

# Strateginė Lietuvos bibliotekų kryptis – susietieji atvirieji duomenys

Regina Varnienė-Janssen, Jūratė Kuprienė

Vilniaus universiteto Komunikacijos fakultetas  
Faculty of Communication at Vilnius University  
[regina.varniene-janssen@kf.vu.lt](mailto:regina.varniene-janssen@kf.vu.lt)  
[jurate.kupriene@gmail.com](mailto:jurate.kupriene@gmail.com)  
ORCID <https://orcid.org/0000-0001-8755-3625>

**Santrauka.** Straipsnio objektas yra *susietųjų atvirųjų duomenų* paradigma, daranti įtaką kultūros politikos pokyčiams, skatinanti transformuoti bibliotekų duomenų modelius ir modernizuoti arba kurti naujas informacijos sistemas naudojant RDF modelį bei semantinio saityno technologijas, grindžiamas OWL, SKOS ar SPARQL. *Pagrindinis tikslas* – identifikuoti *susietųjų atvirųjų duomenų* paradigmą ir reikalavimus, keliamus *susietiesiems atviriesiems duomenims*; nustatyti bibliotekų duomenų modelių plėtrą *susietųjų atvirųjų duomenų* kryptimi ir metodologinius sprendimus, leidžiančius transformuoti dabartinę katalogavimo praktiką bei modernizuoti informacines sistemas, užtikrinančias *susietųjų atvirųjų duomenų* rengimą mūsų šalyje bei jų sklaidą semantiniame saityne. Tyrimas grindžiamas fenomenologine metodologija, kuria remiantis siekiama išaiškinti ir suprasti *susietųjų atvirųjų duomenų* paradigmą kaip fenomeno objektyvų tapsmą ir šio fenomeno poveikį socialinei aplinkai. Tyrimo metu nustatytos *susietųjų atvirųjų duomenų* suformavimo kryptys, tikslai ir metodologinės jos realizavimo galimybės bibliotekų sektoriuje, o išanalizuotos užsienio šalių bibliotekose įgyvendinamos gerosios praktikos *susietųjų atvirųjų duomenų* iniciatyvos patvirtina investicijų svarbą pridėtinę vertę kuriančioms integralioms informacinėms paslaugoms. Tezaurą būtina keisti formalia ontologija, transformuoti bibliotekose įsitvirtinusių katalogavimo praktiką, taikant bibliotekų duomenų modelius, grindžiamus nauja bibliografinė paradigma – *kūrinio, išraiškos, apraiškos ir vieneto* esybių aprašymu ir disponuojančia esybių, atributų ir santykių struktūra, bei aprūpinti informacinę sistemą specialia semantinės paieškos programine įranga. Straipsniu taip pat siekiama inicijuoti platesnę diskusiją ne tik tarp bibliotekų, bet ir atkreipti politikų, atsakingų už kultūros politikos formavimą, dėmesį į tai, kad ir bibliotekų plėtrai skirtame dokumente svarbu strategiškai pagrįsti šalies bibliotekų duomenų valdymo pertvarką *susietųjų atvirųjų duomenų* kryptimi, akcentuojant bibliotekų sektoriaus duomenų integralumą ir numatant reikalingą finansinę paramą šiai pertvarkai įgyvendinti.

**Reikšminiai žodžiai:** susietieji atvirieji duomenys, atvirųjų duomenų kultūros politika, semantinio saityno technologijos, bibliotekų duomenų modeliai, naujoji bibliografinė paradigma.

## Linked Open Data as a Strategic Orientation of Lithuanian Libraries

**Abstract.** The article builds upon research into the paradigm of *linked open data*, which impacts developments of the cultural policy as well as encourages the transformation of library data models and modernization of information systems or setting up new ones by using the RDF model and Semantic Web technologies based on OWL, SKOS or SPARQL. *The article's main purpose* is to identify the structure of the paradigm of linked open data and the requirements for such data, elaborate on the development of library data models focusing on linked open data as well as to identify methodological solutions allowing to transform the current cataloguing practice and modernize information

**Received:** 12/2/2021. **Accepted:** 12/4/2021

Copyright © 2021 Regina Varnienė-Janssen, Jūratė Kuprienė. Published by Martynas Mažvydas National Library of Lithuania

This is an Open Access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

systems that ensure the creation of *linked open data* in Lithuania and their dissemination within the Semantic Web. The research determined guidelines for shaping a paradigm of *linked open data*, its purposes, and the methodological potential for its implementation within the library sector. The analysed best practice initiatives related to *linked open data* undertaken by libraries abroad confirmed the importance of investing into integrated information services which create added value. The research also allowed drawing conclusions concerning the moving of libraries toward tendencies of creating *linked open data*. Thesauri should be replaced by formal ontologies, the entrenched cataloguing practice should be transformed by applying library data models based on the new bibliographic paradigm which requires a description of the entities Work, Expression, Manifestation, and Item as well as possesses a structure of entities, attributes, and relationships, and information systems should be supplied with special software for semantic search. The article also has the purpose to encourage broader discussion among Lithuania's libraries as well as draw the attention of cultural policy makers to the necessity that, in the strategic document for library development, the transformation of library data management be oriented at linked open data by emphasizing the integrity of data within the library sector and that financial support for the realization of such transformation be provided.

**Keywords:** linked open data, open data policy, Semantic Web technologies, library data model, new bibliographic paradigm.

## Įvadas

2016 m. papildytame Lietuvos strateginiame dokumente „Bibliotekų plėtros strateginės kryptys 2016–2022 m.“ akcentuojamas reikšmingas bibliotekų vaidmuo šalies kultūros politikos kontekste, stiprinant visuomenės tapatybę, kūrybingumą, plėtojant socialinį kapitalą. Nors bibliotekos tipologiniu aspektu skiriasi, kiekviena jų rinkinių sklaida ir teikiamomis informacinėmis paslaugomis svariai prisideda prie edukacijos procesų, šalies kultūros paveldo aktualinimo ar vykdomų mokslinių tyrimų. Informacinėms paslaugoms teikti reikia kokybiškų duomenų. Tokiais duomenimis disponuoja bibliotekos, nes čia kaupiamas informacijos turinys įvairiose medijose: garsas, vaizdas, tekstai tradicinėse medijose, taip pat ir skaitmeniniai išteklių, skelbiami kultūros ir mokslo tematika, įvairiausiomis pasaulio šalių kalbomis. Šie išteklių yra kataloguojami bibliotekose dirbančių profesionalų, o metaduomenys yra saugomi informacinėse sistemose, suteikiančiose prieigą prie šios informacijos. Visa tai užtikrina bibliotekose kaupiamų išteklių ir duomenų apie juos patikimumą. Bendradarbiaudamos su švietimo ir mokslo institucijomis, kitais kultūros sektoriais ir verslu, bibliotekos tampa pagrindinėmis mokslo informacijos ir žinių perdavimo proceso dalyvėmis ir svarbiausiomis tautos kultūros paveldo aktualinimo pasaulyje tarpininkėmis. Vienas ryškiausių tokio bendradarbiavimo pavyzdžių kultūros sektoriuje yra Lietuvos nacionalinės Martyno Mažvydo bibliotekos 2005–2008 m. ir 2009–2012 m. įgyvendinta *Virtuali elektroninio paveldo informacijos sistema* (toliau – VEPIS), kuri ir toliau plėtojama taikant semantinio saityno (*Web 3*) technologijas su informacijos sklaidos portalu [www.epaveldas.lt](http://www.epaveldas.lt). Jūratės Kuprienės, Alberto Šermoko ir Reginos Varnienės-Janssen atlikti tyrimai<sup>1</sup> patvirtina tai, jog ši sistema, integruojanti įvairių kultūros sektoriaus ir švietimo institucijų skaitmenintus objektus, atitinka *Nestor* darbo grupės suformuluotus skaitmeninių saugyklų patikimumo kriterijus<sup>2</sup> ir semantinio saityno reikalavimus, nes yra grindžiama OAI<sup>3</sup>, CIDOC CRM<sup>4</sup>

<sup>1</sup> Varnienė-Janssen, R.; Kuprienė, J. Authenticity and Provenance in Long-Term Digital Preservation: Analysis of the Scope of Content. *Informacijos mokslai*, 2018, t. 82, p. 131–160; Varnienė-Janssen, R.; Šermokas, A. Ontologies and Technologies for Integrating and Accessing Digital Cultural Heritage: Lithuanian Approach. *Informacijos mokslai*, 2020, t. 88, p. 66–82; Varnienė-Janssen, R.; Šermokas, A. Provenance in the Context of Digital Cultural Heritage Content the Lithuanian Approach. In: *Authority, Provenance, Authenticity, Evidence: Selected Papers from the Conference and School Authority, Provenance, Authenticity, Evidence, Zadar, October 2016* / Ed. by M. Willer, A. J. Gilliland, M. T. Zadar: Sveučilsteu Zadaru, 2018, p. 213–257.

<sup>2</sup> Catalogue of Criteria for Trusted Digital Repositories. *Nestor Working Group on Trusted Repositories Certification, Version 1. Deutsche Nationalbibliothek*, 2006. Prieiga per internetą: [https://files.dnb.de/nestor/materialien/nestor\\_mat\\_08-eng.pdf](https://files.dnb.de/nestor/materialien/nestor_mat_08-eng.pdf). Žiūrėta 2021-03-15.

<sup>3</sup> ISO 14721:2012 *Space Data and Information Transfer Systems – Open Archival Information System (OAIS) – Reference Model*. Prieiga per internetą: <https://www.iso.org/standard/57284.html>. Žiūrėta 2021-03-15.

<sup>4</sup> *CIDOC Conceptual Reference Model. Version 5.1.2.*, 2013. Prieiga per internetą: [http://www.cidoc-crm.org/sites/default/files/cidoc\\_crm\\_version\\_5.1.2.pdf](http://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_5.1.2.pdf). Žiūrėta 2021-03-15.

ir RDF<sup>5</sup>. OAIS modelis, kuriuo yra grindžiamas VEPIS funkcionalumas, leidžia nuosekliai registruoti sistemoje vykstančius duomenų ir informacijos pokyčius. Šių pokyčių pagrindas yra skaitmeninė prove-niencija – saugomų duomenų patikimumo ir išteklių autentiškumo įrodymas, turintis ypač didelę reikšmę informacijos vartotojui. Kadangi informacijos paieška sistemoje organizuojama ontologijų pagrindu – CIDOC CRM, OWL ir RDF, VEPIS užtikrina laisvai pasirenkamą duomenų apdorojimo būdą nepri-klausomai nuo konkrečios programinės įrangos galimybių. Šios VEPIS galimybės atitinkamai vertinamos trečiu duomenų teikimo brandos lygiu. Duomenys lengvai publikuojami, tačiau duomenų eksportui iš uždarų formatų reikia papildomų įrankių. A. Šermokas ir R. Varnienė-Janssen, pripažindami, kad VEPIS jau yra įsitvirtinusi tarp nacionalinio lygmens sistemų ir įvertinta gana aukštu trečiu duomenų brandos lygiu, taip pat atkreipia dėmesį į tai, kad ši sistema kol kas dar neužtikrina galimybės gauti užklausą papildančios semantiškai artimos informacijos iš nesusijusių su VEPIS sistemų<sup>6</sup>. Tai leidžia daryti išvadą, kad vis dar lieka erdvės tolesnei VEPIS plėtrai, kuri būtų vertinama penktu duomenų brandos lygiu.

Lyginant kitų Lietuvos bibliotekų sistemų funkcionalumus su VEPIS, matomas akivaizdus jų skir-tumas. Priešingai nei VEPIS, kur paieška organizuojama duomenų susiliejimo būdu, panaudojant on-tologijas ir semantinio saityno technologijas, kitų Lietuvos bibliotekų informacinėse sistemose paieška grindžiama modeliu, siūlančiu vartotojui galimybę susipažinti su iš anksto nustatyto duomenų rinki-niu. Tokio modelio sisteminė paieška leidžia surasti bibliotekų fiziniuose ar skaitmeniniuose ištekliuose reikiamą informacijos išteklių, dažniausiai – knygą, žurnalą ar jo straipsnį, mokslinį darbą, kultūros paveldo objekto skaitmeninę kopiją. Paprastai formuluodamas užklausą vartotojas turi įvesti reikšmi-nius žodžius, tiksliausiai nusakančius jo poreikį. Ir tik nuo vartotojo gebėjimo formuluoti užklausą bei bibliotekos išteklių formavimo pagrindų išmanymo priklauso paieškos rezultato sėkmė. Nors bibliotekų paieškos sistemos tobulinamos ne vieną dešimtį metų, paieška šiose sistemose rezultatyviausia tada, kai vartotojui pavyksta pateikti užklausą ne šnekamosios kalbos, bet reikšminiais žodžiais, kurie vartojami bibliotekos išteklių katalogavimo sistemose. Tad tenka pripažinti, kad iki šiol tarp vartotojo ir bibliote-kos išteklių išlieka nepasiekiamumo problema: vartotojui rasti yra sudėtinga, o biblioteka nėra užtikrin-ta, jog ištekliai yra sukataloguoti taip, kad vartotojas juos suras. Šios problemos priežastis yra itin kom-pleksiški ir pernelyg detalūs, standartizuoti ir specifiskai bibliotekų veiklai pritaikyti duomenų modeliai, ne tik apsunkinantys informacinę paiešką vartotojams, bet ir neužtikrinantys šių modelių sąveikumo su kitų institucijų sistemomis semantiniame saityne. Kita vertus, laikytis tarptautiniu mastu suderintų metaduomenų formatų – UNIMARC, MARC21 ar DC, metaduomenų turinio standartų – ISBD, RDA ar AACR, duomenų perdavimo protokolų – OAI-PHM reikalauja siekis užtikrinti bibliotekų informa-cinių sistemų funkcionavimo tvarumą realizuojant vis didėjančią funkcinių Europos ir pasaulio bibliotekų suderinamumą. Standartizuoti duomenų modeliai ir protokolai sudaro galimybę bibliotekoms keistis metaduomenimis, dalyvauti paskirstyto katalogavimo procesuose arba tarptautiniuose projektuose, pavyzdžiui, VIAF ar kultūros paveldo informacijos skaitmeninėje bibliotekoje *Europeana*. Naujos se-mantinio saityno galimybės ir jų pagrindų pakitę vartotojų informaciniai poreikiai skatina bibliotekas imtis esminės informacinių sistemų pertvarkos, kuri, Gordono Dunsire ir Mirnos Willer nuomone, yra pernelyg užsitęsusi<sup>7</sup>. Autoriai mano, jog lėtą pasaulio bibliotekų informacinių sistemų pertvarką lemia daug priežasčių: nepalankios ekonominės ir technologinės sąlygos, konkrečių institucijų apsisprendi-mas užtikrinti jau veikiančių bibliotekinių sistemų tvarų funkcionavimą bei IFLA organizacinio darbo trikdžiai rengiant naujus duomenų ir informacijos modelius, organizuojant tarptautinį jų aptarimą bei

<sup>5</sup> *RDF Schema 1.1*. W3C Recommendation 25 February 2014. Prieiga per internetą: <https://www.w3.org/TR/rdf-schema/>. Žiūrė-ta 2021-03-15.

<sup>6</sup> Plačiau žr. Varnienė-Janssen, R.; Šermokas, A. *Ontologies and Technologies for Integrating and Accessing Digital Cultural Heri-tage: Lithuanian Approach*, p. 66–82.

<sup>7</sup> Dunsire, G.; Willer, M. *Initiatives to Make Standard Library Metadata Models and Structures Available to the Semantic Web. World Library and Information Congress: 76th IFLA General Conference and Assembly, Gothenburg, Sweden – 10–15 August 2010*. Prieiga per internetą: <http://www.ifla.org/en/ifla76>. Žiūrėta 2021-03-03.

patvirtinimą. Informacijos valdymas dabartinėse informacijos sistemose tampa dar sudėtingesne problema, nes bibliotekos, siekdamos kuo geriau patenkinti vartotojų poreikius, pateikia paieškai ne tik savus išteklius, bet ir tuos, kurie prieinami atvirai ar iš dalies atvirai internete: kitų bibliotekų, archyvų bei muziejų išteklius, mokslo duomenų bazes, registrus, virtualių parodų objektus ir kitą turinį, atitinkantį bibliotekos profilį ir išteklių formavimo principus. Kadangi duomenų tipai, laikomi skirtingose informacijos sistemose, paprastai skiriasi, reikalingi specialūs sprendimai, galintys šiuos duomenis sujungti kartu bei įgalinti informacinę sistemą ne tik suprasti vartotojo užklausą, bet ir ją išversti į kalbą, suprantamą kitoms informacinėms sistemoms. Poreikį tobulinti bibliotekų informacines sistemas pagal meto ir vartotojų reikalavimus patvirtina daugelio šalių mokslininkų atlikti tyrimai, gerosios praktikos pavyzdžiai ir pagaliau Europos Komisijos (toliau – EK) atvirųjų duomenų strategija<sup>8</sup>. Mūsų šalies strateginio dokumento, skirto bibliotekų plėtrai 2016–2022 m., 12.3 ir 12.6 uždaviniuose<sup>9</sup> bibliotekos skatinamos kurti naujas elektronines paslaugas panaudojant įveiktą infrastruktūrą ir įsitraukiant į strategines valstybės informacijos išteklių valdymo veiklas bei struktūras. Dokumente nurodoma, jog bibliotekos, disponuojančios patikimais ir didelės apimties duomenų masyvais (angl. *Big data*), galėtų tapti sėkminga ir vartotoją orientuota ir atvirų inovacijų skatinimo grandimi. Kita vertus, didesnis bibliotekų įsitraukimas į strategines valstybės informacijos išteklių valdymo veiklas šiame dokumente nedetalizuojamas ir nesiejamas su EK atvirųjų duomenų strategija, skatinančia esminę viešojo sektoriaus institucijų duomenų pertvarką *susietųjų atvirųjų duomenų* kryptimi. Analogišką išvadą dėl valstybinio požiūrio stokos pertvarkant bibliotekų informacines infrastruktūras *susietųjų atvirųjų duomenų* kryptimi daro ir Diane Rasmussen Pennington ir Laura Cagnazzo, tyrusios platesnį bibliotekų duomenų atvėrimą pritaikant *susietųjų atvirųjų duomenų* technologiją Europos nacionalinėse bibliotekose<sup>10</sup>. Gauti duomenys leido tyrėjams daryti išvadą, jog beveik visose tirtose šalyse dėl *susietųjų atvirųjų duomenų* technologijų platesnio taikymo bibliotekose trūksta valstybinio požiūrio. Štai, pavyzdžiui, Nyderlanduose susietųjų atvirųjų duomenų praktika paliekama atskirų institucijų atsakomybei, o Škotijoje susietųjų atvirųjų duomenų iniciatyva traktuojama kaip specifinė semantinio saityno technologija, kuriai plačiau pritaikyti pakaktų semantinio saityno galimybių suvokimo skatinimo. Svarbu paminėti, kad daugelis šio tyrimo respondentų vienu klausimu sutaria: „jeigu duomenų atvėrimas tampa vyriausybinio lygiu palaikoma praktika, turi būti peržiūrėti prioritetai ir skiriamos atitinkamos lėšos“<sup>11</sup>. Ši nuomonė svarbi ir mūsų šaliai, nes minėtame strateginiame dokumente, skirtame Lietuvos bibliotekų plėtrai 2016–2022 m., semantinio saityno technologijų taikymas šalies bibliotekose apskritai neaktualizuojamas, nors dvejis metais anksčiau (2014 m.) paskelbtame Informacinės visuomenės plėtros komiteto (toliau – IVPK) *Lietuvos viešojo sektoriaus atvirų duomenų architektūros modelyje* jau buvo numatytas ir bibliotekų duomenų platesnis atvėrimas pagal semantinio saityno reikalavimus. Šis modelis yra universalus, tačiau konkrečių sektorių veiklos kontekstas ir naudojami duomenų valdymo standartai yra skirtingi. Todėl bibliotekų sektoriaus duomenų valdymo ypatumai yra aktualūs, keliantys sudėtingus iššūkius, nes bibliotekos privalės užtikrinti ne žemesnį nei trečią atvirųjų duomenų brandos lygį; skirtingų duomenų rinkmenų susiejimą, vertinamą penkių žvaigždučių atvirųjų duomenų teikimo brandos lygiu bei leidžiančiu įgyvendinti semantinę paiešką; atvertų duomenų palaikymą, kokybę ir aktualumą ilguoju laikotarpiu. Visų šių uždavinių kontekste iškyla ir kita problema – papildomų finansinių išteklių poreikis, kuris yra įvardijamas ir jau minėtame IVPK modelyje.

<sup>8</sup> *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Open data An Engine for Innovation, Growth and Transparent Governance*. Brussels, 12.12.2011 COM 882 Final. Prieiga per internetą: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF>. Žiūrėta 2021-04-03.

<sup>9</sup> Bibliotekų plėtros strateginės kryptys 2016–2022 metams. Lietuvos Respublikos kultūros ministro 2016 m. balandžio 29 d. įsakymu Nr. ĮV-344, p. 5–6. Prieiga per internetą: <https://lrkm.lrv.lt/uploads/lrkm/documents/files/Biblioteku%20pletros%20kryptys.pdf>.

<sup>10</sup> Rasmussen, P. D.; Cagnazzo, L. Connecting the Silos Implementations and Perceptions of Linked Data Across European Libraries. *Journal of Documentation*, 2019, Vol. 75, No. 5, p. 643–666.

<sup>11</sup> *Ibidem*.

Apie sistemų galimybes sujungti informaciją internete pirmasis pasauliui pranešė interneto išradėjas Timas Berners-Lee, pristatydamas semantinio saityno kelio žemėlapi, pats autorius įvardijo jį ateities planu, kuris dar nėra išbandytas<sup>12</sup>. Trečiosios kartos saitynas (angl. *Web 3.0*), grindžiamas saityno išplėtos metodologija, t. y. žiniomis apie išteklius, kuriomis operuoja ne tik žmogus, bet ir informacinės sistemos, telkiančios šias žinias į bendrą prasminį susietųjų duomenų tinklą. Tai neabejotinai pasaulio mastu svarbiausias pastarųjų dviejų dešimtmečių įvykis, nes, paties išradėjo T. Berners-Lee žodžiais, tuometiniam saitynui buvo suteikta daugiau prasmės ir motyvacijos skelbti duomenis saityne kompiuteriu skaitomu formatu<sup>13</sup>. Semantinio saityno atsiradimas turėjo didžiulės įtakos vartotojų informacinių poreikių bei jų informacinės elgsenos pokyčiams ir šiuo metu daugelio pasaulio šalių bibliotekų sistemos, įskaitant ir Lietuvos, jau nebetenkina pakitusių informacijos vartotojų poreikių. Tai patvirtina ir minėtas tyrimas<sup>14</sup>, ir kitų mokslininkų publikacijos šiuo klausimu. Pavyzdžiui, ilgąmetė IFLA ISBD peržiūros komiteto pirmininkė Elena Escolano Rodríguez, pristatydama ISBD pokyčių perspektyvas, kartu skatina atsižvelgti į laikmečio keliamus reikalavimus, daugiau domėtis bibliotekose naudojamų metaduomenų modelių pokyčiais bei šių modelių sąsajomis ir sumaniai jas taikyti praktikoje<sup>15</sup>. Analogišką išvadą pateikia ir W3C bibliotekų susietųjų duomenų grupės mokslininkai atlikto tyrimo ataskaitoje (*Library Linked data incubator group <...>*, 2011<sup>16</sup>). Tyrėjai daro išvadą, kad bibliotekų duomenys saityne nėra integruoti, dažnai yra išreikšti natūraliąja kalba, jiems taikomi tik bibliotekų bendruomenėje priimti standartai ir jų technologine plėtra rūpinasi ne informacijos profesionalai, o tiekėjai. Išsakytas nuomones palaiko ir kiti autoriai. Kimas Tallerås pagrindine priežastimi, dėl kurios bibliografiniai įrašai yra izoliuoti, įvardija MARC formato metodologiją, kurią vertinant naujos bibliografinės paradigmos požiūriu yra grindžiama vieneto aprašymo lygiu ir nedisponuojama esybių, atributų ir santykių struktūra. Todėl autorius neatsitiktinai siūlo orientuotis į ontologinį instrumentą, užtikrinantį bibliografinių duomenų parengimą pagal semantinio saityno metodologiją<sup>17</sup>. Su išsakyta mintimi siejasi ir Oreste'o Signore'o nuomonė, kad iki semantinio saityno atsiradimo sukurtose informacijos sistemose naudojami metaduomenų žodynai ir formatai neorganizuoja objektų į hierarchiją. Autoriaus nuomone, duomenų integravimas metaduomenų lygiu „<...> negali išnaudoti visų galimų ryšių gausos tarp skirtingų informacijos elementų. Ryšių mechanizmas taip ir lieka informacijos vartotojo galvoje.“<sup>18</sup> Tokius ryšius, autoriaus nuomone, užtikrina ontologija besiremiantis duomenų organizavimas ir semantinio saityno technologijos, nes naudojant RDF išteklių aprašymo formatą duomenys yra koduojami kaip triadų aibės: *veiksny* – *tarinys* – *papildinys*, kurios grindžiamos URI. Tai užtikrina papildomų duomenų išdavimą pagal pateiktą užklausą. Analogišką išvadą daro ir Rimvydas Laužikas, Jonas Juškys, Vygtintas Vaitkevičius bei R. Varnienė-Janssen, kurių nuomone, duomenis integruojant metaduomenų lygiu nepakankamai užtikrinamas jų struktūravimas, nes skirtingai traktuojami ir pateikiami istorinės geografijos, chronologijos bei biografistikos duomenys<sup>19</sup>.

<sup>12</sup> Berners-Lee, T. *The Semantic Web Road Map*, 1998, p. 1. Prieiga per internetą: [https://www.researchgate.net/publication/200034151\\_The\\_Semantic\\_Web\\_road\\_map](https://www.researchgate.net/publication/200034151_The_Semantic_Web_road_map). Žiūrėta 2021-03-06.

<sup>13</sup> *Ibidem*.

<sup>14</sup> Žr. Rasmussen, P. D.; Cagnazzo, L. Connecting the Silos Implementations and Perceptions of Linked Data Across European Libraries, p. 643–666.

<sup>15</sup> Plačiau žr. Rodríguez Escolano, E. RDA and ISBD: History of a Relationship. *JLIS.it.*, 2016, Vol. 7, No. 2, p. 49–82.

<sup>16</sup> *Library Linked Data Incubator Group Final Report. W3C Incubator Group Report 25 October 2011* / T. Baker, E. Bermès, K. Coyle, G. Dunsire, A. Isaac, P. Murray, P. Panzer, J. Schneider, R. Singer, E. Summers, W. Waiters, J. Young, and Zheng. Prieiga per internetą: [www.w3.org/Consortium/facts](http://www.w3.org/Consortium/facts). Žiūrėta 2021-03-12.

<sup>17</sup> Plačiau žr. Talaras, K. From Many Records to One Graph: Heterogeneity Conflicts in the Linked Data Restructuring Cycle. *Proceedings of the Eighth International Conference on Conceptions of Library and Information Science, Copenhagen, Denmark*, 2013, p. 19–22. Prieiga per internetą: <https://files.eric.ed.gov/fulltext/EJ1044646.pdf>. Žiūrėta 2021-03-12.

<sup>18</sup> Signore, O. *The Semantic Web and Cultural Heritage: Ontologies and Technologies Help in Accessing Museum Information*, 2007. Prieiga per internetą: <https://www.semanticscholar.org/paper/The-Semantic-Web-and-Cultural-Heritage-%3A-Ontologies-Signore/ab4f0ee826bbea097f4d7aa4d5244b67c0caaea6>. Žiūrėta 2021-03-12.

<sup>19</sup> Varnienė-Janssen, R.; Laužikas, R.; Vaitkevičius, V.; Juškys, J. Skaitmeninto kultūros paveldo vientiso turinio kūrimo strateginiai, metodologiniai ir techniniai sprendimai: lietuviškoji koncepcija. *Bibliografija 2010–2011: mokslo darbai*, 2013, p. 43–55.



VEPIS sistemoje išspręsti šias problemas padeda CIDOC CRM ontologija grindžiamas *Bendras asmenvardžių vietovardžių ir istorinės chronologijos tezaurus* (BAVIC), atliekantis informacijos iš įvairių šaltinių integravimo funkciją ir vykdamas informacijos išdavimą vartotojui. Apie platesnį bibliotekų duomenų atvėrimą visuomenei mūsų šalyje prabilta dar 2007 m. paskelbus Johno Byrumo straipsnį Nacionalinės bibliotekos leidinyje *Bibliografija*<sup>20</sup> ir 2010 m. *susietųjų atvirųjų duomenų* analizei paskyrus visą mokslo darbų rinkinį – *Bibliografija 2010–2011*<sup>21</sup>.

Tyrinėjant bibliotekų galimybę pritaikyti *susietųjų atvirųjų duomenų* metodologiją, padedančią gerinti pasikeitimą informacija šalies ir tarptautiniu mastu bei išvėgti kitas naujas galimybes, kyla nemažai klausimų. Kokiais sprendimais reikėtų vadovautis siekiant dabartinę katalogavimo praktiką pertvarkyti *susietųjų atvirųjų duomenų* pagrindu? Kurie gerosios praktikos pavyzdžiai yra jau pasiteisinę racionalizuojant bibliotekų katalogavimo praktiką *susietųjų atvirųjų duomenų* kryptimi? Kokie metaduomenų kūrimo, sąsajų ir sklaidos pokyčiai labiausiai atlieptų šiandieninius informacijos vartotojo poreikius? Kurie metodologiniai ir technologiniai sprendimai leistų paspartinti bibliotekų informacinių sistemų plėtrą kuriant modernias paslaugas, užtikrinančias integraciją su kitomis, pirmiausia šalies ir Europos Sąjungos šalių duomenis kaupiančiomis ir teikiančiomis infrastruktūromis? Atsakyti į šiuos klausimus padeda šis atliktas tyrimas.

*Straipsnio objektas* – *susietųjų atvirųjų duomenų* paradigma, daranti įtaką kultūros politikos pokyčiams, skatinanti transformuoti bibliotekų metaduomenų modelius bei modernizuoti arba kurti naujas bibliotekų informacijos sistemas, naudojant RDF modelį bei semantinio saityno technologijas, grindžiamas OWL, SKOS ar SPARQL.

*Straipsnio tikslas* – identifikuoti *susietųjų atvirųjų duomenų* paradigmos sąrangą ir reikalavimus, keliamus *susietiesiems atviriesiems duomenims*; nustatyti bibliotekų duomenų modelių plėtrą *susietųjų atvirųjų duomenų* kryptimi ir metodologinius sprendimus, leidžiančius transformuoti dabartinę katalogavimo praktiką bei modernizuoti informacines sistemas, užtikrinančias *susietųjų atvirųjų duomenų* rengimą mūsų šalyje ir jų sklaidą semantiniame saityne.

Tyrimas grindžiamas fenomenologine metodologija, kuria remiantis siekiama išaiškinti ir suprasti *susietųjų atvirųjų duomenų* paradigmos kaip fenomeno objektyvų tapsmą ir šio fenomeno poveikį socialinei aplinkai. Remiantis Johnu W. Creswellu, tyrėjo dėmesio centre – ne individai kaip naratyviniame tyrime, bet fenomenas ir tai, kaip šis fenomenas yra patiriamas skirtingų individų<sup>22</sup>. Fenomenologinė metodologinė nuostata, taikoma šiame tyrime, orientuoja ne tiek į paties fenomeno nagrinėjimą, kiek į jo reikšmę, nulemtą žmogiškosios sąmonės veiklos ir darančios poveikį kitoms žmogiškosios veiklos kryptims. Tyrėjas ieško to, kas bendra skirtingoms individų patirtims susidūrus su šiuo fenomenu. Ta bendra patirtis yra kaip realybės dalis ir yra fenomeno esmė. Šis tyrimas konstruojamas remiantis būtent tokia metodologine nuostata. Tyrimą sudaro keletas lygių. Pirmas – *susietųjų atvirųjų duomenų* paradigmos supratimo interpretacijos lygis yra grindžiamas Europos Sąjungos ir Lietuvos strateginių dokumentų, įžvelgiančių duomenų atvėrimo potencialą, jų galimą naudą ir nurodančių jų atvėrimo kryptį, analize. Antras – *susietųjų atvirųjų duomenų* paradigmos supratimo interpretacijos lygis yra grindžiamas reikalavimų, keliamų *susietųjų duomenų* aprašymui, skelbimui bei jų sujungimui internete, analize. Trečias – *susietųjų atvirųjų duomenų* paradigmos supratimo interpretacijos lygis yra grindžiamas įvairių šalių tyrėjų, įskaitant ir šio straipsnio autorės, patirčių, susidūrus su šiuo fenomenu, analize. Tokia tyrimo logika užtikrina perėjimo nuo atskirybės, individualių ir subjektyvių žinių į apibendrintą sąvokinių socialinės tikrovės aprašymo lygį.

<sup>20</sup> Byrum, J. D. Rekomendacijos kuo greičiau patobulinti OPAC ir valstybinės bibliografijos tarnybos vaidmuo sprendžiant šį uždavinį. *Bibliografija 2005–2006: mokslo darbai*, 2007, p. 14–23.

<sup>21</sup> *Bibliografija 2010–2011: mokslo darbai* / mokslinė redaktorė ir sudarytoja R. Varnienė-Janssen. Vilnius: Lietuvos nacionalinė Martyno Mažvydo biblioteka, 2013.

<sup>22</sup> Plačiau žr. Creswell, J. W. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. London: Sage Publications, 2009.

## 1. Susietųjų atvirųjų duomenų paradigma

### 1.1. Atvirųjų duomenų ir informacijos valdymo politikos paradigma

Šiuo metu Europos Sąjungos (toliau – ES) mastu yra įgyvendinama EK atvirųjų duomenų strategija<sup>23</sup>, kuria siekiama paspartinti naujų paslaugų plėtrą, paremtą inovatyviais informacijos derinimo arba naudojimo metodais, skatinančiais ekonomikos augimą ir didinančiais visuomenės aktyvumą. Atvirieji duomenys šioje strategijoje apibrėžiami kaip bendroji informacija, kuria laisvai naudotis, pakartotinai naudotis ir ją platinti gali bet kas – nemokamai ar už nedidelį mokestį. Įgyvendindamas EK atvirųjų duomenų strategiją IVPK inicijavo tyrimą, kuriuo buvo siekiama įvertinti Lietuvos nacionalinio lygmens viešojo sektoriaus teikiamų viešai prieigai duomenų ir metaduomenų struktūros modelių, sistemų funkcionalumus ir sukurti Lietuvos viešojo sektoriaus atvirųjų duomenų architektūros modelį<sup>24</sup>. Tyrimui buvo pasitelkta T. Berners-Lee pasiūlyta „penkių žvaigždučių schema“<sup>25</sup>, leidžianti įvertinti tiriamų rinkmenų atitiktį atviriesiems duomenims keliamiems reikalavimams bei Lietuvos informacijos sistemų gebėjimą generuoti ir vykdyti susietųjų duomenų paiešką. Tyrimo rezultatai atskleidė, kad „<...> Lietuvos skelbiamos informacijos rinkmenos netenkina dabartinių atvirųjų duomenų poreikių ir neatitinka ES lygiu skelbiamų rekomendacijų“. Tyrimu taip pat nustatyta, kad „<...> apie 24 % atvirųjų duomenų iniciatyvų atitinka antrą brandos lygį, t. y. duomenys skelbiami struktūrizuotų duomenų forma, uždarais formatais (pvz., MS Excel), 3 % – trečią, t. y. duomenys skelbiami struktūrizuotų duomenų forma, atvirais formatais (pvz., CSV), o didžioji dauguma Lietuvos atvirųjų duomenų iniciatyvų (apie 73 %) yra pirmo brandos lygio, t. y. duomenys skelbiami nestruktūrizuotais, neredaguojamais formatais (pvz., skenuoti dokumentai, PDF). Taip pat trūksta ne tik atvirųjų formatų ir standartų taikymo praktikos, bet taip pat ir standartų bei formatų integralumo“<sup>26</sup>. Nustatyti trūkumai mažina atvirųjų duomenų teikiamą naudą verslui, visuomenei ir pačioms institucijoms. Šiuo požiūriu *susietųjų atvirųjų duomenų* struktūra ir jų pateikimo pobūdis iš esmės skiriasi nuo ataskaitoje išvardytų duomenų modelių ir metaduomenų struktūros bei jų pagrindu grindžiamos paieškos rezultatų. *Susietųjų atvirųjų duomenų* pagrindu organizuojamos paieškos rezultatai kur kas labiau atitinka informacijos vartotojo lūkesčius. 2014 m. IVPK paskelbti tyrimo rezultatai taip pat leidžia daryti išvadą, jog *susietaisiais atviraisiais duomenimis*, atitinkančiais penktą duomenų brandos lygį, nedisponuoja nė viena mūsų šalies institucija. Akivaizdu, kad papildomo finansavimo prireiks ne tik kuriant bendrą šalies universalią atvirųjų duomenų platformą, bet ir institucijoms, teikiančioms į ją duomenis. Todėl stochastinis procesas, kuris dažniausiai taikomas kultūros projektams finansuoti mūsų šalyje, padėties neišgelbės. Padėtį svariai galėtų pagerinti suformuotas valstybinis požiūris ir numatytas finansavimas, remiantis kultūros sektoriaus institucijų duomenų valdymo pertvarką *susietųjų atvirųjų duomenų* kryptimi. Fondoieškos srityje atminties institucijos sukaupė pakankamai patirties, kai 2004–2009 m. joms pačioms teko ieškoti papildomo finansavimo kultūros paveldo skaitmeninio projektams įgyvendinti. Bendromis atminties institucijų pastangomis buvo parengti strateginių dokumentų *Lietuvos kultūros paveldo skaitmeninio koncepcijos* ir *Lietuvos kultūros paveldo skaitmeninio, skaitmeninio turinio ir prieigos strategijos* projektai. Lietuvos Respublikos kultūros ministerijai aprobavus šiuos projektus ir Vyriausybei juos patvirtinus, atsirado galimybė kreiptis finansinės paramos į ES struk-

<sup>23</sup> The European Commission Policy on Open Data. Prieiga per internetą: <https://digital-strategy.ec.europa.eu/en/policies/open-data-0>; Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Open Data An Engine for Innovation, Growth and Transparent Governance.

<sup>24</sup> Lietuvos viešojo sektoriaus atvirųjų duomenų architektūros modelis. Vilnius: Informacinės visuomenės plėtros komitetas prie Susisiekimo ministerijos, 2014, p. 9. Prieiga per internetą: [https://ivpk.lrv.lt/uploads/ivpk/documents/files/Veikla/Veiklos\\_sritys/Informacijos\\_pakartotinis\\_naudojimas/Atvir%C5%B3%20duomen%C5%B3%20teikimo%20modelis.pdf](https://ivpk.lrv.lt/uploads/ivpk/documents/files/Veikla/Veiklos_sritys/Informacijos_pakartotinis_naudojimas/Atvir%C5%B3%20duomen%C5%B3%20teikimo%20modelis.pdf). Žiūrėta 2021-04-03.

<sup>25</sup> Berners-Lee, T. Linked Data – Design Issues. Retrieved July 23, 2006. Prieiga per internetą: <http://www.w3.org/DesignIssues/LinkedData.html>. Žiūrėta 2021-04-03.

<sup>26</sup> Lietuvos viešojo sektoriaus atvirųjų duomenų architektūros modelis, p. 9.

tūrinius fondus 2004–2007 m. ir 2007–2013 m. programavimo laikotarpiais ir įgyvendinti reikšmingus projektus, užtikrinančius ilgalaikį skaitmeninto turinio išsaugojimą bei aktualinančius mūsų šalies kultūros paveldą. Nurodytų strateginių dokumentų svarba akcentuojama ir vertinant ES struktūrinės paramos poveikį kultūrai (2007–2013 m. ES struktūrinės paramos <...>, 2016). Tenka konstatuoti, kad šiuo metu kultūros sektoriuje nėra strateginio dokumento, aktualizuojančio atminties institucijų duomenų valdymo pertvarką *susietųjų atvirųjų duomenų* kryptimi. Tad galima turėti vilties, kad straipsnyje analizuojamos *susietųjų atvirųjų duomenų* realizavimo bibliotekose galimybės paskatins platesnę diskusiją ne tik tarp šalies bibliotekų, bet ir padės atkreipti politikų, atsakingų už kultūros politikos formavimą, dėmesį į tai, įgyvendinant *Atvirų duomenų iniciatyvos įgyvendinimo modelyje* (2018 m.) numatytus reikalavimus ypač aktualu strateginiame bibliotekų plėtros dokumente strategiškai pagrįsti mūsų šalies bibliotekų duomenų valdymo *susietųjų atvirųjų duomenų* kryptimi pertvarką, akcentuojant bibliotekų sektoriaus duomenų integralumą, ir numatyti finansinę paramą šiai pertvarkai įgyvendinti. Kita galima alternatyva galėtų būti skubiai parengta kompleksinė programa su finansine parama numatytiems pokyčiams kultūros sektoriuje įgyvendinti. Finansavimo perspektyva šiems pokyčiams atsiveria naujai patvirtintose ES programose. Pavyzdžiui, ES. *Skaitmeninės Europos programoje 2021–2027 m.* tarp prioritetinių finansavimo sričių paminėtas platus aukšto lygio skaitmeninių įgūdžių ir skaitmeninių technologijų naudojimas visose ekonomikos srityse ir visuomeniniame gyvenime (ES. *Skaitmeninės Europos programa*, 2019).

## 1.2. *Susietųjų atvirųjų duomenų metodologijos paradigma*

Kol kas nėra vienos *susietųjų atvirųjų duomenų* sampratos, tačiau yra net keletas jos apibūdinimų, grindžiamų skirtingais metodologiniais požiūriais, kurie apibendrinti Marianos Ziku analitinėje apžvalgoje<sup>27</sup>. T. Berners-Lee akcentuoja, kad semantinis saitynas yra ne tik jame paskelbti duomenys, tai yra ir nuorodų sukūrimas<sup>28</sup>. Tokiai nuomonei pritaria ir kiti autoriai. Christianas Bizeris, Tomas Heathas ir T. Barners-Lee *susietuosius duomenis* apibūdina kaip gerosios praktikos rinkinį skelbiant ir sujungiant struktūrinius duomenis internete<sup>29</sup>. Remiantis *susietųjų duomenų* metodologijos autoriais, galima daryti išvadą, kad *susietųjų duomenų* samprata apima ir metodologines ir technologines priemones. Naudojant *Išteklių aprašo modelį* (toliau – RDF)<sup>30</sup> bei *saityno ontologijų kalbas* (toliau – OWL)<sup>31</sup> užtikrinama duomenų iš skirtingų šaltinių bei nesusijusių sistemų heterogeninių duomenų sietis semantiniame saityne. Vienas didžiausių iššūkių kuriant semantinį saityną – įveikti įvairių bendruomenių skelbiamų metaduomenų heterogeniškumą. Siekdamas įgyvendinti šį iššūkį 2006 m. T. Berners-Lee nustatė *susietiesiems duomenims* keliamus reikalavimus ir juos publikavo *Susietųjų duomenų principuose*<sup>32</sup> (toliau – Principai), kurie neabejotinai galėtų būti naudingi kaip gerosios praktikos rekomendacijos projektuojant naujas arba renovuojant senąsias sistemas susietųjų duomenų generavimo kryptimi. Principus sudaro pagrindinės keturios taisyklės:

- sąvokoms apibūdinti naudokite universalųjį išteklių identifikatorių (URI);
- siekdami sukurti galimybę surasti šias sąvokas naudokite HTTP URI;
- ieškantiems URI, suteikite naudingos informacijos pasinaudodami RDF, SPARQL standartais;
- pateikite nuorodų į kitus URI, kurios suteiks vartotojui galimybę atrasti daugiau jį dominančių dalykų.

<sup>27</sup> Žr. Ziku, M. Digital Cultural Heritage and Linked Data: Semantically-Informed Conceptualisations and Open Practices with Focus on Intangible Cultural Heritage. *Liber Quarterly*, 2020, Vol. 30, No. 1, p. 1–16.

<sup>28</sup> Berners-Lee, T. *The Semantic Web Road Map*, p. 1.

<sup>29</sup> Žr. Bizer, Ch.; Heath, T.; Berners-Lee, T. Linked Data – The Story So Far. *International Journal on Semantic Web and Information Systems*, 2009, Vol. 5, No. 3, p. 1–22.

<sup>30</sup> *Resource Description Framework*: Prieiga per internetą: <https://www.w3.org/TR/owl-features/>.

<sup>31</sup> *Web Ontology Language*: Prieiga per internetą: [https://en.wikipedia.org/wiki/Web\\_Ontology\\_Language#OWL\\_sublanguages](https://en.wikipedia.org/wiki/Web_Ontology_Language#OWL_sublanguages).

<sup>32</sup> Berners-Lee, T. Linked Data – Design Issues.



Šios taisyklės ir semantinio saityno konsorciumo (Web3C) dokumentacija<sup>33</sup> suformavo nūdienos semantinio saityno funkcionavimo metodologiją, reikalaujančią kiekvienai sąvokai naudoti *universalųjį išteklių identifikatorių* (toliau – URI), kuris turi būti susijęs su naudinga semantine informacija. Pagrindinis struktūrinis abstrakčios sintaksės komponentas yra triadų aibės, kurias sudaro *universaliojo išteklių identifikatoriumi* besiremiantis *veiksny*, *papildinys* ir *tarinys*. *Veiksnys* identifikuoja tai, apie ką yra teiginys, *tarinys* – tam tikrą apibūdinamo dalyko aspektą, o *papildinys* – šio aspekto reikšmę. Triados pagrindas yra *tarinys*, kuris yra pateikiamas kaip RDF savybė, o triados *veiksnys* ir *papildinys* pateikiami kaip RDF klasių nariai. Klasės apibūdina dalykus, o savybės – dalykų santykius; klasės ir savybės yra pamatiniai RDF elementų tipai. Dalykas, apibūdinamas kaip klasė, gali būti bet koks išteklius arba esybė, apie kurią norima ką nors pasakyti; toks dalykas naudojamas kaip triados veiksnys. RDF yra pamatas kuriant pamatines ontologijas, leidžiančias integruoti konkrečios srities ar interesų lauko sampratas. Kaip tokį pavyzdį galima būtų pateikti CIDOC CRM, kuris buvo sukurtas muziejų sektoriui ir pamažu buvo pritaikytas informacijai koduoti ir skleisti internete daugelyje visuomeninio gyvenimo sričių. 1 lentelėje pateiktas pavyzdys iliustruoja *veiksni – tarinio – papildinio* triadą, remiantis CIDOC CRM grindžiama RDF metodologija. Lentelėje *veiksnys* yra *veikėjas* (angl. *Actor*), kuris yra koduojamas E39 CIDOC CRM klase. *Tarinys* lentelėje yra CIDOC CRM savybė, kuri koduojama <P1 *Yra vadinamas (identifikuojamas)*> (angl. *Identified by (identifies)*); ir *papildinys* atitinkamai lentelėje yra CIDOC CRM klasė E41 *Veikėjo vardas (pavadinimas), identifikuojantis objektą* (angl. *Appellation*).

1 lentelė

CIDOC CRM klasė	Savybė	Atributas
E39 Actor / Veikėjas	P1 Identified by (identifies) / Yra vadinamas (identifikuojamas)	E41 Appellation / Veikėjo vardas (pavadinimas), identifikuojantis objektą

Ch. Bizeris, T. Heathas ir T. Barners-Lee teigia, kad triados *veiksnys* ir *papildinys* gali priklausyti skirtingiems duomenų rinkiniams, kurie naudodami RDF susieja dokumentus. Tai užtikrina heterogeninių duomenų pateikimą RDF formatu kartu su jo semantine išplėta – RDF(S) į semantinį saityną, o tai ypač vertinga perteikiant, jungiant ir interpretuojant valdymo sistemas ir portalus<sup>34</sup>.

Kita svarbi *susietųjų atvirųjų duomenų* principų taisyklė reikalauja užtikrinti sampratų suradimo galimybę. Siekiant šio tikslo kiekvienam *universaliajam išteklių identifikatoriui* (toliau – URI) turėtų būti naudojamas HTTP protokolas. Pats semantinio saityno išradėjas T. Berners-Lee pataria, jog paprasčiausias būdas realizuoti šį sprendimą yra tame pačiame URI faile pateikti nuorodą į kitą URI<sup>35</sup>. Ši priemonė dar yra vadinama nuorodų susiejimo (*de-reference*) technologija, užtikrinančia papildomų duomenų suteikimą į pateiktą užklausą. Tokia nuorodų susiejimo technologija yra vertinama technologijos gebėjimu užtikrinti galimybę žmogui prieinama forma skaityti saityne pateiktą tinklalapį HTML formatu arba mašina – apdorojamus duomenis RDF formatu.

### 1.3. Susietųjų atvirųjų duomenų technologijų paradigma

Kaip aptarta, pagrindinė semantinio saityno kalba yra RDF modelis, kuris yra naudojamas ištekliams aprašyti, suteikiantis priemones, reikalingas ištekliams apibrėžti. Be RDF, semantinis saitynas dar disponuoja saityno ontologijos kalba (OWL), (OWL2, 2012), kuri yra išraiškingesnė ir formalesnė nei RDF ir RDF-S, galinti labai glaudžiai susieti duomenis, kuriai būdinga semantinių savybių gausa ir kuri gali būti

<sup>33</sup> World Wide Web. Prieiga per internetą: <https://www.w3.org/DesignIssues/LinkedData.html>.

<sup>34</sup> Bizer, Ch.; Heath, T.; Berners-Lee, T. Linked Data – The Story So Far, p. 1–22.

<sup>35</sup> Berners-Lee, T. Linked Data – Design Issues.

vartojama atliekant automatines loginės operacijas, pavyzdžiui, galimybę logiškai apdoroti duomenis ir dirbti su heterogeniniais duomenų šaltiniais. Todėl dalykinės rubrikos, klasifikacijų rubrikos ir tezaurai, pateikti RDF naudojant SKOS, suteikia galimybę išreikšti terminų santykius, taip pat siauresnius ir tikslesnius terminų atitikmenis. Naudojant SKOS kiekvienai tam tikro termino sąvokai suteikiamas URI, atliekantis funkcinių termino suderinamumą daugiakalbėje informacinėje erdvėje. Tai suteikia daugiau loginio triadų išvedimo galimybių. Naudojant RDF ir OWL sukuriama ir kita galimybė – tai semantiškai raiškių duomenų modelių, pavyzdžiui, pamatinės ontologijos (*Core ontology*), padedančios integruoti informaciją iš įvairių šaltinių, sukūrimas. Pavyzdžiui, Lietuvoje muziejų ir bibliotekų sektoriuje 2010 m. plėtojant VEPIS buvo pasirinktas CIDOC CRM, nes šis standartas pagrįstas objektiniu požiūriu į tikrovę, o ne į konvencionalias sąvokas, priklausančias nuo skirtingos kalbinės, kultūrinės ar istorinės tradicijos. Ši ontologija buvo pritaikyta VEPIS pamatinei ontologijai – *Bendram asmenvardžių, vietovardžių ir istorinės chronologijos tezaurui*<sup>36</sup> sukurti. Semantinio saityno technologijos grindžiamos OWL, SKOS, SPARQL, nurodančiomis duomenų pateikimo kalbą ir duomenų pateikimo taisykles. Apibendrinant galima daryti išvadą, kad *susietųjų atvirųjų duomenų* paradigmai susiformuoti įtakos turėjo EK *atvirųjų duomenų* strateginiai dokumentai, kuriais siekiama paspartinti naujų paslaugų plėtrą, paremtą inovatyviais informacijos derinimo metodais, skatinančiais ekonomikos augimą ir didinančiais visuomenės aktyvumą; sukurta *Išteklų aprašo modelio* – RDF ir *Susietųjų duomenų principų* metodologija, leidžianti aprašyti duomenis ir informaciją, teikti ir sieti semantiniame saityne; bei sukurta semantinių saityno OWL, SKOS ar SPARQL grindžiamų technologijų, kuriomis užtikrinama duomenų iš skirtingų šaltinių ir nesusijusių sistemų heterogeninių duomenų sietis semantiniame saityne. Tuo pagrindu sąlygiškai galima išskirti tris paradigmas: *atvirųjų duomenų ir informacijos valdymo politikos, susietųjų atvirųjų duomenų metodologijos ir susietųjų atvirųjų duomenų technologijos*.

Tyrimas taip pat leidžia daryti išvadą, jog *susietųjų atvirųjų duomenų* pagrindu organizuojama paieška leidžia pagerinti informacijos tikslumą, atsižvelgiant į informacijos vartotojo ketinimus ir konteksto terminų reikšmę, o ne į atskirų žodžių prasmę užklausoje. Pagrindinis tokios paieškos siekis yra gauti klausimą atitinkančius atsakymus.

## 2. Standartizuotų bibliotekų metaduomenų modelių ir formatų plėtra *susietųjų atvirųjų duomenų* kryptimi

Nuo semantinio saityno sukūrimo bibliotekų sektoriuje buvo pasiūlyta nemažai iniciatyvų, siekiančių bibliotekose naudojamus metaduomenų modelius suderinti su semantiniu saitynu. Ir mūsų šalies bibliotekose siekiant išteklų katalogavimo pertvarkos tikslinga geriau susipažinti su įvairių IFLA padalinių iniciatyvomis modernizuojant jau senstelėjusius bibliotekų metaduomenų modelius. Moksliskai argumentuotą informaciją apie bibliotekų metaduomenų modelių ir struktūrų iniciatyvas, siekiančias pritaikyti šiuos modelius semantiniame saitynui, daugelį metų skelbia Mirna Willer ir Gordonas Dunsire<sup>37</sup>. Nuolatinė šių tyrėjų veikla įvairiuose IFLA komitetuose ir kituose tarptautiniuose forumuose plėtojant

<sup>36</sup> Laužikas, R; Vaitkevičius, V. *Bendras asmenvardžių, vietovardžių ir istorinės chronologijos tezauras. Virtualios elektroninio paveldo informacinės sistemos (VEPIS) specifikacija*. 2011.

<sup>37</sup> Žr. Willer, M. Name Authority Control: From Conceptual Model to the Semantic Web and Linked Data. In: *Summer School in the Study of Historical Manuscripts: Proceedings* / Ed. by M. Willer and M. Tomić. Zadar: Sveučilište u Zadru, 2013, p. 243–264; Dunsire, G.; Willer, M. Standard Library Metadata Models and Structures for the Semantic Web. *Library Hi Tech News*, 2011, Vol. 28, No. 3, p. 1–12; Dunsire, G.; Willer, M. UNIMARC and Linked Data. *International Federation of Library Associations and Institutions*, 2011, Vol. 37, No. 4, p. 314–326; Willer, M.; Dunsire, G.; Bosačić, B. ISBD and the Semantic Web. *JLIS.it: Italian Journal of Library and Information Science*, 2010, Vol. 1, No. 2, p. 213–236; Willer, M.; Dunsire, G. ISBD, the UNIMARC Bibliographic Format, and RDA: Interoperability Issues in Namespaces and the Linked Data Environment. *Cataloging & Classification Quarterly*, 2014, Vol. 52, No. 8, p. 888–913; Dunsire, G.; Willer, M. Authority versus Authenticity: The Shift from Labels to Identifiers. *Authority, Provenance, Authenticity, Evidence: Selected Papers from the Conference and School Authority, Provenance, Authenticity, Evidence, Zadar, October 2016*, p. 87–115.

UNIMARC formatą bibliografiniams ir autoritetiniams įrašams, ISBD – bibliografinių įrašų turinio standartą, *Funkcinių reikalavimų bibliografiniams įrašams* (FRBR) ir *Funkcinių reikalavimų autoritetiniams įrašams* (FRAR) modelius, RDA sudarė jiems galimybę ne tik užfiksuoti IFLA metaduomenų modelių plėtros raidą, atskleisti šių modelių sąsajas, pateikti išsamią šios IFLA veiklos analizės panoramą, bet ir atsakyti į esminių bibliotekų bendruomenei rūpimą klausimą: kokiais sprendimais reikėtų vadovautis siekiant dabartinę katalogavimo praktiką pertvarkyti *susietųjų atvirųjų duomenų* pagrindu. 2011 m. publikuotame straipsnyje autoriai pateikė net keletą šios pertvarkos būdų. Vienas jų – bibliotekų standartus naudoti kaip pagrindą naujų *triadų susietiesiems atviriesiems duomenims* rengti sudarant atitinkamus elementus RDF forma; kitas galimas sprendimas – atitikčių (angl. *mapping*) esamiems elementams rengimas. Rengiant tokias atitiktis RDF savybės galėtų būti pasirenkamos iš įvairių standartų bei ontologijų ir derinamos vienoje programinėje įrangoje, atsižvelgiant į jos funkcinius reikalavimus naudojant saityno ontologiją. Trečias sprendimas, kurį ypač rekomenduoja autoriai, – naudoti su semantinio saityno reikalavimais suderintus metaduomenų modelius<sup>38</sup>. Tokių modelių plėtros ištakos, kaip nurodo M. Willer ir G. Dunsire, – 2006 metai, kai IFLA ISBD peržiūros sekcija ėmėsi kurti XML schemą, skirtą ISBD<sup>39</sup>. ISBD peržiūros darbo grupė vienu svarbiausių uždavinių laikė identifikuoti ISBD ryšį su FRBR modeliu, siekdama priderinti ISBD terminiją prie FRBR esybių – *kūrinio, išraiškos, apraiškos ir vieneto aprašymo*. Tačiau po ilgų svarstymų šios idėjos 2003 m. buvo atsisakyta. Ir tik po ilgos pertraukos 2021 m. šis projektas buvo atgaivintas ir plėtojamas. Rašant šį straipsnį ISBD yra transformuojamas derinant jį su nauja bibliografinė paradigma, išreikšta IFLA LRM – *kūriniu, išraiška, apraiška* ir *vienetu* bei orientuojantis į *apraiškos* lygmenį. Užbaigus šį projektą, ISBD turėtų būti pritaikytas ir nepublikuotiems ištekliams, išplėtotą *apraiškos* sąvoka ją derinant su LRM bei suformuotas *ISBD apraiškos atributų elementų ir ryšio rinkinys*. Taip bus užtikrinta ISBD transformacija ir sąveikumas su LRM *apraiškos* lygiu, papildantis ISBD nepublikuotų išteklių aprašymo reikalavimais bei atitinkantis *susietųjų atvirųjų duomenų* lygį. Toje pačioje ISBD peržiūros grupės ataskaitoje<sup>40</sup> akcentuojami ir kiti IFLA projektai, kurie neabejotinai yra svarbūs *susietųjų atvirųjų duomenų* kontekste. Tai ISBD žodynai, kurie dabar yra IFLA vardų erdvių svetainėje, leidžiantys juos naudoti bibliotekose, archyvuose ir muziejuose pritaikant juos semantinio saityno ir *susietųjų atvirųjų duomenų* programose. IFLA standartų plėtros nuoseklumą skatina metaduomenų turinio modelio – ISBD ir formos modelio – UNIMARC sietis. Dėl šios priežasties šie du metaduomenų modeliai buvo plėtojami lygiagrečiai: bibliografinių įrašų – UNIMARC/B 2008<sup>41</sup> ir autoritetinių įrašų – 2009 m. UNIMARC/A<sup>42</sup> formatai. Be minėtų modelių, IFLA darbo grupės lygiagrečiai vykdė tyrimus plėtodamos FRAD<sup>43</sup> ir FRASAD<sup>44</sup>, kurie sukurti kaip į esybes ir jų santykius orientuoti modeliai ir kurie vėliau visi kartu sudarė koncepcinių modelių šeimą. Svarbu atkreipti dėmesį į tai, kad trečiame UNIMARC/A leidime, publikuotame 2009 m., jau buvo integruotos ontologinės FRAD modelio ypatybės, reglamentuojančios esybių, atributų ir santykių struktūrą bei priartinančios bibliotekų metaduomenis prie semantinio saityno ir jo paslaugų. Stebint tolesnę koncepcinių modelių šeimos plėtrą svarbu pažymėti, jog šie modeliai buvo transformuojami į objektą orientuotą FRBRoo modelį, kuris daro įtaką ir kitų modelių plėtrai. Todėl neatsitiktinai, daugelio tyrėjų nuomone, tobulinant bibliotekų informacines sistemas semantinio saityno kryptimi, tikslinga naudoti nauja bibliografinė paradigma paremtų modelių

<sup>38</sup> Dunsire, G.; Willer, M. UNIMARC and Linked Data, p. 315.

<sup>39</sup> *Ibidem*, p. 315.

<sup>40</sup> Update on the Work of the ISBD Review Group. *Committee on Standards ISBD Review Group*. Prieiga per internetą: <https://www.ifla.org/isbd-rg>.

<sup>41</sup> *UNIMARC Manual: Bibliographic Format. 3rd edition* / Edited by A. Hopkinson. München: K. G. Saur, 2008.

<sup>42</sup> *UNIMARC Manual: Authorities Format. 3rd edition* / Edited by M. Willer. München: K. G. Saur, 2009.

<sup>43</sup> *Functional Requirements for Authority Data (2009)*. By IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). Prieiga per internetą: <https://www.ifla.org/publications/functional-requirements-for-authority-data>.

<sup>44</sup> *Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model. (2010)*. *IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR)*. Editors: Marcia Lei Zeng, Maja Žumer, Athena Salaba. Prieiga per internetą: <https://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>. Žiūrėta 2021-03-24.

struktūras, galinčias iš esmės pakeisti nuo laikmečio reikalavimų atsilikusią katalogavimo praktiką bibliotekose. Pat Riva ir Maja Žumer nurodo du iš jų: tai į esybes ir jų santykius orientuotas IFLA pavyzdinis bibliotekų *Koncepcinis bibliografinės informacijos modelis* (toliau – LRM) (*Library Reference Model*) ir į objektą orientuotas modelis – *Definition of FRBRoo: A Conceptual Model for Bibliographic Information in Object-Oriented Formalism* (toliau – FRBRoo)<sup>45</sup>. Pirmasis jų – LRM, kuris oficialiai buvo patvirtintas 2017 m., o 2020 m. išverstas į lietuvių kalbą, grindžiamas esybių ir jų santykių modeliavimo principu, apimantis visus bibliografinio universumo aspektus. Pasak autorių, šis modelis sujungia tris modelius (FRBR, FRAD ir FRSAD) į vieną pamatinį modelį, integruojantį visus bibliografinės ir autoritetinės informacijos aspektus. Tokia modelių integracija svariai praplėtė LRM taikymo erdvę, kuri tapo kur kas platesnė nei jo pirmtakų – FRBR, FRAD ir FRSAD modelių. Kitas modelis – FRBRoo yra į objektą orientuota FRBR modelio versija, deranti su CIDOC koncepciniu pamatiniu modeliu (CIDOC CRM), kurį suformavo 2003 m. sutelkta Tarptautinė FRBR ir CIDOC CRM darninimo darbo grupė, siekianti išreikšti IFLA FRBR modelį, grindžiamą CIDOC CRM metodologija ir kodavimo nuostatomis. Analizuojant FRBR darninimą su CIDOC CRM peršasi nuomonė, kad FRBRoo plėtros posūkį CIDOC CRM kryptimi nulėmė pastarojo modelio ypatybės, leidžiančios aukštesniu lygiu struktūrintai, sistemingai ir semantiškai korektiškai perteikti realaus pasaulio vaizdą konceptualizuojant pagal standartizuotą būdingą viršklasių struktūrą, leidžiančią teikti savybes – tolygias santykiams ir atributams, dažniausiai taikomiems poklasių lygmeniu. Atsižvelgiant į tai, kad pagrindinis FRBRoo darninimo su CIDOC CRM siekis yra sukurti naująjį į objektą orientuotą modelį, buvo pakeistas ir jo pavadinimas į LRMoo. Šių modelių kūrėjų nuomone, ir LRM, ir LRMoo turi privalumų ir konkrečią paskirtį, tad dėl modelio pasirinkimo nusprendžia kiekviena šalis ar konkreti institucija. Kaip teigia P. Riva ir M. Žumer (Riva, 2018), LRMoo geriausiai tiktų bibliotekų ir muziejų informacijai integruoti ir galėtų būti metodologinis pagrindas ren-giant *susietuosius atviruosius duomenis* bibliotekose<sup>46</sup>.

Atsižvelgiant į tai, kas pasakyta, galima daryti išvadą, jog, pereinant prie bibliotekų metaduomenų modelių ir duomenų struktūros pertvarkos *susietųjų atvirųjų duomenų* kryptimi, reikia žengti tris pagrindinius žingsnius: teaurą būtina keisti formalia ontologija, transformuoti bibliotekose įsitvirtinusią katalogavimo praktiką taikant bibliotekų duomenų modelius, grindžiamus nauja bibliografinė paradigma – *kūrinio, išraiškos, apraiškos ir vieneto* esybių aprašymu ir disponuojančia esybių, atributų ir santykių struktūra bei aprūpinti informacinę sistemą specialia semantinės paieškos programine įranga.

### 3. Susietųjų atvirųjų duomenų geroji praktika pasaulio šalių bibliotekose

Mokslinėje literatūroje, be teorinių *susietųjų atvirųjų duomenų* klausimų, jau publikuojama ir gerosios praktikos pavyzdžių perkeliant duomenis iš bibliotekinių formatų į RDF formatą. Apie bibliotekų standartų naudojimą kaip pagrindą naujiems triadų susietiesiems duomenims rengti, RDF atitikčių (angl. *mapping*) esamiems elementams rengimą arba su semantinio saityno reikalavimais suderintų metaduomenų modelių BIBFRAME, IFLA LRM naudojimą įvairių šalių bibliotekų specialistai dalijasi patirtimi kasmetinėse IFLA konferencijose, tarptautinėje skaitmeninių bibliotekų konferencijoje TPDL (angl. *Theory and Practice of Digital Libraries*)<sup>47</sup> ir kituose tarptautiniuose forumuose, kurių apžvalga pateikiama šiame poskyryje.

2012 m. Kongreso bibliotekoje (toliau – KB) buvo paskelbta pirmoji BIBFRAME modelio (angl. *Bibliographic Framework*)<sup>48</sup> specifikacijos versija, keičianti požiūrį į bibliografinių duomenų kūrimą ir

<sup>45</sup> Riva, P.; Žumer, M. *FRBRoo, the IFLA Library Reference Model, and now LRMoo: A Circle of Development*, 2018. Prieiga per internetą: <http://library.ifla.org/2130/1/074-riva-en.pdf>. Žiūrėta 2021-03-24.

<sup>46</sup> *Ibidem*.

<sup>47</sup> *Theory and Practice of Digital Libraries*. Prieiga per internetą: <http://www.tpd.eu/>. Žiūrėta 2021-04-02.

<sup>48</sup> *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services*. Washington: Library of Congress, DC November 21, 2012. Prieiga per internetą: <https://www.loc.gov/bibframe/pdf/marclcd-report-11-21-2012.pdf>.



sklaidą bei deranti su semantiniu saitynu ir susietųjų duomenų technologijomis. Antroji šio modelio – BIBFRAME 2.0 versija<sup>49</sup>, kurios pagrindinis tikslas – konvertuoti bibliografinius įrašus MARC21 formatu į *susietuosius duomenis* vadovaujantis ir adaptuojant RDF modeliavimo praktiką, pasirodė 2016 m. Vadovaujantis šiuo modeliu KB informacijos ištekliams buvo nustatyti unikalieji identifikatoriai, įvardytos jų savybės ir nustatyti tarpusavio ryšiai. Netrukus BIBFRAME modelis imtas taikyti Vokietijos ir Švedijos nacionalinėse bibliotekose, Vengrijos nacionalinio muziejaus bibliotekoje, Aleksandrijos bibliotekoje transformuojant katalogų, nacionalinių bibliografijų ir autoritetinius įrašus į atvirusius duomenis.

Bibliotekinės programinės įrangos tiekėjo *Ex Libris* programinėje įrangoje (toliau – PĮ) *Alma*, kurią naudoja daugiau nei 1,9 tūkst.<sup>50</sup> viso pasaulio bibliotekų, taip pat yra pritaikytas BIBFRAME modelis. Šio modelio pagrindu gali būti saugoma, redaguojama ir skelbiama bibliografinė informacija. Naudojant šį modelį *Alma* sistemoje esantys bibliografiniai įrašai gali būti peržiūrėti ar eksportuojami BIBFRAME schema, taip pat prieinami per unikalios URI (*Ex Libris*. BIBFRAME).

JAV ne pelno organizacija *Online Computer Library Center* (OCLC, vienijanti daugiau nei 10 tūkst.<sup>51</sup> viso pasaulio bibliotekų) taip pat eksperimentuoja taikydama BIBFRAME *WorldCat* katalogo įrašams transformuoti. Kaip nurodo Nathanas B. Putnamas, atliekant šį eksperimentą 11 mln. *WorldCat* MARCXML įrašų buvo konvertuota į BIBFRAME naudojant automatizuotus konverterius<sup>52</sup>. Kita Kongreso bibliotekos *susietųjų atvirųjų duomenų* iniciatyva – autoritetinių įrašų ir žodynų platforma – *Id.loc.gov*. Nuo 2019 m. panaudodama *Wikidata* identifikatorius KB per metus sukūrė per 1,2 mln. ryšių tarp bibliotekos ir *Wikidata* įrašų. Šiuo projektu siekiama paskatinti naudoti *Wikidata* identifikatorius ir kartu padidinti *id.loc.gov* naudojimą ir matomumą, kadangi į *Wikidata* taip pat perduoti bibliotekos identifikatoriai. Mattas Milleris atkreipia dėmesį į tai, kad įgyvendinant tokio pobūdžio projektus susiduriama su iššūkiu užtikrinti proceso tęstinumą. Kadangi *Wikidata* duomenys nuolat keičiasi, įrašų lyginimo procesas abiejose sistemose vyksta nuolat<sup>53</sup>.

Prie modelių, suderintų su semantinio saityno reikalavimais, taikymo susietųjų atvirųjų duomenų rengimo atvejų galima būtų priskirti ir Lietuvos pavyzdį. 2021 m. modernizuojamoje *Lietuvos integruojamoje bibliotekų informacinėje sistemoje* (toliau – LIBIS) pradedamas taikyti IFLA LRM modelis, kurio pagrindu naujai modeliuojamas išteklių katalogavimas, identifikuojant ir pateikiant *kūrinio, išraiškos ir apraiškos* esybių aprašymo taisyklės<sup>54</sup>. *Susietųjų atvirųjų duomenų* požiūriu vertinga yra tai, kad šios taisyklės apima UNIMARC/A ir UNIMARC/B naujų laukų papildymus, ISBD pakeitimus ir tai, kad 2009 m. trečio UNIMARC/A leidimo papildymuose yra integruotos ontologinės FRAD modelio ypatybės, reglamentuojančios esybių, atributų ir santykių struktūrą<sup>55</sup>. Tai priartins mūsų šalies bibliotekų LIBIS programine įranga kuriamus metaduomenis prie semantinio saityno ir jo paslaugų.

<sup>49</sup> Overview of the BIBFRAME 2.0 Model. *Library of Congress*, 2016. Prieiga per internetą: <https://www.loc.gov/bibframe/docs/bib-frame2-model.html>. Žiūrėta 2021-03-28.

<sup>50</sup> *Library Technology Guides*. Prieiga per internetą: <https://librarytechnology.org/product/alma/>. Žiūrėta 2021-04-02.

<sup>51</sup> *WorldCat*. Prieiga per internetą: <https://www.worldcat.org/>. Žiūrėta 2021-04-02.

<sup>52</sup> Putnam, N. B. *OCLC Research BIBFRAME 2.0 Analysis. LC BIBFRAME Update, ALA Annual 2018*. Prieiga per internetą: <https://www.loc.gov/bibframe/news/pdf/oclc-converter-alaal2018.pdf>. Žiūrėta 2021-03-28.

<sup>53</sup> Miller, M. *Library of Congress. Id.loc.gov and Wikidata, One Year Later. SWIB20: Semantic web in Libraries, 2020*. Prieiga per internetą: [https://youtu.be/OvcBBRZ\\_sUM](https://youtu.be/OvcBBRZ_sUM). Žiūrėta 2021-03-13.

<sup>54</sup> Žr. *LIBIS išteklių įrašų sudarymo vadovas: integruojamasis išteklius* / Parengė: T. Auškalnis, R. Bagušytė, L. Buckienė, V. Čeglienė, D. Gedminienė, G. Kubilienė, I. Kvietkuvienė, A. Markevičiūtė, I. Miliauskienė, D. Norkūnienė, D. Pečiulienė, G. Rutkauskienė, D. Sipavičiūtė, R. Skačkauskienė, R. Švabauskaitė, A. Ušpalienė, dr. V. Vaitkevičiūtė-Verbickienė; atsakomoji redaktorė dr. R. Januševičienė. Vilnius, Lietuvos nacionalinė Martyno Mažvydo biblioteka, 2021. Prieiga per internetą: <https://journals.lnb.lt/metodologijos-paskyra/article/view/738/733>.

<sup>55</sup> Žr. *Functional Requirements for Bibliographic Records: Final Report. IFLA Study Group on the Functional Requirements for Bibliographic Records*. München: Saur, 1998. Amended 2009. Prieiga per internetą: <https://www.google.com/search?q=%3A+http%3A%2F%2Fwww.ifla.+org%2Fen%2Fpublications%2Ffunctional-requirements-for-bibliographic-records.&oq=%3A+http%3A%2F%2Fwww.ifla.+org%2Fen%2Fpublications%2Ffunctional-requirements-for-bibliographic-records.&aqs=chrome..69i57.1570j0j7&sourceid=chrome&ie=UTF-8>. Žiūrėta 2021-04-03.

Geriau susipažinti su *susietųjų atvirųjų duomenų* iniciatyvomis, įgyvendinamomis įvairių šalių bibliotekose, padeda 2019 m. Hamburge įvykusi konferencija *Semantinis saitynas bibliotekose* (SWIB19)<sup>56</sup> ir 2020 m. nuotoliniu būdu įvykusi SWIB20 konferencija<sup>57</sup>.

2019 m. vykusioje konferencijoje René Voorburg pristatė Nyderlandų nacionalinės bibliotekos 2018 m. paskelbtą *susietųjų atvirųjų duomenų* publikavimo platformą *data.bibliotheken.nl*<sup>58</sup>. Pasak pranešėjos, siekiant pagrindinio tikslo – atverti bibliotekos duomenis internete ir darant sietis, reikėjo ne tik pradėti naudoti RDF, bet parengti duomenis eksportuoti, modeliuoti, konvertuoti ir transformuoti. Ir vis dėlto svarbiausi žingsniai, kuriant *susietuosius atviruosius duomenis*, pranešėjos nuomone, yra semantinis modeliavimas, nes nuo jo priklauso šios platformos panaudojimo vertė. Pasak pranešėjos, *data.bibliotheken.nl* platformos semantinį modelį ateityje siekiama išplėtoti publikuojamiems duomenims pritaikant unifikotą metaduomenų schemą, paremtą IFLA-LRM ontologija ir PREMIS schema. Platforma grindžiama SKOS, RDF, URI. Duomenys yra publikuojami Nyderlandų nacionalinės bibliotekos specialistų sukurta schema naudojant RDF<sup>59</sup>.

Kitas *susietųjų atvirųjų duomenų* gerosios praktikos pavyzdys – *Datos.bne.es*<sup>60</sup>. Tai Ispanijos nacionalinės bibliotekos katalogas. Jis buvo sukurtas transformavus katalogo įrašus MARC21 formatu į RDF ir duomenų modelį, grindžiamą FRBR. *Datos.bne.es* eksperimentas leidžia testuoti neįprastus metaduomenų kūrimo procesus. Pasak Ricardo Santoso, vienas didžiausių šio projekto pasiekimų – *Wikidata* duomenų įkėlimas į 80 000 bibliotekos katalogo įrašų apie autorius, įskaitant lytį, gimimo vietą, išsimokslinimą, profesiją, kalbą, narys ir pan. Bibliotekos įrašai su *Wikidata* duomenimis buvo siejami per *Wikidata* ir OCLC VIAF (*Virtual International Authority File*)<sup>61</sup> autoritetinių įrašų identifikatorius<sup>62</sup>.

2011 m. Prancūzijos nacionalinė biblioteka paskelbė *susietųjų atvirųjų duomenų* paslaugą – *data.bnf.fr*<sup>63</sup>, padedančią vartotojams rasti reikalingos informacijos bibliotekos kolekcijose. Tačiau tuo nebuvo apsiribota. Raphaëlle Lapôtre, Marie Destandau ir Emmanuel Pietriga akcentavo, jog 2017 m. bibliotekoje subrendo idėja transformuoti metaduomenis naujos bibliografinės paradigmos pagrindu<sup>64</sup>. Taip buvo sukurta vizualinė rekomendavimo sistema, kuri gali būti grindžiama *susietųjų atvirųjų duomenų* technologijomis. 2019 m. (kai buvo skaitomas pranešimas) metaduomenys buvo vaizduojami grafinėmis sietimis, žinių grafikuose. Pranešėjų nuomone, naudojant RDF rezultatai yra akivaizdūs: užtikrinama duomenų sietis su *Europeana*, Prancūzijos suvestiniu katalogu, VIAF (*Virtual International Authority File*), KB žodynais, kitomis ontologijomis ir tezaurais, pavyzdžiui, *Agrovoc*, *Geonames*, *Thesaurus W*. Be to, naudojant URI, duomenis galima atsisiųsti RDF formatu.

Maxo Plancko istorijos mokslų tyrimų institute Vokietijoje kuriama infrastruktūros platforma, skirta humanitarinių mokslų tyrimams. Konferencijoje pristatydami šios platformos paskirtį Florianas Kräutlis ir Esther Chen pabrėžė, jog ji turėtų užtikrinti efektyvesnį skaitmeninių tyrimų rezultatų suradimą, prieigą ir daugkartinį jų panaudojimą. Vieną instituto skaitmeninių objektų – šaltinių, anotacijų ar pačių

<sup>56</sup> SWIB19: *Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs08>. Žiūrėta 2021-04-03.

<sup>57</sup> SWIB20: *Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: <https://swib.org/swib20/programme.html#abs18>. Žiūrėta 2021-04-03.

<sup>58</sup> <http://data.bibliotheken.nl/>. Žiūrėta 2021-04-03.

<sup>59</sup> Voorburg, R. Publishing Linked Data on Data. Bibliotheken.nl. SWIB19: *Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs08>. Žiūrėta 2021-03-15.

<sup>60</sup> <http://datos.bne.es/inicio.html>. Žiūrėta 2021-03-15.

<sup>61</sup> OCLC. VIAF – *Virtual International Authority File*. Prieiga per internetą: <http://viaf.org/>. Žiūrėta 2021-03-16.

<sup>62</sup> Santos, R. In and out: Workflows between Library Data and Linked-data at the National Library of Spain. SWIB19: *Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs10>. Žiūrėta 2021-03-15.

<sup>63</sup> <https://data.bnf.fr/>. Žiūrėta 2021-03-15.

<sup>64</sup> Lapôtre, R.; Destandau, M.; Pietriga, E. Proposing Rich Views of Linked Open Data Sets: The S-paths Prototype and the Visualization of FRBRized Data in Data.BNF.Fr. SWIB19: *Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs21> <https://www.youtube.com/watch?v=1ih40dwivmA>. Žiūrėta 2021-03-20.

mokslo duomenų bazių prieigą turėtų užtikrinti sumanytas žinių žemėlapis. Didžiausias iššūkis kuriant šią platformą, pasak pranešėjų, yra sujungti duomenis iš įvairių šaltinių, siekiant išlaikyti originalų kontekstą ir detales. Gaunamiems duomenims sujungti pasitelktas CIDOC-CRM, kuris naudojamas žinių žemėlapiui modeliuoti ir duomenims pasikeisti. Infrastruktūrai realizuoti naudojama *ResearchSpace* programinė įranga, sukurta semantinių duomenų platformos *Metaphactory* pagrindu<sup>65</sup>.

JAV istorinių įrašų diskografijos duomenų bazė (*The Discography of American Historical Recordings* – toliau DAHR) yra kuriama Kalifornijos universitete, joje sukaupta per 300 000 plokštelių, kuriose yra daugiau kaip 60 000 atlikėjų įrašai, sudarantys per 40 000 garso failų internete. Siekiant suteikti kuo daugiau konteksto šiems įrašams atliekant mokslinius tyrimus, buvo nuspręsta generuoti susietuosius duomenis, metaduomenis papildant duomenimis iš kitų šaltinių. Pasak Davido Seuberto, Shawno Averkampio ir Michaelo Lashutkos, šio sprendimo rezultatai džiugina, nes DAHR šiuo metu yra susieta su KB autoritetinių įrašų duomenų baze, *Wikidata*, *MusicBrainz*. Vartotojai *Wikipedijoje* dabar gali rasti ne tik biografijas ir nuotraukas, bet ir nuorodas į *Allmusic* sistemos garso įrašų apžvalgas bei *Spotify* sistemos muzikos įrašus. Siekdami paskatinti dar didesnę DAHR naudojimą ir matomumą, naujai sukurtais DAHR URI įrašais yra pildomi *Wikidata* ir *MusicBrainz*. Be išvardytų priemonių, DAHR pildoma VIAF, *Getty*, FAST identifikatoriais bei autoritetinių įrašų informacija, kuri svariai papildo muzikos įrašų duomenų bazę DAHR<sup>66</sup>.

Suomijos nacionalinė biblioteka 2019 m. transformavo suomiškąjį YSA tezaurą, kuriame pateikta per 36 000 žodžių, ir jo atitikmenį švediškąjį *Allårs* tezaurą į ontologiją ir pradėjo naudoti kaip pagrindinę suomiškąją YSO ontologiją – suomių, švedų ir anglų kalbomis. Bibliotekoje subrendus idėjai MARC įrašams konvertuoti naudoti RDA ir URI, buvo parengtos įrašų konvertavimo taisyklės ir įsigyta reikalinga programinė įranga. Konferencijoje pristatydami šį projektą Matias Frosterus, Jarmas Saarikkas ir Okkas Vainonenas akcentavo, jog MARC įrašai buvo konvertuojami ne tik *Melinda* nacionaliniame suvestiniame kataloge, bet ir kitose bibliotekų duomenų bazėse bei vyriausybinių institucijų sistemose, naudojančiose YSA ir *Allårs*. Taip buvo sukurtos atviro kodo konvertavimo programos, kurias gali naudoti bibliotekos ar bibliotekinių sistemų tiekėjai<sup>67</sup>.

Suomijos nacionalinė biblioteka taip pat plėtoja kitą projektą – daugiakalbį automatizuotą dalykinio indeksavimo (dalykinimo) įrankį, sukurtą 2017 m. Pasak Osmos Suomineno, Monos Lehtinenos ir Juhos Inkinenos, sukurti tokį įrankį skatino siekis sutrumpinti laiką, skirtą dalykiniam indeksavimui<sup>68</sup>, bei poreikis turėti daugiakalbį ir įvairius indeksavimo algoritmus palaikantį žodyną. Sėkmingai įgyvendinus projektus šis įrankis per REST API jau taikomas daugelio institucijų veikloje: universitetų bibliotekų dokumentų talpyklose, knygų tiekėjo metaduomenų procesuose, *Dissemin* leidybos paslaugoje, kuria naudojantis publikuojami akademiniai darbai atviros prieigos talpyklose bei automatinio dalykinio indeksavimo paslaugoje *Finto AI*, veikiančioje nuo 2020 m. kaip *Finto* tezauro ir ontologijos paslaugų dalis. Įkeliant viso teksto dokumentą į talpyklą *Finto AI* iš karto pasiūlomi reikšminiai žodžiai, kurie gali būti priskiriami dokumentui ir įrašomi į metaduomenis. Kitas susietųjų duomenų kelias į biblioteką – per knygų tiekėją, kuris gauna iš leidėjų ir dalijasi su biblioteka knygų trumpaisiais metaduomenimis ir santraukomis. Knygų tiekėjas taip pat naudoja *Finto AI* – automatiškai patikrinami iš leidėjų gauti metadu-

<sup>65</sup> Kräutli, F.; Chen, E. Digital Sources and Research Data: Linked and Usable. *SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs21> <https://www.youtube.com/watch?v=1ih40dwivmA>. Žiūrėta 2021-04-23.

<sup>66</sup> Seubert, D.; Averkamp, S.; Lashutka, M. Changing the Tires while Driving the Car: A Pragmatic Approach to Implementing Linked Data. *SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: <https://swib.org/swib20/slides/04-04-seubert.pdf>. Žiūrėta 2021-04-23.

<sup>67</sup> Frosterus, M.; Saarikko, J.; Vainonen, O. 20 Million URIs and the Overhaul of the Finnish Library Sector Subject Indexing. *SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs09>; <https://youtu.be/D2pz5iwi034>. Žiūrėta 2021-04-23.

<sup>68</sup> Bibliotekininkystės vadovėliuose įsitvirtinusi „dalykinimo“ sąvoka; taip vadinamas procesas, kai dokumento turinys išreiškiamas žodine forma – dalykine rubrika ar reikšminiais žodžiais vartojant natūralios kalbos žodžius. Šiame tekste apibūdinamam procesui išreikšti vartojamas lietuviškas atitikmuo „dalykinis indeksavimas“ (angl. *Subject indexing*).

menys, santraukos ir pateikiami reikšminių žodžių ar dalykinių rubrikų pasiūlymai, kuriuos pasirenka metaduomenų specialistai tiekėjo organizacijoje. Tiekėjas dalijasi aprašais ne tik su bibliotekomis, bet ir su knygyvais. Siekiant aukštesnės kokybės biblioteka daug dėmesio skiria indeksavimo algoritmams testuoti. Pasak pranešėjų, dalykinio indeksavimo kokybės tyrimas parodė, kad bibliotekinių priskirtos dalykinės rubrikos ar reikšminiai žodžiai yra tikslesni, bet jiems nedaug nusileidžia automatinis dalykinis indeksavimas<sup>69</sup>.

Automatiniam dalykinių rubrikų ar reikšminių žodžių priskyrimui naudojant kontroliuojamus žodynus ir šiems procesams sukurta speciali PĮ yra taikoma ir Karlo III universitete Madride. Pasak Ricardo Eito-Bruno, atliekant automatinį dalykinį indeksavimą programinėmis priemonėmis ieškoma reikšminių žodžių tekstuose ir atitinkamų dalykinių rubrikų kontroliuojamame žodyne. Programinė įranga grindžiama atvirosios inovacijos platforma (angl. *open innovation platforms*). Atvirų inovacijų platformos – tai interneto puslapiai, veikiantys kaip kompanijų, žmonių ir kt. „direktorijos“, kurių tikslas palaikyti bendradarbiavimo procesą kaupiant duomenis apie autorius ir jų grupes, organizacijas, taip pat jų darbus ir pasiekimus. Pranešėjo nuomone, duomenims indeksuoti pradėjus naudoti *MeSH* įrašus, atsirado puiki galimybė autorių metaduomenis papildyti reikšminiais žodžiais iš metaduomenų apie pačių autorių darbus. Skirtingų sistemų semantinei sąveikai užtikrinti naudojama OAI-PMH, tekstinių įrašų ir SKOS tezauro terminų atitikmenims nustatyti naudojamas *PoolParty* įrankis<sup>70</sup>.

Vokietijos centrinės bibliotekų asociacijos (*Verbundzentrale des GBV*) iniciatyva sukurta internetinė PĮ *Cocoda*, užtikrinanti bibliotekų klasifikavimo sistemų, autoritetinių failų, žinių žemėlapių, tezaurų, ontologijų ir kitų panašių žinių organizavimo schemų pildymą ir jų valdymą šalies mastu. Jacobas Voß'as, konferencijoje pristatydamas *Cocoda*, akcentavo ir kitus šios įrangos funkcionalumus, leidžiančius integruoti daugelį tokių schemų iš įvairiausių šaltinių – visas jas galima peržiūrėti ar užregistruoti registre „Basic Register of Thesauri, Ontologies & Classifications“ (BARTOC) ir naudoti jo identifikatorius. Registre jau yra užregistruota Dewey dešimtainė klasifikacija, *Europeanos* maisto ir gėrimų klasifikavimo schema, KB autoritetinių įrašų ir dalykų žodynai, turintys URI, tad yra galimos jų tarpusavio sąveikos per *Cocoda*. BARTOC yra patvirtintas Tarptautinės žinių organizavimo bendruomenės (angl. *International Society for Knowledge Organization – ISKO*<sup>71</sup>). Susietųjų atvirųjų duomenų kontekste PĮ *Cocoda* funkcionalumai neabejotinai yra svarus indėlis praplečiant šiuolaikinių bibliotekų paieškos galimybes, nes suteikia galimybę ieškoti įvairiose žinių schemose reikiamo reikšminio žodžio, o radus tinkamą suteikia galimybę susieti su savo žodynu, palygina ir pateikia galimų žodžių ar sąvokų ryšių rekomendacijas, nustato, kas jau turi ryšius, o kam dar reikia sukurti.

J. Voß'as ypač akcentavo ryšių tarp skirtingų schemų galimybės sukūrimą bei galimybę nustatyti ryšių tipus, kurių šiuo metu yra keletas. Pavyzdžiui, galima išsaugoti savo sukurtus žodžių ar sąvokų ryšius internete, atsisiųsti kaip JSKOS CSV failą, o ateityje bus galima išsaugoti tiesiai ir savo duomenų bazėje. Sukurta PĮ naudoja JSKOS duomenų formatą ir API. Yra galimybė į *Cocoda* įkelti žodynus CSV, RDF, MARCXML ir kitais formatais<sup>72</sup>.

2020 m. kalbėdamas tos pačios konferencijos tęsinyje J. Voß'as kartu su bendraautoriumi Stefanu Peteru pristatė kitą *susietųjų atvirųjų duomenų* aspektu reikšmingą Vokietijos centrinės bibliotekų asociacijos iniciatyvą. Tai infrastruktūra, leidžianti bibliotekoms keistis žinių organizavimo sistemomis-klasifikacijomis, autoritetiniais failais, tezaurais, ontologijomis. Ši infrastruktūra orientuota į Vokietijos bibliotekas

<sup>69</sup> Suominen, O; Lehtinen, M; Inkinen, J. Annif and Finto AI: DIY Automated Subject Indexing from Prototype to Production. *SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: <https://swib.org/swib20/slides/01-02-suominen.pdf>; <https://youtu.be/nzK97hzPMNE>. Žiūrėta 2021-04-23.

<sup>70</sup> Eito-Brun, R. Automatic Indexing of Institutional Repository Content Using SKOS. *SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: <https://swib.org/swib20/slides/01-01-brun.pdf>; <https://youtu.be/KIDbM3VaRHl>. Žiūrėta 2021-04-24.

<sup>71</sup> <http://www.isko.org/>. Žiūrėta 2021-04-23.

<sup>72</sup> Voß, J. Controlled Vocabulary Mapping with Cocoda. *SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html>. Žiūrėta 2021-04-23.



ir leidžia rasti ir sukurti ryšius tarp bibliotekų sistemų, naudojančių Dewey dešimtainę klasifikaciją, *Regensburg* klasifikaciją, *Basisklassifikation*<sup>73</sup>.

## Išvados

1. Atliktas tyrimas leidžia daryti išvadą, kad *susietųjų atvirųjų duomenų paradigmą* suformavo lygia-grečiai besiformuojančios kryptys, kartu kuriančios naują perspektyvą – *susietuosius atviruosius duomenis*. Sąlygiškai galima išskirti tris *susietųjų atvirųjų duomenų paradigmas*:

- ✓ *atvirųjų duomenų ir informacijos valdymo politikos paradigma*, išreikšta *Skaitmeninėje darbotvarkėje*, EK atvirųjų duomenų strategijoje, *VSI direktyvoje 2013/37/ES* ir kituose strateginiuose ES ir Lietuvos dokumentuose, išvėlgiančiuose duomenų atvėrimo potencialą, jų galimą naudą ir nurodanti jų atvėrimo kryptį;
- ✓ *susietųjų atvirųjų duomenų metodologijos paradigma*, išreikšta *RDF – Išteklių aprašo modelyje* ir *Susietųjų duomenų principuose*, leidžianti aprašyti, skelbti ir sujungti struktūrintus duomenis internete;
- ✓ *susietųjų atvirųjų duomenų technologijų paradigma*, išreikšta *OWL, SKOS ir SPARQL* semantinio saityno kalbomis, užtikrinančiomis duomenų iš skirtingų šaltinių bei nesusijusių sistemų heterogeninių duomenų sietį semantiniame saityne.

2. Tyrimas patvirtina tai, jog *susietųjų atvirųjų duomenų* paradigma tapo svarbiausiu metodologiniu pagrindu, skatinančiu:

- ✓ daugelyje šalių parengti strateginius atvirųjų duomenų iniciatyvų įgyvendinimo dokumentus, užtikrinančius platesnį viešojo sektoriaus duomenų atvėrimą visuomenei;
- ✓ modernizuoti arba kurti naujas informacijos sistemas naudojant *RDF* ir semantinio saityno technologijas, grindžiamas *OWL, SKOS ar SPARQL* semantinio saityno kalbomis, leidžiančias pagerinti įvairių sistemų duomenų sietį, padidinti informacinės paieškos tikslumą ir iki minimumo sumažinti duomenų dubliavimą.

3. Tyrimu nustatyta, kad pereinant prie bibliotekų duomenų ir metaduomenų modelių pertvarkos *susietųjų atvirųjų duomenų* kryptimi reikia žengti tris pagrindinius žingsnius:

- ✓ nuo tezauro pereiti prie formalios ontologijos;
- ✓ pakeisti dabartinę bibliotekose įsivertinusių požiūrį į išteklių aprašymą bei transformuoti katalogavimo procesą taikant bibliotekų duomenų modelius, grindžiamus nauja bibliografinė paradigma – *kūrinio, išraiškos, apraiškos ir vieneto* esybių aprašymu ir disponuojančia esybių, atributų ir santykių struktūra;
- ✓ informacinėms sistemoms būtina sukurti arba pritaikyti specialią semantinės paieškos programinę įrangą, grindžiamą semantinio saityno technologijomis.

4. Gauti tyrimo duomenys leidžia daryti išvadą, kad Lietuvoje dėl duomenų pertvarkos *susietųjų atvirųjų duomenų* kryptimi bibliotekų sektoriuje trūksta valstybinio požiūrio:

- ✓ *Atvirųjų duomenų iniciatyvos įgyvendinimo modelyje* (2018) tarp viešojo sektoriaus institucijų numatytas ir platesnis bibliotekų duomenų atvėrimas visuomenei, tačiau strateginiame dokumente *Bibliotekų plėtros strateginės kryptys 2016–2022 m.* ši veikla neaktualizuojama;
- ✓ *Atvirųjų duomenų iniciatyvos įgyvendinimo modelis* yra universalus, tačiau konkrečių sektorių veiklos kontekstas ir naudojami duomenų valdymo standartai yra skirtingi. Todėl bibliotekų sektoriaus duomenų valdymo ypatumai yra svarbūs, tam reikalingas valstybinis požiūris, strategiškai pagrindžiantis mūsų šalies bibliotekų duomenų valdymo *susietųjų atvirųjų*

<sup>73</sup> Voß, J.; Peters, S. Making Use of the Coli-conc Infrastructure for Controlled Vocabularies. *SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: [https://docs.google.com/presentation/d/1UYp9krriWE42lFuCjGG62qajNSLP6\\_HKDJHlOcvkRd0/edit#slide=id.gac736ad0a8\\_0\\_1](https://docs.google.com/presentation/d/1UYp9krriWE42lFuCjGG62qajNSLP6_HKDJHlOcvkRd0/edit#slide=id.gac736ad0a8_0_1). Žiūrėta 2021-03-13.

duomenų kryptimi pertvarką, akcentuojant bibliotekų sektoriaus duomenų integralumą bei numatant finansinę paramą šiai pertvarkai įgyvendinti.

5. Mokslinės literatūros ir *susietųjų atvirųjų duomenų* gerosios praktikos iniciatyvų, įgyvendinamų pasaulio šalių bibliotekose, analizė patvirtina tai, jog investicijos į susietųjų atvirųjų duomenų generavimą bibliotekose sukuria pridėtinę vertę:

- ✓ Su semantinio saityno reikalavimais suderintų duomenų modelių, grindžiamų nauja bibliografinė paradigma, taikymas modeliuojant duomenis ir modernizuojant informacines sistemas sudaro pagrindą automatizuoti reikšminių žodžių priskyrimą, naudoti reikšminius žodžius ir kitą semantiškai susijusią informaciją iš kitų organizacijų žodynų, ontologijų bei įvairių duomenų bazių, suteikiančių galimybę įgyvendinti prasmingesnę ir efektyvesnę paiešką, papildant turimą informaciją išoriniais šaltiniais;
- ✓ *susietųjų atvirųjų duomenų* pagrindu organizuojama paieška leidžia pagerinti informacijos tikslumą, atsižvelgiant į informacijos vartotojo ketinimus ir konteksto terminų reikšmę, o ne į atskirų žodžių žodžių prasmę užklausoje. Pagrindinis tokios paieškos siekis yra gauti klausimo prasmę atitinkančius atsakymus. Svarus *susietųjų atvirųjų duomenų* paieškos pranašumas – sutaupyti laiko analizuojant ir filtruojant pateiktus užklausoje rezultatus.

## Literatūra ir šaltiniai

- Berners-Lee, Tim. Linked Data – Design Issues. Retrieved July 23, 2006. Prieiga per internetą: <http://www.w3.org/DesignIssues/LinkedData.html>. Žiūrėta 2021-04-03.
- Berners-Lee, Tim. *The Semantic Web Road Map*, 1998, p. 1. Prieiga per internetą: [https://www.researchgate.net/publication/200034151\\_The\\_Semantic\\_Web\\_road\\_map](https://www.researchgate.net/publication/200034151_The_Semantic_Web_road_map). Žiūrėta 2021-03-06.
- Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services*. Washington: Library of Congress, DC November 21, 2012. Prieiga per internetą: <https://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>
- Bibliografija 2010–2011: mokslo darbai* / mokslinė redaktorė ir sudarytoja R. Varnienė-Janssen. Vilnius: Lietuvos nacionalinė Martyno Mažvydo biblioteka, 2013.
- Bibliotekų plėtros strateginės kryptys 2016–2022 metams*. Patvirtinta Lietuvos Respublikos kultūros ministro 2016 m. balandžio 29 d. įsakymu Nr. ĮV-344, p. 5–6. Prieiga per internetą: <https://lrkm.lrv.lt/uploads/lrkm/documents/files/Biblioteku%20pletros%20kryptys.pdf>
- Bizer, Christian; Heath, Tom; Berners-Lee, Tim. Linked Data – The Story So Far. *International Journal on Semantic Web and Information Systems*, 2009, Vol. 5, No. 3, p. 1–22. DOI: <https://doi.org/10.4018/jswis.2009081901>
- Byrum, John D. Rekomendacijos kuo greičiau patobulinti OPAC ir valstybinės bibliografijos tarnybos vaidmuo sprendžiant šį uždavinį. *Bibliografija 2005–2006: mokslo darbai*, 2007, p. 14–23.
- Catalogue of Criteria for Trusted Digital Repositories. *Nestor Working Group on Trusted Repositories Certification, Version 1*. Deutsche Nationalbibliothek, 2006. Prieiga per internetą: [https://files.dnb.de/nestor/materialien/nestor\\_mat\\_08-eng.pdf](https://files.dnb.de/nestor/materialien/nestor_mat_08-eng.pdf). Žiūrėta 2021-03-15.
- CIDOC Conceptual Reference Model. Version 5.1.2*, 2013. Prieiga per internetą: [http://www.cidoc-crm.org/sites/default/files/cidoc\\_crm\\_version\\_5.1.2.pdf](http://www.cidoc-crm.org/sites/default/files/cidoc_crm_version_5.1.2.pdf). Žiūrėta 2021-03-15.
- Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Open data An Engine for Innovation, Growth and Transparent Governance*. Brussels, 12.12.2011 COM 882 Final. Prieiga per internetą: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0882:FIN:EN:PDF>. Žiūrėta 2021-04-03.
- Creswell, John W. *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*. London: Sage Publications, 2009.

- Dunsire, Gordon; Willer, Mirna. Standard Library Metadata Models and Structures for the Semantic Web. *Library Hi Tech News*, 2011, Vol. 28, No. 3, p. 1–12. DOI: <http://dx.doi.org/10.1108/07419051111145118>
- Dunsire, Gordon; Willer, Mirna. Authority versus Authenticity: The Shift from Labels to Identifiers. In: *Authority, Provenance, Authenticity, Evidence: Selected Papers from the Conference and School Authority, Provenance, Authenticity, Evidence, Zadar, October 2016* / Edited by M. Willer, A. J. Gilliland, M. Tomič. Zadar: Sveučilišteu Zadaru, 2018, p. 87–115.
- Dunsire, Gordon; Willer, Mirna. UNIMARC and Linked Data. *International Federation of Library Associations and Institutions*, 2011, Vol. 37, No. 4, p. 314–326. DOI: <https://doi.org/10.1177/0340035211430008>
- Dunsire, Gordon; Willer, Mirna. Initiatives to Make Standard Library Metadata Models and Structures Available to the Semantic Web. *World Library and Information Congress: 76th IFLA General Conference and Assembly, Gothenburg, Sweden – 10–15 August 2010*. Prieiga per internetą: <http://www.ifla.org/en/ifla76>. Žiūrėta 2021-03-03.
- Eito-Brun, Ricardo. Automatic Indexing of Institutional Repository Content Using SKOS. *SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: <https://swib.org/swib20/slides/01-01-brun.pdf>; <https://youtu.be/KIDbM3VaRHI>. Žiūrėta 2021-04-24.
- Frosterus, Matias; Saarikko, Jarmo; Vainonen, Okko. 20 Million URIs and the Overhaul of the Finnish Library Sector Subject Indexing. *SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs09>; <https://youtu.be/D2pz5iwi034>. Žiūrėta 2021-04-23.
- Functional Requirements for Authority Data (2009)*. By IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). Prieiga per internetą: <https://www.ifla.org/publications/functional-requirements-for-authority-data>
- Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model. (2010). *IFLA Working Group on the Functional Requirements for Subject Authority Records (FRSAR)*. Editors: Marcia Lei Zeng, Maja Žumer, Athena Salaba. Prieiga per internetą: <https://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>. Žiūrėta 2021-03-24.
- Functional Requirements for Bibliographic Records: Final Report. *IFLA Study Group on the Functional Requirements for Bibliographic Records*. München: Saur, 1998. Amended 2009. Prieiga per internetą: <https://www.google.com/search?q=%3A+http%3A%2F%2Fwww.ifla.+org%2Fen%2Fpublications%2Ffunctional-requirements-for-bibliographic-records.&oq=%3A+http%3A%2F%2Fwww.ifla.+org%2Fen%2Fpublications%2Ffunctional-requirements-for-bibliographic-records.&aqs=chrome..69i57.1570j0j7&sourceid=chrome&ie=UTF-8>. Žiūrėta 2021-04-03.
- ISO 14721:2012 Space Data and Information Transfer Systems – Open Archival Information System (OAIS) – Reference Model*. Prieiga per internetą: <https://www.iso.org/standard/57284.html>. Žiūrėta 2021-03-15.
- Kräutli, Florian; Chen, Esther. Digital Sources and Research Data: Linked and Usable. *SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs21> <https://www.youtube.com/watch?v=1ih40dwivmA>. Žiūrėta 2021-04-23.
- Lapôtre, Raphaëlle; Destandau, Marie; Pietriga, Emmanuel. Proposing Rich Views of Linked Open Data Sets: The S-paths Prototype and the Visualization of FRBRized Data in Data.BNE.Fr. *SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs21>; <https://www.youtube.com/watch?v=1ih40dwivmA>. Žiūrėta 2021-03-20.
- Laužikas, Rimvydas; Vaitkevičius, Vygintas. Bendras asmenvardžių, vietovardžių ir istorinės chronologijos tezauras. *Virtualios elektroninio paveldo informacinės sistemos (VEPIS) specifikacija*. 2011.
- LIBIS išteklių įrašų sudarymo vadovas: integruojamasis išteklius* / Parengė: T. Auškalnis, R. Bagušytė, L. Buckienė, V. Čeglienė, D. Gedminienė, G. Kubilienė, I. Kvietkuvienė, A. Markevičiūtė, I. Miliauskienė, D. Norkūnienė, D. Pečiulienė, G. Rutkauskienė, D. Sipavičiūtė, R. Skačkauskienė, R. Švabauskaitė, A. Ušpalienė, dr. V. Vaitkevičiūtė-Verbickienė; atsakomoji redaktorė dr. R. Januševičienė. Vilnius, Lietuvos nacionalinė Martyno Mažvydo biblioteka, 2021. Prieiga per internetą: <https://journals.lnb.lt/metodologijos-paskyra/article/view/738/733>
- Library Linked Data Incubator Group Final Report. W3C Incubator Group Report 25 October 2011* / T. Baker, E. Bermès, K. Coyle, G. Dunsire, A. Isaac, P. Murray, P. Panzer, J. Schneider, R. Singer, E. Summers, W. Waiters, J. Young, and Zheng. Prieiga per internetą: [www.w3.org/Consortium/facts](http://www.w3.org/Consortium/facts). Žiūrėta 2021-03-12.

- Library Technology Guides*. Prieiga per internetą: <https://librarytechnology.org/product/alma/>. Žiūrėta 2021-04-02.
- Lietuvos viešojo sektoriaus atvirų duomenų architektūros modelis*. Vilnius: Informacinės visuomenės plėtros komitetas prie Susisiekimo ministerijos, 2014, p. 9. Prieiga per internetą: [https://ivpk.lrv.lt/uploads/ivpk/documents/files/Veikla/Veiklos\\_sritys/Informacijos\\_pakartotinis\\_naudojimas/Atvir%C5%B3%20duomen%C5%B3%20teikimo%20modelis.pdf](https://ivpk.lrv.lt/uploads/ivpk/documents/files/Veikla/Veiklos_sritys/Informacijos_pakartotinis_naudojimas/Atvir%C5%B3%20duomen%C5%B3%20teikimo%20modelis.pdf). Žiūrėta 2021-04-03.
- Miller, Matt. Library of Congress. Id.loc.gov and Wikidata, One Year Later. *SWIB20: Semantic web in Libraries, 2020*. Prieiga per internetą: [https://youtu.be/OvcBBRZ\\_sUM](https://youtu.be/OvcBBRZ_sUM). Žiūrėta 2021-03-13.
- OCLC. VIAF – *Virtual International Authority File*. Prieiga per internetą: <http://viaf.org/>. Žiūrėta 2021-03-16.
- Overview of the BIBFRAME 2.0 Model. *Library of Congress*, 2016. Prieiga per internetą: <https://www.loc.gov/bibframe/docs/bibframe2-model.html>. Žiūrėta 2021-03-28.
- Putnam, Nathan B. *OCLC Research BIBFRAME 2.0 Analysis. LC BIBFRAME Update, ALA Annual 2018*. Prieiga per internetą: <https://www.loc.gov/bibframe/news/pdf/oclc-converter-alaal2018.pdf>. Žiūrėta 2021-03-28.
- Rasmussen, Pennington, Diane; Cagnazzo, Laura. Connecting the Silos Implementations and Perceptions of Linked Data Across European Libraries. *Journal of Documentation*, 2019, Vol. 75, No. 5, p. 643-666. DOI: <https://doi.org/10.1108/JD-07-2018-0117>
- RDF Schema 1.1*. W3C Recommendation 25 February 2014. Prieiga per internetą: <https://www.w3.org/TR/rdf-schema/>. Žiūrėta 2021-03-15.
- Resource Description Framework*: Prieiga per internetą: <https://www.w3.org/TR/owl-features/>
- Riva, P.; Žumer, M. *FRBRoo, the IFLA Library Reference Model, and now LRMoo: A Circle of Development*, 2018. Prieiga per internetą: <http://library.ifla.org/2130/1/074-riva-en.pdf>. Žiūrėta 2021-03-24.
- Rodríguez Escolano, E. RDA and ISBD: History of a Relationship. *JLIS.it.*, 2016, Vol. 7, No. 2, p. 49–82. DOI: <http://dx.doi.org/10.4403/jlis.it-11703>
- Santos, Ricardo. In and out: Workflows between Library Data and Linked-data at the National Library of Spain. *SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019*, Hamburg, Germany. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs10>. Žiūrėta 2021-03-15.
- Seubert, David; Averkamp, Shawn; Lashutka, Michael. Changing the Tires while Driving the Car: A Pragmatic Approach to Implementing Linked Data. *SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: <https://swib.org/swib20/slides/04-04-seubert.pdf>. Žiūrėta 2021-04-23.
- Signore, O. *The Semantic Web and Cultural Heritage: Ontologies and Technologies Help in Accessing Museum Information*, 2007. Prieiga per internetą: <https://www.semanticscholar.org/paper/The-Semantic-Web-and-Cultural-Heritage-%3A-Ontologies-Signore/ab4f0ee826bbea097f4d7aa4d5244b67c0caea6>. Žiūrėta 2021-03-12.
- SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html#abs08>. Žiūrėta 2021-04-03.
- SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: <https://swib.org/swib20/programme.html#abs18>. Žiūrėta 2021-04-03.
- Suominen, Osma; Lehtinen, Mona; Inkinen, Juho. Annif and Finto AI: DIY Automated Subject Indexing from Prototype to Production. *SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: <https://swib.org/swib20/slides/01-02-suominen.pdf>; <https://youtu.be/nzK97hzPMNE>. Žiūrėta 2021-04-23.
- Taleras, K. From Many Records to One Graph: Heterogeneity Conflicts in the Linked Data Restructuring Cycle. *Proceedings of the Eighth International Conference on Conceptions of Library and Information Science, Copenhagen, Denmark*, 2013, p. 19–22.
- Theory and Practice of Digital Libraries*. Prieiga per internetą: <http://www.tpd.eu/>. Žiūrėta 2021-04-02.
- The European Commission Policy on Open Data*. Prieiga per internetą: <https://digital-strategy.ec.europa.eu/en/policies/open-data-0>. Žiūrėta 2021-04-03.
- UNIMARC Manual: Bibliographic Format. 3rd edition* / Edited by A. Hopkinson. München: K. G. Saur, 2008.
- UNIMARC Manual: Authorities Format. 3rd edition* / Edited by M. Willer. München: K. G. Saur, 2009.
- Update on the Work of the ISBD Review Group. *Committee on Standards ISBD Review Group*. Prieiga per internetą: <https://www.ifla.org/isbd-rg>



- Varnienė-Janssen, Regina; Kuprienė, Jūratė. Authenticity and Provenance in Long-Term Digital Preservation: Analysis of the Scope of Content. *Informacijos mokslai*, 2018, t. 82, p. 131–160.
- Varnienė-Janssen, Regina; Laužikas, Rimvydas; Vaitkevičius, Vyktintas; Juškys, Jonas. Skaitmeninto kultūros paveldo vientiso turinio kūrimo strateginiai, metodologiniai ir techniniai sprendimai: lietuviškoji koncepcija. *Bibliografija 2010–2011: mokslo darbai*, 2013, p. 43–55.
- Varnienė-Janssen, Regina; Šermokas, Albertas. Ontologies and Technologies for Integrating and Accessing Digital Cultural Heritage: Lithuanian Approach. *Informacijos mokslai*, 2020, t. 88, p. 66–82. DOI: <https://doi.org/10.15388/Im.2018.82.9>
- Varnienė-Janssen, Regina; Šermokas, Albertas. Provenance in the Context of Digital Cultural Heritage Content the Lithuanian Approach. In: *Authority, Provenance, Authenticity, Evidence: Selected Papers from the Conference and School Authority, Provenance, Authenticity, Evidence, Zadar, October 2016* / Ed. by M. Willer, A. J. Gilliland, M. Tomič. Zadar: Sveučilšteu Zadaru, 2018, p. 213–257.
- Voß, Jakob; Peters, Stefan. Making Use of the Coli-conc Infrastructure for Controlled Vocabularies. *SWIB20: Semantic Web in Libraries. Held Online on 23–27 November 2020*. Prieiga per internetą: [https://docs.google.com/presentation/d/1UYp9krriWE42lFuCjGG62qajNSLP6\\_HKDJHlIoCvkRd0/edit#slide=id.gac736ad0a8\\_0\\_1](https://docs.google.com/presentation/d/1UYp9krriWE42lFuCjGG62qajNSLP6_HKDJHlIoCvkRd0/edit#slide=id.gac736ad0a8_0_1). Žiūrėta 2021-03-13.
- Voß, Jakob. Controlled Vocabulary Mapping with Cocoda. *SWIB19: Semantic Web in Libraries. Held on 25–27 November 2019, Hamburg, Germany*. Prieiga per internetą: <https://swib.org/swib19/programme.html>. Žiūrėta 2021-04-23.
- Web Ontology Language*: Prieiga per internetą: [https://en.wikipedia.org/wiki/Web\\_Ontology\\_Language#OWL\\_sublanguages](https://en.wikipedia.org/wiki/Web_Ontology_Language#OWL_sublanguages)
- Willer, Mirna; Dunsire, Gordon. ISBD, the UNIMARC Bibliographic Format, and RDA: Interoperability Issues in Namespaces and the Linked Data Environment. *Cataloging & Classification Quarterly*, 2014, Vol. 52, No. 8, p. 888–913. DOI: <https://doi.org/10.1080/01639374.2014.921260>
- Willer, Mirna; Dunsire, Gordon; Bosčić, Boris. ISBD and the Semantic Web. *JLIS.it: Italian Journal of Library and Information Science*, 2010, Vol. 1, No. 2, p. 213–236. DOI: <http://dx.doi.org/10.4403/jlis.it-4536>
- Willer, Mirna. Name Authority Control: From Conceptual Model to the Semantic Web and Linked Data. *Summer School in the Study of Historical Manuscripts: Proceedings* / Edited by M. Willer and M. Tomič. Zadar: Sveučilište u Zadru, 2013, p. 243–264.
- WorldCat*. Prieiga per internetą: <https://www.worldcat.org/>. Žiūrėta 2021-04-02.
- World Wide Web*. Prieiga per internetą: <https://www.w3.org/DesignIssues/LinkedData.html>
- Ziku, Mariana. Digital Cultural Heritage and Linked Data: Semantically-Informed Conceptualisations and Open Practices with Focus on Intangible Cultural Heritage. *Liber Quarterly The Journal of the Association of European Research Libraries*, 2020, Vol. 30, No. 1, p. 1–16. DOI: <https://doi.org/10.18352/lq.10315>