

Mogućnosti ostvarenja otvorenog rodoslovlja semantičkim tehnologijama - Acta Croatica 3.0

Davor Lauc, Donat Sorić, Darko Vitek

Tražeci ishodište interesa za povijesnu znanost i historiografiju, nerijetko ćemo pozornost usmjeriti k rodoslovnim istraživanjima. Izgradnja historiografskog narativa o predcima, u ne tako davnoj prošlosti, bila je rezervirana za osobe plemenita roda, koje su imale tradiciju pismenosti i koje su svojim rodoslovljima često dokazivali i potvrđivali svoj povlašteni društveni položaj. Masovnim opismenjavanjem, demokratizacijom školstva, ali isto tako i preobrazbom državnih i crkvenih institucija te razvojem administracije, nastaju mnogobrojni zapisi o pojedincima, koji postaju dostupni i širim društvenim slojevima. To će rezultirati činjenicom da se znanje o predcima naglo povećava, što će voditi i većem interesu za genealoška istraživanja. Sliku arhiva, u kojem sjede malobrojni znanstvenici, zadubljeni u povijesne spise, danas je zamijenila slika arhivskih korisnika, koji sjede za računalima i pretražuju zapise o svojim predcima. Velikom interesu za genealogije svjedoče i mnogobrojni internetski servisi, koji nude mogućnosti istraživanja i izrade rodoslovnih stabala. Dapače, genealoški sadržaji na Internetu nalaze se u samom vrhu traženosti i uvelike su nadmašili interes iskazan, primjerice, za sport ili politiku.

U takvim, promijenjenim okolnostima, postavljaju se i nova pitanja vezana uz rodoslovlje. Jedno od njih je i mogu li se sintetizirati mnogobrojna pojedinačna genealoška istraživanja, objediniti nepregledni podaci o različitim osobama te pružiti mogućnost, kako

amaterima da saznaju nešto više o svojim predcima starijim od treće ili četvrte generacije, do kada najčešće obiteljska predaje seže, tako i profesionalcima da valoriziraju svoja dosadašnja istraživanja i obogate ih novim podacima. Odnosno, može li se konstruirati otvoreno obiteljsko stablo, koje bi moglo obuhvatiti pojedinačna genealoška istraživanja i omogućiti istovremeni rad velikom broju korisnika na zajedničkom obiteljskom stablu.

Odgovor na ta pitanja potrebno je, između ostaloga, tražiti u novim informacijskim tehnologijama i njihovim implementacijama u genealoškim istraživanjima.

Otvoreno rodoslovlje – izazovi

Veliko, globalno genealoško stablo je zamisao za čiju je realizaciju potrebno konceptualno odgovoriti na nekoliko problema.

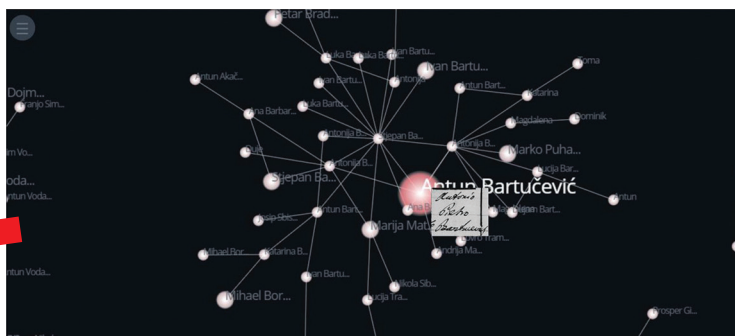
U prvom redu potrebno je izgraditi platformu, koja će biti otvorena prema različitim tipovima izvora, što je preduvjet izgradnje otvorenog genealoškog stabla. Nedvojbeno, kao primarni izvor treba istaknuti digitalizirane arhivske zbirke, koje u ovom slučaju možemo okarakterizirati kao djelomično strukturirane izvore, no potrebno je otvoriti mogućnosti učitavanja velike količine tiskane građe i njezino računalno prepoznavanje, a ne treba zanemariti niti digitalizirane privatne zbirke dokumenata, za koje ne možemo pretpostaviti neku određenu strukturu.

Nadalje, kako profesionalni i amaterski korisnici nemaju iste zahtjeve, potrebno je omogućiti pristup podacima različitim razinama korištenja. Amaterski korisnici, čiji interes za obiteljsko stablo staje na razini obiteljske predaje, nemaju niti približne tehničke ni sadržajne zahtjeve kao profesionalci, koji koriste, primjerice, razinu strukturiranosti podataka na razini GRAMPS-a (**Genealogical Research and Analysis Management Programming System**), koji je vjerojatno najzastupljeniji u profesionalnim genealoškim istraživanjima.

Na razini organizacije podataka, budući da je očekivano da globalno obiteljsko stablo sadrži velike količine podataka, nužno je razlikovati izvore od zaključaka izvedenih iz izvora te pravilno valorizirati iste. Primjerice, nerijetko zaključivanje o predcima na temelju parcijalnih podataka (poput, imena i prezimena), a ne uzimajući u obzir demografske pravilnosti, kao što je prosječna životna dob, prosječno doba ženidbe, rađanja i sl., može voditi do krivih zaključaka. U tom smislu je nužno atomizirati sve podatke i razdvojiti objekte od relacija te se koristiti tehnikama umjetne inteligencije, kojima bi se nudila vjerojatnost točnosti pojedinih zaključaka.

Možda najkompleksniji segment arhitekture otvorenog genealoškog stabla je suradnja velikog broja korisnika na istim podacima. Naime, u okolnostima rada više korisnika na istom podatku postavlja se pitanje kontradiktornosti podataka (o istoj osobi iz izvora mogu biti izvedeni suprotstavljeni zaključci), što otvara problematiku opovrgavanja i potvrđivanja tvrdnje, odnosno zaključka. Kao polazišno načelo u rješenju tog kompleksnog problema je identificiranje sljedbenosti podataka, odnosno nužno je svakoj atomičnoj informaciji znati izvorište i svako njezino korištenje i interpretiranje.

Acta
Croatica 3.0 -
Vizualizacije
rodoslovlja Visa



 **INFO**

 Portal Acta Croatica: <https://actacroatica.com/hr/>

Na kraju, a ne manje važno, je i pitanje zaštite osobnih podataka i sigurnosti. Reguliranje unosa podataka o živim osobama te dijeljenje tih podataka s drugim korisnicima otvara mnogobrojne probleme, čije rješenje iziskuje pažljivu organizaciju dostupnosti podataka. Usklađenost s zakonodavnim okvirom je samo jedan segment tog problema.

O kompleksnosti rješenja prikazanih izazova za realizaciju otvorenog obiteljskog stabla svjedoči i činjenica da ono još uvijek nije realizirano, unatoč velikom interesu iskazanom za njega. Najbliže funkcionalnom otvorenom obiteljskom stablu se približio projekt WikiTree, kojega je 2008. pokrenuo Chris Whitten.

Semantički web u kontekstu genealogije

Preobrazba web tehnologije, koja se kretala od ručno pisanih HTML stranica, a karakterizira ih statičnost, preko dinamičnih stranica generiranih računalom namijenjenih za pretraživanje, odnosno logičko procesuiranje od strane ljudi, do treće generacije weba, koja omogućuje strojno učenje informacija dostupnih na Internetu (web 3.0 ili semantički web), otvorila je i nove mogućnosti u konstruiranju otvorenog obiteljskog stabla.

Osnovne značajke semantičke tehnologije možemo opisati kao prelazak sa dokumenta na manji entitet, podatak, koji se povezuje putem relacija u različite mreže podataka (Heath i Bizer, 2011). U tom smislu semantički web nije orijentiran na odnose mrežnih stranica, nego na odnose entiteta i njihovih svojstava, koji se utvrđuju pomoću njihovih metapodataka. Na primjeru običnog web pretraživanja tehnologija semantičkog weba omogućila bi pregledavanje vjerojatnih rezultata za upit o, primjerice banu Josipu Jelačiću, na temelju upita strukturiranog kroz metapodatke; muškarac koji je živio u prvoj polovici 19. stoljeća, general i hrvatski ban, rođen u Petrovaradinu, dok bi u klasičnom web okruženju za zadovoljavajući rezultat pretrage bilo potrebno u tražilicu upisati „Josip Jelačić“.

U kontekstu genealoških istraživanja u semantičkom web okruženju prilikom pretraživanja nepoznatih predaka upit bi trebalo strukturirati kroz poznate metapodatke, primjerice, lokacija, približno vrijeme življenja, zanimanje, ime i prezime sina ili unuka i sl., što bi u kombinaciji s prepoznavanjem istraživačke prakse i nekim osnovnim demografskim i genealoškim pravilima rezultiralo ponuđenim popisom osoba koji zadovoljavaju u određenoj mjeri postavljene kriterije pretrage.

Pogodnosti semantičke tehnologije u genealoškim istraživanjima je nemjerljiva i sve više prepoznata. O semantičkoj tehnologiji se razmišlja i među vodećim genealoškim servisima poput FamilySearch, no njezina potpuna implementacija još uvijek nije zaživjela.

Projekt Acta Croatica 3.0

Projekt Acta Croatica 3.0 nadovezuje se na postojeći projekt ActaCroatica.com, koji je uz podršku ključnih institucija u zemlji i inozemstvo pokrenuo digitalizaciju i diseminaciju genealoških podataka u domovinskoj i iseljenoj Hrvatskoj. U sljedećoj je fazi cilj stvoriti tehnološke i sadržajne pretpostavke izgradnje hrvatskog otvorenog rodoslovlja. Ideja je omogućiti suradnju profesionalnih istraživača, entuzijasta, kao i povremenih korisnika na izgradnji zajedničkog otvorenog obiteljskog stabla svih Hrvata i stanovnika Hrvatske. U ožujku 2018. godine, na adresi vis.actacroatica.com, pokreće se probno otvoreno rodoslovlje Otoka Visa, koje će se tijekom godine proširiti na cijelu Hrvatsku. ■

Acta Croatica 3.0 - Vizualizacije rodoslovlja Visa



Primjer pristupa matičnoj knjizi, tradicionalnom rodoslovnom izvoru



| Della Parrocchia di S. Pietro e Paolo | | Grado di Dalmazia | | ATTI DI NASCITA | | Distretto di Dubrovnik | |
|---------------------------------------|--|--------------------|----------------------------|----------------------|----------------------|------------------------|-------------|
| Numero | Data, Ora, Nome e Cognome del Padre | INSEDIAMENTO | Nome e Cognome della Madre | INDICAZIONE | DEI GENITORI | RELIGIONE | NOTE, GONDO |
| 17 | Ludovica Vincenza Maria Mardesich (Bukurega Vicka Marija Mardesic) | St. Pietro e Paolo | Vukari | Di S. Pietro e Paolo | Di S. Pietro e Paolo | Cattolica d'Occidente | FABRINI |