

Rekomendācijas Atsauču datu semantiskās sadarbības veicināšanai starp Latvijas kultūras mantojuma institūcijām

DOI: <https://doi.org/10.52197/IBIV6839>

Kopsavilkums

Latvijas Nacionālās bibliotēkas Bibliogrāfijas institūts, 2021.

Izstrādes darba grupa: Mārīte Apenīte, Uldis Bojārs, Anita Goldberga, Maira Kreislere, Anita Rašmane, Jogita Sauka.

Rekomendācijas sagatavotas Latvijas Nacionālās bibliotēkas (LNB) un Latvijas Universitātes projektā [“Latvijas atmiņas institūciju dati digitālajā telpā: vienojot kultūras mantojumu”](#) (Nr. Izp-2019/1-0365), ko Fundamentālo un lietišķo pētījumu programmas ietvaros finansē Latvijas Zinātnes padome.

Rekomendācijas attiecas tikai uz kultūras mantojuma atveidei nepieciešamajām kontrolētajām vērtībām, ko nodrošina autoritatīvie dati un citas vērtību vārdnīcas, un norobežojoties no kultūras mantojuma objektus aprakstošajiem metadatiem, kas nav šo rekomendāciju tvērumā.

Izvērtējot datus no atsevišķām nozaru sistēmām, to datu modeļus un izmantotos standartus, autori secina, ka arhīvos, bibliotēkās un muzejos tiek izmantoti atšķirīgi datu formāti ar dažādu elementu strukturētības pakāpi, daļēji atšķirīgas atslēgentītes, kas ierobežo sadarbību vienotu kontrolētu datu elementu vērtību izmantošanā, un vērojamas citas nekonsekvences. Sadarbībai nosaukto entīšu vienveidīgai un kontrolētai lietošanai Vienotajā kultūras mantojuma digitālo objektu pārvaldības un izplatīšanas platformā, konkrēti Atsauču datu risinājuma izveidē, nepieciešama vienošanās par jaunu datu modeli, datu standartiem un formātiem, kas atbilst semantiskā tīmekļa principiem.

Jaunākajos arhīvu, bibliotēku un muzeju datu konceptuālajos modeļos datu elementu semantika izvērtēta no lietotāju vajadzību viedokļa. Rekomendācijās tiek analizētas trīs saskarnozaru konceptuālo modeļu (Records in Context Conceptual Model, IFLA Library Reference Model un CIDOC Conceptual Reference Model) entītes, kas ir būtiskas sadarbībai Latvijas Vienotajā kultūras mantojuma digitālo objektu pārvaldības un izplatīšanas platformā vienveidīgai un kontrolētai dažādu vārdu/nosaukumu, identificēšanai un lietošanai. Analīze tiek veikta, lai identificētu kopīgās un atšķirīgās

atslēgentītes un to galvenos atribūtus un raksturīgākās attiecības, kas varētu palīdzēt praktiskā datu elementu kartēšanā.

Ņemot vērā esošās situācijas analīzi Latvijas atmiņas institūcijās un datu konceptuālo modeļu atslēgentīšu salīdzinājuma rezultātus, Kultūras mantojuma entīšu reģistru (KMER, kas ir Atsauču datu risinājuma pamats) nepieciešams veidot, balstoties uz LNB un Latvijas Nacionālā arhīva autoritatīvajiem datiem, jo tie ietver informāciju par būtiskajām kultūras mantojuma entītēm. Pakāpeniski jāpāriet uz galveno nosaukto entīšu – personu, institūciju, dzimtu, vietu, notikumu, darbu, tematu un formas/žanra – kontrolētu lietojumu arī citu atmiņas institūciju datos. Saistībā ar jaunu datu modeli nepieciešams vienoties:

- kuras datu elementu vērtības iekļaut KMER datu modelī, kuras ņemt no vērtību vārdnīcām;
- par kopīgiem noteikumiem, kas novērstu pieraksta atšķirības (tas attiecas uz personu vārdiem, datējumiem u.c. entītēm un to īpašībām);
- par datu granularitāti (piemēram, vārds, uzvārds; institūcija, tās apakšstruktūra u.tml.);
- par datu kvalitātes/uzticamības līmeņiem, kas ļautu datu veidotājiem atšķirt uzticamas un pārbaudītas datu vērtības no mazāk uzticamām;
- par datu oriģinālās izcelsmes informācijas saglabāšanu, t.sk. identifikatoriem, jo izcelsmes informācija kalpo kvalitātes un atbilstības izvērtēšanai;
- par starpnozaru datu, kas attiecas uz KMER, kartēšanas principiem, pārejot uz entīšu pārvaldību;
- par datu modeļa paplašināšanas iespējām;
- par iespējām attīstīt un ieviest praksē saistīto datu formātus, jo neviena no esošajām kultūras mantojuma institūciju sistēmām tiešā veidā nenodrošina autoritatīvo datu izguvi saistīto datu formātos, lai tos varētu atvērt plašai izmantošanai gan savstarpēji, gan ārpus kultūras nozares.

Semantiskās tehnoloģijas piedāvā jaunus informācijas strukturēšanas principus, lai radītu ne tikai mašīnlasāmus, bet arī mašīninterpretējamus datus – RDF, SPARQL, SKOS u.c. standarti nodrošina datu struktūras atkārtotai izmantošanai. Saistītos datus var izmantot, lai pievienotu kontekstuālu informāciju un izveidotu savienojumu tīklu, kas labāk atspoguļo zināšanas reālajā pasaulē. Konteksts un savienojumi var palīdzēt pētniekiem gūt pilnīgāku izpratni par materiāliem, aicinot sabiedrības grupas vairāk iesaistīties to izmantošanā. Saistītie dati ir jaudīgi arī tāpēc, ka tie nav saistīti ne ar vienu konkrētu sistēmu, bet gan integrē starpkolekciju saturu, tādējādi vadot lietotāju ne tikai starp kolekcijām, bet arī starp dažādām krātuvēm.

Rekomendācijās apkopoti arī labās prakses piemēri no risinājumiem ārpus Latvijas (SNAC, Finto/FINNA, Wikidata, EDM, Interim), no kuriem izriet apkopotie ieteikumi turpmākajai sadarbībai kultūras mantojuma politikas veidotājiem un ieviesējiem Latvijā. Būtiskākās izmaiņas, kas atšķiras no arhīvu, bibliotēku un muzeju ierastās prakses un kas jāievieš, veidojot vienotu kultūras mantojuma Atsauču datu risinājumu ar entīšu reģistru, ir autoritatīvo ierakstu (LNA, LNB) entifikācija, izcelšana no slēgtām sistēmām un atvēršana publiskai piekļuvei atvērtu vērtību vārdnīcu/ontoloģiju formā. Citiem vārdiem – kultūras mantojuma datu kvalitātes kontrolē jāpāriet no autoritatīvo ierakstu datu modeļa uz entīšu datu modeli, kur katru entītes instanci un tās īpašības (piemēram, nosaukumus dažādās valodās un rakstībās) un attiecības jeb saistību definē, izmantojot mašīnlasāmus identifikatorus. Tādējādi aprakstošo metadatu veidotājiem savās institūciju pamatsistēmās atliks vien pievienot saiti uz attiecīgo entīti, un konkrētā

informācija vairs nebūs manuāli jāpārraksta. Galveno entīšu klašu informācija jāglabā atsevišķi no kultūras mantojuma objektu aprakstošajiem metadatiem, t.i., jāglabā KMER.

Šāda vienota kvalitātes kontrole nozīmē, ka no jauna jārada kultūras mantojuma entīšu pārvaldībai un attīstībai atbilstoša infrastruktūra, kurā var definēt entītes tipu un atsevišķi aprakstīt katras entītes īpašības, norādot tās saistību ar citām entītēm. Jābūt iespējai entītes informāciju mainīt un papildināt, nezaudējot saikni ar pašiem kultūras mantojuma objektiem, kas var būt aprakstīti institūciju pamatsistēmās.

Rekomendācijas ir pirmais mēģinājums aprakstīt vajadzības jaunam starpnozaru sadarbības datu modelim, lai no ierakstiem slēgtās sistēmās pārietu uz atvērtu kultūras mantojuma galveno entīšu un to īpašību-attiecību risinājuma izveidi, izmantojot atvērto saistīto datu standartus. Tikai šāds risinājums atvērtu vērtīgu zināšanu bāzi izmantošanai gan kultūras mantojuma nozarēs, gan citiem interesentiem, kā arī uzlabotu piekļuvi pētniekiem, kā arī radošo industriju, izglītības, tūrisma u.c. nozaru pārstāvjiem. Atsauču datu risinājumam jāklūst par kultūras mantojuma nozares digitālās transformācijas un digitālo objektu pārvaldības un izplatīšanas platformas pamatu.

LNB kā nacionālais agregators nodrošina kultūras mantojuma nozaru datu piegādi Europeana. Līdz ar to pāreja uz entīšu sistēmu, kas savietojama ar Europeana Data model (EDM), ir būtiska, lai nodrošinātu datu piegādi Europeana ar pēc iespējas minimālākiem datu kvalitātes zudumiem.

[Pilns rekomendāciju teksts](#)