism with Schema.org, 2018. URL: https://www.scitepress.org/Papers/2018/67551/67551.pdf (2021-04-05) [62] Usp. Isto, str. 3-4. [63] Isto, str. 6-7.

[55] Usp. Barker, Phil; Campbell, Lorna M. Nav.dj. Usp.

prihvaća nove verzije.[56] Obje grupe shvaćaju Schema.org kao projekt koji zahtjeva neprestani razvoj, ali i kao projekt koji tijekom godina sve više drugih organizacija i tražilica počinje koristiti. Sama Schema.org organizacija smatra da će se razvoj usmjeriti na identificiranje novih vrsta strukturiranih podataka koje će osigurati kvalitetnije rezultate pretraživanja te ugradnju drugih rječnika u Schema.org rječnik uz poticanje drugih zajednica na kreiranje vlastitih odgovarajućih rječnika. Također, Schema.org organizacija smatra kako će rječnik uskoro biti dostupan i na drugim jezicima osim engleskog.[57] Iako je Schema.org, odnosno njezin rječnik trenutno dostupan samo na engleskom jeziku, označavanje se može koristiti na bilo kojoj mrežnoj stranici na bilo kojem jeziku zbog čega rješavanje te poteškoće i nije prioritet. S obzirom da se rječnik koristi u velikom broju aplikacija i mrežnih stranica, očekivano je da ne ispunjava specifične zahtjeve. Upravo zbog jedinstvenih zahtjeva neke organizacije izrazile su nezadovoljstvo dosadašnjim razvojem Schema.org projekta što je rezultiralo različitim novim inicijativama. Posljedice tih reakcija su uvođenje rNews standarda IPTC-a (engl. *International Press and Telecommunications Council*) za potrebe novinarstva i telekomunikacija. To je ujedno i prva organizacija koja je objavila proširenje uz partnerstvo Scheme.org što je postavilo temelj za suradnju Schema.org projekta s brojnim drugim organizacijama.[58] Ovakve suradnje očekuju se zasigurno i u budućnosti koja će dovesti do daljnjih Schema.org proširenja. Što se tiče dosadašnjih proširenja koja su uspješno implementirana kao Schema.org dodatna proširenja, smatra se da će se i ona u budućnosti razvijati zbog potreba različitih organizacija koje sve više koriste proširenja kako bi se dodali detaljniji pojmovi te odvojili pojedini termini s ciljem izbjegavanja preklapanja i preusko definiranih pojmova.[59] Također, dosadašnja proširenja prilikom označavanja koriste oznake koje upućuju na to da se koristi proširenje koje još nije prihvaćeno u Schema.org rječniku. To će se u budućnosti promijeniti tako da će oznake postati jednostavnije ili će ih se izostaviti, jer u današnjoj praksi dolazi do miješanja takvih oznaka. Financijska i bankovna proširenja će uključiti dodatne i izmijeniti postojeće vrste sadržaja i svojstava te smanjiti ograničenja, a proširenja koja su u tijeku će vremenom prestati biti neriješena ili će ih se izbaciti iz projekta ovisno o kontekstu, dok će vanjska proširenja prolaziti isti postupak.[60] Kako bi Schema.org postala još više korištena unaprijedit će se i dosadašnje brojne prednosti kao što su uključivanje URL-a srodne ili iste mrežne stranice,

[56] Usp. Schema.org. About. Nav.dj.
[57] Usp. Isto.
[58] Usp. Ronallo, Jason. Nav.dj.
[59] Usp. Schema.org. Extensions. Nav.dj.
[60] Usp. Schema blog. Nav.dj..

primjerice filma i *trailera* filma ili knjige i sažetka knjige, što ujedno pomaže tražilicama za bolje razumijevanje sadržaja koji se opisuje na mrežnoj stranici. Uz navedeno, prednost je što postoji više elemenata za dodavanje poveznice na mrežnu stranicu, kao što su <a> i <link> element, a ukoliko je sadržaj teško opisati zbog njegova prikaza, može ga se opisati pomoću elementa <meta> uz atribut @content. Također, Schema.org rječnik pruža hijerarhiju za vrste podataka koji se opisuju kao što su datum (engl. *Date*), Boolean tip podataka (engl. *Boolean*), Broj (engl. *Number*), Tekst (engl. *Text*) i slično.[61] U budućnosti, hijerarhija vrsta sadržaja i svojstava bi trebale brže usvajati svoje izmjene i preinake, jer postoje domene koje su obuhvaćene Schema.org-om, ali se u njezinim hijerarhijama ne mogu pronaći odgovarajući postavljeni podaci.[62] Uz to, hijerarhije trebaju biti vidljivije te se treba osigurati njihovo olakšano pretraživanje, jer tražilice nerijetko brže pronalaze danas zastarjeli rječnik mikropodataka koji se koristio prije pojave Schema.org rječnika..[63] Schema.org danas pruža otprilike desetak promjena na godinu, ali će u budućnosti uvelike promijeniti svoj rječnik u smislu kombinacije s već postojećim GS1 rječnikom. Na taj bi se način osigurao bogatiji opis proizvoda na webu upravo zbog uporabe dva bogata rječnika na najbolji mogući način.[64] Od prvog korištenja u vidu Googleovih kratkih isječaka s pridodanom vrijednošću pa sve do brojnih proširenja, Schema.org često je korištena zbog čega ju je potrebno neprestano unaprjeđivati. Schema.org će se nastaviti razvijati u budućnosti s obzirom na potrebe korisnika, jer se ona ne razvija kao model koji teži savršenstvu, već kao model koji odgovara na potrebe onih koji žele dati dodatnu vrijednost sadržaju.

[61] Usp. Isto.

[62] Usp. Paulheim, Heiko. Nav.dj., str. 5.

[63] Usp. Isto.

[64] Schema blog. Nav.dj.
Usp.

4. Zaključak

Razvojem informacijske tehnologije razvili su se različiti projekti, inovacije, tehnologije i koncepti koji omogućuju daljnji razvoj weba, a koji se i dalje razvijaju. Kao jedna od takvih tehnologija koristi se model mikropodataka koji je svoju primjenu pronašao na mrežnim stranicama. Mikropodaci opisuju sadržaj mrežne stranice odgovarajućim oznakama i atributima te su često povezani s odgovarajućim tehnologijama kako bi njihova primjena bila učinkovitija. Prednost mikropodataka uvidjele su brojne zajednice koje na temelju mikropodataka kreiraju rječnike. Jedan je od takvih rječnika Schema.org koji služi za opis strukturiranih podataka na webu kako bi tražilice prikazale relevantnije rezultate pretraživanja. Korištenje rječnika zahtjeva uporabu sintakse mikropodataka uz dodavanje vrsta ili tipova sadržaja te odgovarajućeg svojstva, a svi strukturirani podaci koji su označeni mikropodacima mogu se kreirati i provjeriti pomoću raznih alata koje su razvile organizacije za vlastite potrebe. Kako bi se izbjeglo korištenje jednog sveobuhvatnog rječnika za sve zajednice, kreirana su Schema.org proširenja. Trenutno su proširenja podijeljena na vanjska i unutarnja, a unutar dvije glavne skupine nalaze se proširenja za specifične zajednice kao što su medicina ili automobilska industrija. Iako ne postoje proširenja za svaku zajednicu, Schema.org rječnik koriste i druge zajednice zbog njegovih brojnih prednosti, a razvojem Schema.org rječnika nastat će brojna druga proširenja. Kao još jedna od prednosti ističe se to što je Schema.org rječnik kreiran na univerzalnoj razini, za svakoga. Iz svega navedenoga može se zaključiti kako je ovo područje vrlo razvijeno, ali i da teži neprestanom daljnjem razvoju.

Literatura

1. Antoniou, Grigoris; van Harmelen, Frank. A semantic Web primer. Massachusetts: Institute of Technology, 2004. Str. 25. URL: https://books.google.hr/books?hl=hr&lr=&id=cIYAiS1wbRsC&oi=fnd&pg=PR13&dq=semantic+web&ots=icwOBtRfBB&sig=k_2zVZ5nqtUkbfaz0eB45yQ3fw8&redir_esc=y#v=onepage&q&f=true (2021-04-02)
2. Barker, Phil; Campbell, Lorna M. What is schema.org?: A Cetis Briefing paper form LRMI, 2014. URL: http://publications.cetis.org.uk/2014/960 (2021-04-12) 3. Drupal: Schema.org. URL: https://www.drupal.org/project/schemaorg (2021-04-20)
4. Google: Alat za tesitranje strukturiranih podataka. URL: https://search.google.com/structured-data/testing-tool (2021-04-20)
5. Google: Test obogaćenih rezultata. URL: https://search.google.com/test/rich-results (2021-04-20)
6. GS1 Web Vocabulary, 2021. URL: https://www.gs1.org/voc/ (2021-04-14)
7. Guha, R.V.; Brickley, Dan; Macbeth, Steve. Schema.org: Evolution of Structured Dana on the Web. // ACM Queue 13, 9(2015), str. 1-28. URL: https://dl.acm.org/doi/pdf/10.1145/2857274.2857276 (2021-04-07)
8. Microdata Generator: Local Business Schema Generator. URL: https://microdatagenerator.org/localbusiness-microdata-generator/ (2021-04-20)
9. Microsoft Bing: URL Inspection. URL: https://www.bing.com/webmasters/help/urlinspection-55a30305 (2021-04-20)
10. Panasiuk, Oleksandra… [et al.]. Representing GeoData for Tourism with Schema.org, 2018. URL: https://www.scitepress.org/Papers/2018/67551/67551.pdf (2021-04-05)
11. Paulheim, Heiko. What the Adoption of schema.org Tells About Linked Open Data, 2015. URL: http://ceur-ws.org/Vol-1362/PROFILES2015_paper6.pdf (2021-04-14)
12. Ronallo, Jason. HTML5 Microdata and Schema.org. // Code4Lib Journal 16 (2012). URL: https://journal.code4lib.org/articles/6400 (2021-04-02)
13. Schema.org. URL: https://schema.org/ (2021-04-12)
14. Schema.org. About. URL: https://schema.org/docs/about.html (2021-04-12)
15. Schema.org. Data Model. URL: https://schema.org/docs/datamodel.html (2021-04-08)
16. Schema.org. Documentation for health/medical types. URL: https://schema.org/docs/meddocs.html (2021-04-16)

17. Schema.org. Extensions. URL: https://schema.org/docs/extension.html (2021-04-14)
18. Schema.org. Getting Started. URL: https://schema.org/docs/gs.html (2021-04-12) 19. Schema.org. How We Work. URL: https://schema.org/docs/howwework.html (202104-14).
20. Schema.org. Markup for Autos. URL: https://schema.org/docs/automotive.html (2021-04-16)
21. Schema.org. Markup for Banks and Financial Institutions. URL: https://schema.org/docs/financial.html (2021-04-16)
22. Schema.org. Markup for Hotels. URL: https://schema.org/docs/hotels.html (2021-0416)
23. Schema.org Section: auto, 2021. URL: https://schema.org/docs/auto.home.html (202104-14)
24. Schema.org Section: bib, 2021. URL: https://schema.org/docs/bib.home.html (202104-14)
25. Schema blog: GS1 Web vocabulary: welcoming the first schema.org external extension, 2016. URL: http://blog.schema.org/2016/02/gs1-milestone-first-schemaorgexternal.html (2021-04-14)
26. Schema Generator & Structured Data Testing Tool. URL: https://seoscout.com/tools/schema-generator (2021-04-20)
27. Structured Data Markup Helper. URL: https://support.google.com/webmasters/answer/3069489?hl=en (2021-04-20)
28. Van de Rakt, Marieke. SEO basics: What is structured dana?, 2020. URL: https://yoast.com/what-is-structured-data/ (2021-04-06)
29. Yandex Webmaster: Structured data validator. URL: https://webmaster.yandex.com/tools/microtest/ (2021-04-20)
30. WebMoves: Internet Strategists. URL: https://www.webmoves.net/tools/microdata (2021-04-21)
31. What is linked data? URL: https://jsld.org/ (2021-04-02)
32. Wordpress.org: plugins. URL: https://wordpress.org/plugins/tags/schemaorg/ (2021-04-20)