

Standartinių bibliotekų metaduomenų modelių ir struktūrų pateikimo semantiniame saityne iniciatyvos

Gordon DUNSIRE

Strathclyde universitetas, Glazgas, Škotija, Jungtinė Karalystė, el. p. gordon@gordondunsire.com

Mirna WILLER

Zadaro universitetas, Kroatija, el. p. mwiller@unizd.hr

Straipsnyje aptariamos naujausios bibliotekų metaduomenų modelių ir struktūrų, tarp jų – IFLA standartų „Functional Requirements for Bibliographic Records“ (FRBR) („Funkciniai reikalavimai bibliografiniams įrašams“), „Functional Requirements for Authority Data“ (FRAD) („Funkciniai reikalavimai autoritetiniams duomenims“) ir „International Standard Bibliographic Description“ (ISBD) („Tarptautinis standartinis bibliografinis aprašas“) kartu su juos palaikančia infrastruktūra pateikimo semantiniame saityne iniciatyvos.

FRBR peržiūros grupė, pasitelkdama RDF, RDF schemą (RDF Schema – RDFS), Paprastą žinių organizavimo sistemą (Simple Knowledge Organization System – SKOS) ir Saityno ontologijų kalbą (Web Ontology Language – OWL), šiuo metu rengia FRAD bei FRBR entiteto ir santykio modelio pateiktis „Išteklių aprašo modeliu“ (Resource Description Framework – RDF) grindžiamoms taikomosioms programoms, atitinkamai atvejais nustatydamas abiejų modelių sankirtas. ISBD/XML tyrimų grupė tiria ISBD pateikimo RDF galimybes. Pagal IFLA Vardų erdvių projektą kuriama administracinė ir techninė infrastruktūra, kuri padėtų įgyvendinti šias iniciatyvas ir skatintų kitas tarnybas perimti šiuos standartus.

Straipsnyje apibūdinamos ir panašios tokių susijusių išorinių standartų, kaip RDA, REICAT (naujosios Italų katalogavimo taisyklės) ir CIDOC CRM, iniciatyvos. DCMI RDA laikina darbo grupė kartu su RDA jungtiniu vykdomuoju komitetu rengia su FRBR ir FRAD suderintas RDA struktūrinių elementų semantinio saityno pateiktis ir kontroliuojamus metaduomenų turinio žodynus. REICAT irgi grindžiamas FRBR, o į objektą orientuota FRBR versija susieta su CRM, kuris, savo ruožtu, pateikiamas RDF. CRM būtinumas iš pradžių buvo grindžiamas muziejų bendruomenei skirtų metaduomenų reikalingumu; dabar siekiama jį pritaikyti ir archyvams, galiausiai sukuriant modelį, kuris būtų bendras visoms pagrindinėms kultūros informacijos sritims: archyvams, bibliotekoms ir muziejams. Žodynų sutapties schemos (Vocabulary Mapping Framework – VMF) projekto metu sukurta semantinio saityno priemonė, automatiškai nustatanti skirtingų informacijos bendruomenių, įskaitant leidėjus, metaduomenų modelių sankirtas. Priemonės pagrindą sudaro keletas standartų, tarp kurių – CRM, FRAD, FRBR, MARC21 ir RDA.

Straipsnyje aptariama šių iniciatyvų svarba, susijusi su gausių turtingų profesionaliai sukurtų minėtus standartus atitinkančiais formatais (pvz., UNIMARC ir MARC21) saugomų metaduomenų pateikimu kaip susietų duomenų, taip pat su tokiomis problemomis, kaip statistinėms ir semantinėms išvadoms daryti būtina kritinė masė, integravimas su vartotojų ir kompiuterių kuriamais metaduomenimis, autentiškumas, teisingumas bei patikimumas. Straipsnyje aptariamos ir iniciatyvos, kuriomis siekiama kaip susietus duomenis teikti kontroliuojamiems žodynams, tokiems kaip Dewey dešimtainė klasifikacija (Dewey Decimal Classification – DDC), ISBD, Kongreso bibliotekos vardų autoritetinis failas (Library of Congress Name Authority File – LCNAF), Kongreso bibliotekos dalykinės rubrikos (Library of Congress Subject Headings – LCSH), Rameau (Prancūzų dalykinės rubrikos), Universalioji dešimtainė klasifikacija (Universal Decimal Classification – UDC) ir Virtualus tarptautinis autoritetinis failas (Virtual Authority File – VIAF). Ir, galiausiai, straipsnyje aptariama galima visų šių iniciatyvų įtaka metaduomenų kūrimo procesams ir jų valdymo sistemoms.

Reikšminiai žodžiai: IFLA standartai ir modeliai; kontroliuojami žodynai; metaduomenys; semantinis saitynas.

Parengiamoji veikla ir įžanga

Pastaruoju metu iškelta daug IFLA parengtų standar-

tinių bibliotekų metaduomenų modelių, struktūrų ir žodynų pateikimo semantiniame saityne iniciatyvų. Tuo tikslu buvo rengiami ir susijusiems kitų organizacijų prižiūri-

miems standartams skirti projektai. IFLA veiklą šioje srityje skatinęs pradinis veiksnys buvo būtent tokio pobūdžio išorinis projektas, kuriam atsirasti padėjo „RDA: išteklių aprašo ir prieigos“¹, „Dublin Core metaduomenų iniciatyvos“ (DCMI)², IEEE mokymosi objektų metaduomenų (*Learning Object Metadata* – IEEE LOM) ir SKOS³ atstovų pasitarimas. Britų bibliotekoje Londone 2007 m. surengtame pasitarime, skirtame duomenų modeliams⁴, buvo rekomenduotos kelios veiksmų kryptys, kurių vienas iš naudingumo aspektų būtų „metaduomenų standarto, suderinamo su saityno sandara ir pasižyminčio labai gera sąveika su kitomis semantinio saityno iniciatyvomis“ pateikimas bibliotekų bendruomenei. Šios veiklos kryptys apima RDA bibliografinių entitetų, santykių ir kontroliuojamos turinio terminijos plėtrą siekiant didesnio jų suderinamumo su *Dublin Core* bei semantiniu saitynu. Pasitarimo metu buvo suformuota DCMI RDA darbo grupė⁵, kuriai pavesta prižiūrėti ir dalyvauti šioje veikloje, daugiausia apimančioje RDF⁶, kuris yra semantinio saityno duomenų modelis, klasių (atitinkančių entitetus) ir savybių (atitinkančių entitetų santykius ir atributus) formulavimą.

RDA pagrindą sudaro IFLA metaduomenų modeliai „Funkciniai reikalavimai bibliografiniams įrašams“⁷ (FRBR) ir „Funkciniai reikalavimai autoritetiniams duomenims“⁸ (FRAD); jame dažnai remiamasi šiuose modeliuose apibūdinamais bibliografiniais entitetais, santykiais ir atributais. Todėl RDA elementų RDF formuluotėse privalo atsispindėti FRBR ir FRAD terminai ir apibrėžtys, kontroliuojamos ne tiek Jungtinio vykdomojo RDA plėtros komiteto (*Joint Steering Committee for Development of RDA (JSC)*), kiek FRBR peržiūros grupės. FRBR peržiūros grupė 2007 m. Durbane (Pietų Afrika) Pasaulio bibliotekų ir informacijos kongreso metu vykusiame pasitarime⁹ atsižvelgė į Pasitarimo, skirto duomenų modeliams, rezultatus ir nusprendė inicijuoti projektą „FRBR (entiteto ir santykio) atitinkančių vardų erdvių apibrėžimas RDF ir kitose sintaksėse“. Vardų erdvė – tai RDF pateikčių publikavimo ir prieigos prie jų teikimo kitoms semantinio saityno programoms būdas, todėl JSC, užuot užsiėmęs savų versijų publikavimu ir tvarkymu, gali naudoti FRBR vardų erdvės elementus.

IFLA Katalogavimo sekcijos ISBD peržiūros grupė neseniai ėmėsi įgyvendinti Medžiagos apibūdinimo tyrimų grupės rekomendaciją parengti ISBD skirtą XML schemą. ISBD peržiūros grupė laikė būtina ištirti ISBD sąvokas ir patį standartą taikant saityno technologijas galimybes, galiausiai išplėtoję šį standartą iki priemonės, atviros semantinio saityno technologijoms ir paslaugoms. 2008 m. Kvebeke Pasaulio bibliotekų ir informacijos kongreso metu buvo įkurta ISBD/XML tyrimų grupė. Ji pasiūlė inicijuoti projektą, kuriam pritarė Profesinė taryba ir kuris vykdomas jau dvejus metus. ISBD/XML tyrimų grupės pagrindinis uždavinys – pasiekti, kad ISBD

taptų svarbiu struktūruotos bibliografinės informacijos vertinimo pasaulio informacijos kontekste veiksnium. Tuo tikslu Tyrimų grupė buvo įpareigota parengti ISBD XML schemą; tačiau, atsižvelgus į veiklą, atliktą derinant RDA ir FRBR su RDF, buvo nutarta, apeinant bendrąją XML, pereiti tiesiai prie RDF/XML aplinkos. Tokio projekto rezultatas turėtų būti ISBD RDF/XML. Tyrimų grupės pasitarime Milane, vykusiame 2009 m. Pasaulio bibliotekų ir informacijos kongreso metu, buvo nustatytos ir patvirtintos tokios veiklos kryptys: remti pastangas, kuriomis siekiama suburti IFLA III skyriaus sekcijų ir kitų padalinių atstovus į specialią arba bendrą darbo grupę įpareigojant ją pasiekti, kad IFLA standartai ir modeliai taptų autoritetingais dokumentais, naudingais semantinio saityno paslaugoms ir priemonėms (išsamiau apie tai žr. toliau), pritari Tyrimų grupės veiklai VMF projekte, analizuoti technines ir su modeliavimu susijusias ISBD taikymo RDF/XML aplinkoje problemas (žr. toliau), apibrėžti su RDF/XML sintakse suderinto ISBD taikymą ir jo funkcijas, parengti ISBD RDF/XML schemas projektą, išanalizuoti ir apibrėžti ISBD elementų funkcijas remiantis FRBR, UNIMARC, MARC bei DC/XML schemomis, naujomis katalogavimo taisyklėmis (tokiomis kaip RDA ir REICAT) bei Suomų katalogavimo taisyklėmis, išanalizuoti susietų duomenų koncepciją ir jai pritari bei aiškinti pardavėjams tokių duomenų svarbą, siekiant skatinti naujos kartos bibliotekų informacijos sistemų plėtrą¹⁰.

2008 m. Kvebeko (Kanada) ir 2009 m. Milano (Italija) pasaulio bibliotekų ir informacijos kongresų metu apie FRBR vardų erdvės projektą buvo diskutuojama kitose už bibliografinius standartus atsakingose IFLA grupėse ir tarp šių grupių. Todėl IFLA Klasifikavimo ir dalykinimo sekcija pasiūlė sudaryti specialią darbo grupę ir įpareigoti ją parengti reikalavimų ir pasirinkimo galimybių dokumentą, kuris atspindėtų IFLA pritarimą IFLA standartų pateikimui formatais, tinkamais taikyti semantiniame saityne. Šiam pasiūlymui pritarė IFLA Bibliografijos, Katalogavimo ir Informacinių technologijų sekcijos, ir 2009 m. pabaigoje buvo suburta speciali IFLA Vardų erdvių grupė. 2010 m. kovo mėn. ši grupė pristatė IFLA savo pranešimą.

IFLA standartai

Daugelis IFLA standartinių bibliografinių modelių ir programinių priemonių yra formaliai arba neformaliai susijusios. Formalus santykis dažniausiai pasireiškia tuo, kad vienas standartas tiesiogiai remiasi kitu, arba skirtingų standartų elementų sankirtomis, patvirtintomis šiuos standartus prižiūrinčių IFLA grupių. Neformali sietis dažniausiai pasireiškia sutaptimi, nustatyta IFLA nepriklausančių grupių ar pavienių asmenų. Svarbus yra laiko aspektas: gali praeiti keletas metų, kol kokia nors tarptautinė ir didžiajai dalimi savanoriška (kas tipiška IFLA) grupė parengs stan-

dar, ir dar keletą metų gali užtrukti jo peržiūra, taisymas bei susiejimas su kitais standartais, kurių rengimo etapas gali būti visai kitas. Ryškus pavyzdys gali būti tai, kas dabar žinoma kaip „Funkcinių reikalavimų“ šeima arba „FRBR modelių šeima“. FRBR buvo paskelbti 1998 m., ir rengėjų jie buvo apibūdinti kaip „pirminis bandymas“; juose toliau pažymima, kad „modelis galėtų būti praplėstas, siekiant apimti duomenis, paprastai pateikiamus autoritetiniuose įrašuose. Ypač būtina išsamiau ištirti entitetus, kurie labai svarbūs dalykų autoritetiniams duomenims, tezaurams ir klasifikacijos sistemoms, taip pat – šių entitetų santykiams.“ FRAD, kurie susiję būtent su pirmuoju siūlymu dėl praplėtimo autoritetiniais įrašais, buvo paskelbti praėjus 11 metų – 2009-aisiais. Tikimasi, kad 2010 m. bus paskelbti *Functional Requirements for Subject Authority Data – (FRSAD)* („Funkciniai reikalavimai dalykų autoritetiniams duomenims“), susiję su antruoju siūlymu – praplėsti FRBR dalykų autoritetiniais duomenimis.

FRBR šeima

Parengiamuoju FRBR vardų erdvių projekto etapu, siekiant susipažinti su RDF sąvokomis ir įvertinti FRBR elementų RDF pateiktis, buvo pasinaudota Nacionalinės mokslinės skaitmeninės bibliotekos (*National Science Digital Library – NSDL*)¹¹ Metaduomenų registro eksperimentine baze. FRBR peržiūros grupė savo pasitarime 2008 m. Pasaulio bibliotekų ir informacijos kongreso metu apsvarstė šios veiklos ataskaitą¹², joje šiai Grupei rekomenduojama parengti savas vardų erdves, kurios būdu kaip nors išskirtinai paženklintos arba nurodoma jų nuosavybė. 2009 m. kongreso metu IFLA svetainės administratorius pritarė saityno srities, kuri apimtų visas IFLA vardų erdves, įsigijimui, ir po kiek laiko šiam tikslui buvo užregistruota sritis „iflstandards.info“. Tuo pačiu metu buvo pasiūlyta metodika, kaip pritaikyti šią sritį specifiniams IFLA standartams. Ta metodika vėliau buvo taikoma NSDL Metaduomenų registro produkcijos srityje rengiant FRBR RDF pateiktis. FRBR_{er} modelio elementų rinkinys¹³ jau beveik baigtas rengti – tam, kad jis būtų patvirtintas, Peržiūros grupei beliko išspręsti keletą problemų. Pavyzdžiui, keletas šio modelio atributų atitinka FRBR pirmosios grupės elementų potipius (pvz., sumažinimo santykis, taikomas tik apraiškoms, kurios yra mikroformos). Norint tai pateikti RDF, tektų sukurti savybę „turi-sumažinimo-santykį“ ir klasę „mikroforma-apraiška“, kuri būtų „apraiškos“ poklasis, tačiau FRBR nėra entiteto „mikroformos apraiška“ apibrėžimo. Kita problema susijusi su FRBR etikečių įvairove. RDF šios etiketės sudaro pasirinktos etiketės, suteikiamos klasės arba savybės URI, kuris daugiau skirtas vartotojui, o ne apdoroti kompiuteriu, pagrindą. Nors etiketės nedaro įtakos RDF pateikčių apdorojimui kompiuteriu, jų įvairovė gali su-

kelti abejonių dėl modelio autoritetingumo ir apsinkinti anglišku etikečių vertimą į kitas kalbas. Šios ir kitos problemos dažniausiai kyla verčiant žmogaus skaitomą dokumentą, kuriame, siekiant padaryti jį suprantamesnį, pateikiami etikečių variantai, į kompiuterio skaitomą „dokumentą“ arba RDF pateikčių rinkinį, kuriam būtinos išsamios ir nuoseklios apibrėžtys bei etiketės. Kaip kiekvienos RDF klasės arba savybės URI pagrindas naudojama vardų erdvė „http://iflstandards.info/ns/fr/frbr/frbr/“. Vardų erdvė, kiek tai įmanoma, atitinka FRBR formuluotę ir kaip informacijos šaltiniu išskirtinai remiasi FRBR baigiamąja ataskaita – taip užtikrinama, kad RDF pateikties formuluotė nebūtų iškraipoma atsižvelgiant į ankstesnes interpretacijas arba pritaikant ją kitų bendruomenių reikmėms. Registre¹⁴ pateikiamas RDF išvedinys neatitinka FRBR modelio visumos, net Peržiūros grupei patvirtinus visus elementus, nes Registras nesuderinamas su Saityno ontologijų kalbos (OWL)¹⁵ savybėmis, kurios būtinos norint pateikti sudėtingus iš FRBR entitetų, santykių ir atributų išvestų RDF klasių ir savybių ryšius. Šie santykiai apriboja klases, taip užtikrinant, kad kas nors, pateiktas kaip FRBR „kūrinys“, nebūtų pateikiamas ir kaip FRBR „išraiška“, „apraiška“ ar „vienetas“ arba laikomas tokiais, ir savybių santykius, taip nustatant, kurios savybės sudaro ekvivalentines poras, įgalindamos vieną trijų sandų derinį logiškai išvesti iš kito. Šiems sudėtingiems FRBR entiteto ir santykio modelio aspektams atspindėti bus naudojamos papildomos, už Registro ribų parengtos pateiktys. Peržiūros grupė numato sujungti šias pateiktis su Registro failu, ir išsamią FRBR modelio RDF versiją paskelbti IFLA svetainėje.

Laikotarpiu tarp 2003 ir 2009 m. buvo parengti FRBR_{oo}¹⁶ – į objektą orientuota FRBR versija, atitinkanti ir praplečianti CIDOC CRM¹⁷, skirtą muziejų bendruomenės metaduomenims. Vienas iš šios iniciatyvos siekinių buvo sudaryti galimybę FRBR taikyti RDF programoms ir kitoms semantinio saityno technologijoms. 2009 m. buvo paskelbta dalinė CRM RDFS (RDF schemas) pateiktis¹⁸, tačiau ji neapima FRBR_{oo} išplėtos. Kartu su FRBR_{oo} paskelbta FRBR_{oo} ir FRBR_{er} elementų sutaptis savo laiku bus panaudota FRBR_{er} ir FRBR_{oo} elementams RDF aplinkoje palyginti bei FRBR_{er}, FRBR_{oo} ir CRM modelių sąsajai tobulinti.

Jau pradėta įgyvendinti FRAD RDF pateiktis laikantis panašios nuostatos, kaip ir FRBR atveju. Kad nebūtų painiojama su FRBR_{er} modeliu, naudojama atskira vardų erdvė „http://iflstandards.info/ns/fr/frad/“. Vėlgi RDF neapima entitetų, santykių ir atributų, FRAD apibrėžiamų kaip FRBR elementai; FRBR peržiūros grupė svarsto galimybę tam tikslui atitinkamai panaudoti FRBR_{er} vardų erdvės klases ir savybes. Tačiau FRAD praplečia kai kurių FRBR elementų apibrėžtis ir taikymo sritį; ten, kur dėl šios priežasties atsiranda semantiškai svarbių skirtumų, FRBR elementui atstovauja nauja FRAD vardų

erdvės klasė arba savybė. FRBR peržiūros grupė, tęsdama numatytą FRBR modelių šeimos konsolidavimo veiklą, atsižvelgia į tas sritis, kurioms būdinga FRBR_{cr} ir FRAD modelių išvestis vienas iš kito, sutaptis ir sąsaja.

Paskutinis modelis šioje šeimoje yra FRSAD. FRSAD pasiekė paskutinįjį plėtros etapą. Panašu, kad, jį taikant RDF pateiktims rengti, bus laikomasi tos pačios nuostatos, kaip ir FRBR bei FRAD atvejais. Vėlgi tikimasi, kad galiausiai tai paveiks būsimą modelių konsolidavimo veiklą.

ISBD

2010 m. gegužės 10 d. pasaulinei peržiūrai buvo pateikta Jungtinės ISBD laidos parengiamoji versija¹⁹, taip ISBD/XML tyrimų grupei suteikiant galimybę naudotis ja kaip pakankamai brandžiu tekstu pradedant RDF pateikčių testavimą NSDL Registro eksperimentinėje bazėje. Testavimo metu būtina apibrėžti ir išnagrinėti dėl RDF taikymo kylančias problemas, atitinkamais atvejais remtis RDA patirtimi, nors tuo pat metu ypač svarbu pasiekti ir tai, kad ISBD rastų savo vietą tarp FRBR ir (UNI) MARC bei DC/XML schemų.

ISBD statusas tarp IFLA dokumentų kelia strategiškai (ir politiškai) svarbų klausimą: „kas sudaro kieno pagrindą?“ Taikymo lygmenyje FRBR yra konceptualus modelis, grindžiamas entiteto ir santykio sąveikos metodologija, kurios esminė savybė yra pritaikomumas RDF, kai tuo tarpu ISBD yra duomenų standartas. Todėl FRBR ir ISBD santykis arba, tiksliau tariant, ISBD ir FRBR santykis ISBD bendruomenėje ir už jos ribų sukėlė daug diskusijų, po kurių buvo priimti tam tikri su šiuo sąryšiu susiję sprendimai ir galiausiai paskelbta ISBD žodyno RDF pateiktis.

„ISBD peržiūros grupė laikė būtinu IFLA uždaviniu išaiškinti ryšį tarp ISBD ir FRBR modelio. Bandant priderinti ISBD terminiją prie FRBR terminų „kūrinys“ („work“), „išraiška“ („expression“), „apraiška“ („manifestation“) ir „vienetas“ („item“) ir pakeisti tokius terminus kaip „leidinys“ („publication“), grupei iškilo sunkumų, didžiausia dalimi todėl, kad FRBR vartojami terminai apibrėžiami entiteto ir santykio modelio kontekste aukštesniu abstrakcijos lygmeniu negu ISBD apibūdinimai. Frankfurte vykusio IFLA ekspertų pasitarimo, skirto tarptautinėms katalogavimo taisyklėms parengti (IME-ICC), pranešime buvo pasakyta, kad „FRBR terminija *neturi* būti paprasčiausiai įterpiama į ISBD ir katalogavimo taisykles, tačiau [pastarosios] turi išlaikyti savo specifinę terminiją ir tiksliai apibrėžti, kaip šios specifinės terminijos kiekvienas terminas konceptualiai susijęs su FRBR terminija“. Peržiūros grupė atsižvelgė į IME-ICC patarimą ir 2003 m. nusprendė ISBD vengti vartoti FRBR terminus“²⁰.

Kitas strateginis uždavinys yra nustatyti linijinės ISBD sandaros, tiksliau – sudėtinių, arba sancaupinių,

metaduomenų pateikčių devyniose ISBD srityse (pradedant 0 sritimi ir baigiant 8 sritimi) ir RDF sutaptį. Todėl rengiant ISBD RDF pateiktį, svarbu atsižvelgti į šiuos labai svarbius ir aktualius dalykus:

1. Sankaupinių duomenų traktavimas apibrėžtame sričių elementų skaičiuje siekiant nustatyti, kas yra RDF klasė ir kas RDF savybė.

2. Būtinų ir pasirenkamų elementų bei sričių traktavimas.

3. Sričių ir elementų jose eilės tvarka.

4. Besikartojančios sritys ir elementai, kuriems taikomos nekintamos iš anksto numatytos taisyklės, nors šios sritys ir elementai gali būti arba jų gali nebūti (atsižvelgiant į išteklaus informaciją).

5. Skyrybos ir dvejos jos funkcijos traktavimas. Pirmoji funkcija – srities elementų identifikavimas, tačiau kadangi sričių negalima vienareikšmiškai identifiukuoti pagal specifinę skyrybą, pagal skyrybą neįmanoma logiškai išvesti ir srities bei elementų „reikšmės“. Antroji funkcija susijusi su elementų, t. y. bibliografinių duomenų, atvaizdavimu; kadangi duomenys aprašomame ištekliuje gali būti arba jų gali nebūti, atvaizdavimas labai priklauso nuo to, kaip ISBD apibrėžia būtinus ir pasikartojančius elementus, ir išdėstymo tvarkos, nuo kurios priklauso duomenų užrašymas.

Pripažįstama, kad kai kurias iš šių problemų, ypač būtinų elementų naudojimą bei skyrybos ir atvaizdavimo apibrėžtis, būtina spręsti už RDF konteksto ribų ISBD įrašo sudarymo metu naudojant XML pateiktis ir programų profilius²¹.

UNIMARC

UNIMARC yra konvertavimo formatas, skirtas keitimuisi bibliografiniais duomenimis tarp sistemų, naudojamų nacionalinėse bibliotekose ir kitose tarnybose. Jis nenurodo, kokia turi būti tam tikrose sistemose naudojamų metaduomenų *struktūra* ar turinys, tačiau yra glaudžiai suderintas su ISBD. Tokios dermės pasekmė yra antrinės sutaptys su keliais kitais šiame dokumente aptariamais standartais ir modeliais, įskaitant FRBR ir RDA bei, savo ruožtu, su FRAD ir CIDOC CRM²². UNIMARC grindžiami daugelio šalių nacionalinių bibliotekų, sudarančių „oficialiąją“ ir autoritetinę (todėl patikimą) nacionalinių bibliografinių metaduomenų versiją, katalogai ir bibliografija; jis taip pat naudojamas ir kaip vietinė jų sistemose esančių įrašų *struktūra*. Todėl labai svarbu parengti UNIMARC RDF pateiktį kaip metaduomenų *struktūros* schemą, kuri būtų taikoma išskiriant įrašų turinį kaip semantiniam saitynui skirtus susietus duomenis. Tai apibūdinama toliau. RDF pateiktis padėtų papildyti ir UNIMARC bei kitų standartų sankirtas, kurios per pastaruosius metus buvo ženkliai pakoreguotos. Į šiuos aspektus buvo atkreiptas Nuolatinio UNIMARC komiteto dėmesys.

Susiję standartai

DCMI RDA laikina darbo grupė siekia trijų tikslų: apibrėžti RDA modeliavimo objektus kaip RDF savybių ir klasių žodyną; identifikuoti įterptųjų reikšmių žodynus, kurie vėliau būtų pateikti RDFS arba SKOS; parengti RDA skirtą „*Dublin Core* taikomąjį profilį“, kurio pagrindą sudarytų FRBR ir FRAD. Laikina darbo grupė, rengdama RDA žodyną²³ RDF pateiktis, naudojami NSDL Metaduomenų registru. Antrasis tikslas jau beveik pasiektas: visi kontroliuojami metaduomenų turinio – pradedant taikomąja medžiaga ir baigiant vaizdo formatais – žodynai pateikti SKOS. Sudėtingiau įgyvendinti pirmąjį tikslą, nes tam būtina metodologija, skirta RDA ir ISBD būdingiems sudėtiniais arba sancaupiniams metaduomenims. Sancaupinių duomenų pavyzdžiu galėtų būti terminas „publikavimo duomenys“, sudarytas iš elementų „publikavimo vieta“ ir „leidėjas“. Grupė parengė metodologiją, kuri paskelbta dokumente, apžvelgiančiame jos nutarimus ir veiklos rezultatus iki 2009 m. pabaigos²⁴. Ši metodologija ir kitos problemos, iškilusios pateikiant RDA elementus, kaip RDF klases ir savybes, bus aptariamoms ISBD/XML tyrimų grupės. Laikiniai darbo grupei sunkiai sekasi įgyvendinti trečiąjį uždavinį – parengti „taikomąjį profilį“. Iš pradžių buvo ketinta parengti ką nors, kas praplėstų arba pakeistų „*Dublin Core* bibliotekų taikomąjį profilį“ (*Dublin Core Library Application Profile – DC-Lib*)²⁵, tačiau buvo nuspręsta, kad naudingiau būtų specifinėms bibliografinių išteklių rūšims rengti atskirus profilius; diskusijos apie tai tebesitęsia iki šiol.

Kita DCMI RDA laikinai darbo grupei iškilusi problema yra RDA ir FRBR bei RDA ir FRAD santykis RDF kontekste, ir ypač tai, ar RDA vardų erdvė turėtų apimti savas FRBR ir FRAD klasių ir savybių RDA pateiktis, ar galbūt į šias pateiktis turėtų būti nukreipiama FRBR ir FRAD vardų erdvė. FRBR ir FRAD elementų buvimas RDA vardų erdvėje buvo būtinas iš pradžių, nes FRBR vardų erdvės plėtra buvo užsitęsusi ilgiau negu tikėtasi, o FRAD vardų erdvė nebuvo stabili (ji tokia nėra iki šiol). Vėluojant paskelbti RDA, FRBR ir FRAD buvo suteikta galimybė pasivyti, tačiau laikina darbo grupė vis dėlto siūlo į RDA vardų erdvę įtraukti atskiras FRBR ir FRAD pateiktis, taip išlaikant lankstumą ir sumažinant priklausomybę nuo IFLA iniciatyvų. Po kurio laiko, naudojant OWL savybes, būtų galima pasiekti, kad RDA ir IFLA pateiktys taptų tapačiomis, arba atsakyti RDA versijų IFLA versijų naudai.

Švedijos nacionalinė biblioteka parengė MARC21 įrašų pateikimo RDF metodologiją ir pritaikė ją LIBRIS – Švedijos suvestiniam katalogui²⁶. Būta ir kitų panašių MARC21 metaduomenų pateikimo semantiniame saityne pasiūlymų²⁷. Plačiai taikomas formatas yra MARC21, ypač angliškai kalbančių šalių na-

cionalinėse ir kitose bibliotekose. Parengta labai daug MARC21 formato įrašų – vien tik OCLC *WorldCat* jų yra keletas šimtų milijonų.

Italų katalogavimo taisyklės REICAT²⁸ yra pirmosios katalogavimo taisyklės, parengtos atsižvelgiant į FRBR, FRAD ir „Tarptautinių katalogavimo principų išdėstymą“²⁹. Taisyklės grindžiamos kūrinio ir unifikuotos antraštės koncepcijomis; jose kalbama apie FRBR kūrinio ir išraiškos koncepcijas, tačiau ne visada taikomos FRBR apibrėžtys ir vartojami FRBR terminai^{30 31}. Kol kas nėra informacijos apie Nuolatinio Italų katalogavimo taisyklių peržiūros komiteto ketinimus skelbti REICAT RDF pateiktis.

Jungtinės Karalystės jungtinis informacijos sistemų komitetas (*Joint Information Systems Committee – JISC*) finansavo Žodynų sutapties schemos (*Vocabulary Mapping Framework – VMF*) projekto³² kryptį, skirtą ženkliai praplėsti RDA/ONIX išteklių klasifikavimo schemą³³ siekiant sukurti priemonę, automatiškai nustatančią metaduomenų standartų žodynų sankirtas, kurias galėtų taikyti JISC bendruomenė mokslo, mokymo ir mokymosi srityse. Pradinę RDA/ONIX schemą parengė RDA atstovai ir leidėjai; ją sudaro su bibliografinių išteklių turiniu ir laikmenomis susijusių atributų ontologija ir jos taikymo metodologija, kuria siekiama sukurti aukšto lygio žmogaus skaitomas turinio ir laikmenos tipų žymas³⁴. Šia schema grindžiami RDA laikmenos, turinio ir medijų tipų žodynai.

VMF projekte daugiausia dėmesio kreipiamą į bibliografinio išteklių ir bibliografinio veiksnio (grupės) kategorijas ir jų santykius, pripažįstant, kad daugelį aiškių bibliografinio veiksnio vaidmenų, tokių kaip redaktorius ar gamintojas, geriau traktuoti kaip santykius: „<kieno>-redaguota“ arba „yra-<kieno>-gamintojas“. Projekto metu buvo išanalizuoti atitinkami įvairių sričių žodynai ir parengta priemonė – VMF matrica, kurią galima taikyti automatiškai nustatant tiksliausias skirtingų metaduomenų schemų žodynų sankirtas. Pati matrica – tai veiklos rūšių ir funkcijų RDF pateiktys, apimančios visus galimus išteklių ir veiksmų santykius. Pasitelkiant OWL savybes, į matricą gali būti įterpiami išorinių žodynų terminai, ir kadangi matricoje pateikiami visi galimi bet kokius du terminus jungiantys keliai, įmanoma suprogramuoti trumpiausią (t. y. optimalų) „atstumą“ tarp šių dviejų terminų.

Pamatinis rinkinys, kuriuo naudojantis buvo kuriama matrica, apėmė CIDOC CRM, FRAD, FRBR, MARC21 ir RDA žodynų terminus; projekte buvo atstovaujami taip pat ir ISBD bei UNIMARC interesai. Matricą galima parsisiųsti iš projekto tinklalapio. Nors tikėtina, kad IFLA žodynai jų priežiūros grupių bus tiesiogiai ir tiksliai susieti rengiant jų RDF pateiktis, juos konsoliduojant ir pakartotinai nustatant sutaptis, VMF matrica galėtų būti labai naudinga IFLA žodynus derinant ir siejant su kitų sričių žodynais.

Susieti duomenys ir semantinis saitynas

Susieti duomenys – tai elementų triados, kuriomis konkrečioms entitetams atstovaujančių elementų savybėms priskiriamos konkrečios vertės. Triadų šaltiniai gali būti labai įvairūs; jas naudojant gali būti sutelkiamos duomenų apie tą patį entitetą (t. y. metaduomenis) sandaupos, pasitelkiant entiteto identifikatorius arba URI. Be to, RDF suteikia galimybę esamų triadų pagrindu daryti logines išvadas ir taip savo ruožtu sukurti naujas triadas.

Tradicinį katalogo įrašą sudaro skirtingų su bibliografiniu entitetu susijusių savybių reikšmės, pavyzdžiui, jo antraštė ir materialusis aprašas. Tokį įrašą įmanoma išskaidyti į elementų triadas, kurių visų subjektas būtų tas pats URI. Kiekvienos triados išraiškos „subjektas-predikatas-objektas“ predikato sandas turi būti savybę identifikuojantis URI, kaip reikalauja RDF. Objekto sandas atspindi savybės vertę. Tai gali būti tiesioginė vertė, pavyzdžiui, ženklų eilutė arba skaičius, ir tai atitiktų daugumą katalogo įrašų elementų. Objekto vertė gali būti ir kitas URI, pavyzdžiui, identifikuojantis SKOS pateikto kontroliuojamo žodyno terminą, todėl gali būti aprėpiami ir autoritetinių įrašų kontroliuojami kreipties elementai, jeigu autoritetinis failas išreikštas RDF.

Pateiksime triadų, į kurias gali būti išskaidytas konkretus įrašas, pavyzdį:

- ?:šisIšteklis ?:turi-antraštę „UNIMARC format for authority records“;
- ?:šisIšteklis ?:turi-autorių <viaf:29776655>;
- ?:šisIšteklis ?:turi-publikavimo-datą „2004“.

Svarbu tai, kad šios triados pateikiamos standartiniu formatu: URI išskaidytas į daugybinį vardų erdvės elementą ir po jo einantį specifinį identifikatoriaus elementą, kurie atskiriami dvitaškiu. Neapibrėžtai daugybinei vardų erdvei nurodyti vartojamas klausukas (?). URI, sudarytas iš konkretaus skaitinio specifinio identifikatoriaus, apskliaudžiamas kampiniais skliaustais. Tiesioginės vertės pateikiamos kabutėse.

To paties išteklius kitame kataloge esantis įrašas gali būti išskaidytas taip:

- ?:tasIšteklis ?:turi-antraštę „UNIMARC format for authority records“.
- ?:tasIšteklis ?:turi-leidėją „Howarth Press“.

Tačiau RDF leidžia abu šiuos išteklius išreikšti taip pat ir viena triada:

- ?:šisIšteklis owl:tasPatsKaipIr ?:tasIšteklis.

Šių trijų triadų pagrindu logiškai gali būti išvedama tokia triada:

- ?:šisIšteklis ?:turi-leidėją „Howarth Press“.

Jeigu iš naujo surinktume pirmojo katalogo įrašą, agreguodami visas triadas, kurių subjektas yra „:šisIšteklis“, jis turėtų leidėją identifikuojančią vertę, kurios nebuvo pirminiame įrašė.

Katalogų įrašų išskaidymas į RDF triadas bibliografinius metaduomenis neabejotinai daro išsamesnius. Šį poveikį dar labiau sustiprina šių triadų integravimas į semantinio saityno „katilą“ ir atitikimo tarp skirtingų to paties bibliografinio entiteto URI nustatymas: įtraukiamos kitų šaltinių – leidėjų, knygų pardavėjų, internetinių enciklopedijų, socialinių tinklų – triados iš naujo įvairiai agreguojant metaduomenis. Iš naujo atliekama agregavimo procesą galima užprogramuoti taip, kad būtų pašalinamos nepageidaujamos savybės ir jų vertės bei triados iš nepatikimų arba nereikšmingų šaltinių. Kai kurie tokio agregavimo aspektai gali būti pavaldūs ir galutiniam vartotojui, pavyzdžiui, pateikiamų savybių arba standartinio įrašo formato pasirinkimas.

Būtina įrašų konvertavimo į triadas išankstinė sąlyga – kiekvienos savybės URI išskyrimas. Savybių, kurios naudojamos tam tikro formato atributams ir santykiams apibrėžti, URI pateikiami RDF reiškiamuose metaduomenų modeliuose ir formatuose. Apdorojant pagal ISBD sudarytus įrašus, kiekvieną ISBD atitinkantį atributą būtina pateikti kaip RDF savybę su savu URI. Šia linkme dirba ISBD/XML tyrimų grupė, naudodama URI iš kurios nors su ISBD susijusios vardų erdvės.

Formatų RDF pateiktys naudingos ir kitais aspektais. Jos pateikia struktūrinį planą, padedantį išskaidyti esamus tam tikro formato įrašus ir iš naujo agreguoti elementų triadas, taip vėl sudarant to paties formato įrašus. Jos gali būti naudojamos ir klaidų kontrolei, siekiant išvengti trūkstamų ar neteisingai koduotų laukų.

Svarbu, kad kiekvienam formatui būtų priskirta sava vardų erdvė, taip išvengiant panašioms atributams, santykiams ir entitetams atstovaujančių savybių ir klasių painiavos. Tarkime, tokia galėtų būti triada, apimanti FRBR atributą „publikavimo arba platinimo data“:

- ?:šisIšteklis frbrer?:turi-publikavimo-arba-platinimo-datą „1973“.

RDF teikiamas šios savybės ribojimas leidžia kompiuteriui daryti loginę išvadą, kad:

- ?:šisIšteklis rdf:tipas frbrer:Apraiška.

Triadoje nurodoma, kad išteklius turi būti FRBR apraiška (o ne kūrinys, išraiška ar vienetas). Jeigu būtų panaudota panaši savybė iš kitos vardų erdvės, loginė išvada būtų kitokia arba jos nebūtų.

Daugeliui triadų su tokiais pat arba lygiaverčiais subjekto URI gali būti taikomas kitoks loginis išvedimas. Jeigu skirtinguose tam tikro entiteto katalogo įrašuose pateikiamos skirtingos vertės, galimi statistiniai metodai, nustatantys svarbą ar tendencingumą nuomonės apie tam tikro bibliografinio entiteto tam tikros savybės vertę. Tarkime, jei devynios iš dešimties skirtingų šaltinių triadų pateikia vertę „X“, o viena – vertę „Y“, galima daryti išvadą, kad tikroji vertė yra „X“, o „Y“ yra neteisingo interpretavimo, įvesties klaidos ar kitokių

katalogavimo taisyklių taikymo pasekmė. Panašiu metodu grindžiama eksperimentinė OCLC *Classify*³⁵ paslauga, konkrečiam kūriniui priskirianti Dewey dešimtainės klasifikacijos arba Kongreso bibliotekos klasifikacijos indeksus: pasitelkiant FRBR besiremiančią statistinę analizę, grupuojami to paties kūrinio įrašai, tik indeksai nustatomi ne pagal triadas, o pagal visą įrašą. Iš milijonų kataloguose esančių įrašų galima gauti nesuskaičiuojamą daugybę triadų, todėl panašu, kad statistinis loginis išvedimas gali būti naudingas daugeliu atvejų.

Kontroliuojami žodynai

Tokių terminijos sričių, kaip dalykinės rubrikos, klasifikacijų rubrikos ir indeksai bei tezaurai, elementų vertės gali būti pateikiamos RDF vartojant SKOS. SKOS suteikia galimybę išreikšti paprastus terminų santykius, pavyzdžiui, platesnius, siauresnius arba tikslus terminų atitikmenis; to paprastai pakanka, kad būtų perteikiama šių terminijos sričių vidinė struktūra. Vartojant SKOS, kiekvienai tam tikro termino reiškiamai sąvokai suteikiamas URI. SKOS leidžia vartoti tą patį URI terminą vėčiant į kitas kalbas, taip automatiškai užtikrinant funkcinį termino ir jo vertinių suderinamumą daugiakalbėje aplinkoje, o tos pačios sąvokos skirtingoms vardų erdvėms priklausantys URI RDF gali būti susiejami ir nustatomos jų sankirtos. SKOS reiškiami terminai gali būti skelbiami kaip susieti duomenys, taip suteikiant daugiau galimybių loginiam triadų išvedimui.

Kontroliuojamų žodynų, kaip susietų duomenų, viešas pateikimas skatina jais naudotis kitų sričių bendruomenes. Kokiam nors žodynui skirta infrastruktūra, tokia kaip tvarkymo ar galutiniam vartotojui teikiamų paslaugų sąsaja, gali būti pritaikoma ir kitam žodynui.

Kaip jau minėjome, nustatyti RDA turinio žodynai jau pateikti SKOS ir, tikėtina, bus paskelbti 2010 m. Šie žodynai nėra daugialgiai – jiems nebūdinga terminų hierarchija. Jie sudaryti iš anglišku terminų, tačiau kai kurie iš jų Vokietijos nacionalinės bibliotekos išversti į vokiečių kalbą išlaikant tą patį URI. Prieiga prie šių terminų nebus ribojama. Kiekvieno URI pagrindinis elementas rodys, kad terminas priklauso RDA, o ne kokiam nors kitam žodynui. Bendras termino ir jo vokiško atitikmens URI palengvins integruotą paiešką dvikalbėje – angliškoje ir vokiškoje – aplinkoje.

Per pastaruosius kelerius metus paskelbtos ir keletu kitų bibliografinių žodynų RDF bei jų, kaip susietų duomenų, pateiktys. Dauguma jų sudaryti dalykinių kreipties elementų ir klasifikacijos sistemų pagrindu.

Kongreso bibliotekos dalykines rubrikas (LCSH) kaip susietus duomenis galima parsisiųsti iš Kongreso bibliotekos autoritetinių įrašų ir žodynų tarnybos³⁶. Jie gali būti interaktyviai naudojami ir teikiant OCLC bandomąją

terminologijos paslaugą³⁷. Ši paslauga apima ir taikomąją programą *Faceted Application of Subject Terminology* (FAST), medicinos dalykines rubrikas *Medical Subject Headings* (MESH), grožinės literatūros formos ir žanro rubrikas bei grafinės medžiagos tezaurą *Thesaurus for Graphical Materials* (TGM).

Kaip susieti duomenys paskelbtos ir prancūzų RAMEAU dalykinės rubrikos³⁸. Kaip susieti duomenys taip pat prieinamos ir RAMEAU bei LCSH dalykinių rubrikų sankirtos, parengtos MACS projekto metu³⁹.

Veikia eksperimentinė OCLC *Dewey* dešimtainės klasifikacijos paslauga *Dewey.info*⁴⁰, kuria naudojantis rubrikai pagal jai priskirtą DDK indeksą pateikiami susieti duomenys. Šis indeksas yra kiekvienos temos URI unikalaus komponento sudedamoji dalis, o URI atitinka šabloną, pagal kurį, žinant DDK indeksą, galima logiškai išvesti temos URI. Kol kas paslauga apsiriboja DDK suvestinėmis, kurias sudaro pirmasis tūkstantis indeksų (t. y. indeksai iš 3 sveikų skaitmenų). Paslauga leidžia vietoj indeksų, kurie dėl savo neinformatyvumo paprastai nerodomi galutiniam vartotojui atliekant dalykinę paiešką, pateikti aiškias rubrikas. Susieti duomenys pajėgūs pasiūlyti patirtį, sukauptą užsakant bibliotekų fonduose saugomą fizinę medžiagą, ir taip perimti turtingų dalykinių metaduomenų vertę.

UDK konsorciumas kaip UDK suvestinę paskelbė elektroninį maždaug 2000 UDK klasių 16 kalbų rinkinį⁴¹. Konsorciumas ketina pateikti UDK suvestinę RDF ir kitais formatais bei jos sankirtas su žinių sisteminimo sistemomis⁴². Suvestinės turiniui suteikta *Creative Commons* licencija.

VIAF yra asmenų ir kolektyvų vardų kontroliuojamų žodynų (autoritetinių įrašų) rinkinys, tvarkomas įvairių šalių bibliografijos tarnybų. Vieno žodyno terminai susieti su atitinkamais kitų žodynų įvediniais, o visų terminų bei jų santykių aibė pateikiama kaip susieti duomenys⁴³.

ISBD nustato, kad įrašant 0 srities, skirtos turinio formai ir medijų tipui, duomenis turi būti pasitelkiami kontroliuojami žodynai. Visiems trimis šios srities elementams – turinio formai, turinio apibūdinimui ir medijų tipui – apibrėžti terminai⁴⁴ imami iš nustatytų sąrašų.

Metaduomenų valdymas

Tikėtina, kad bibliografinių metaduomenų, atitinkančių profesionalius standartus, buvimas semantinio saityno aplinkoje labai paveiks katalogavimo procesus ir metaduomenų valdymą.

Dėmesys bus sutelkiamas nebe į visą bibliografinį įrašą – išteklių apibūdinančių metaduomenų rinkinį, apibrėžiamą tam tikro formato ir katalogavimo taisyklių, – o tik į metaduomenų pateiktį. Pagal dabartinę praktiką katalogavimas atliekamas bibliografinio įrašo

lygmeniu. Prireikus padaryti kokį nors vieną pakeitimą, tarkime, pridėti katalogavimo taisyklių reikalaujamus metaduomenis, vietinius metaduomenis, tokius kaip dalykinės rubrikos, arba ištaisyti klaidą, manipuluojama visu įrašu. Įrašas paprastai įkeliamas į redagavimo sąsają, pataisoma viena jo dalis ir pakoreguotu įrašu pakeičiamas originalus įrašas.

Paskirstyto katalogavimo paslaugos ir konsorciškai grindžiami ne tiek įrašų sudedamųjų dalių, kiek įrašų, kaip visumos, sanaukėmis ir keitimusi. Viso to pasekmė gali būti didelis besidubliuojančių metaduomenų skaičius sanaukoje, nes sunku pakankamai tiksliai automatiškai identifikuoti to paties bibliografinio ištekliaus įrašus ir patenkinti dalyvaujančių organizacijų poreikius. Todėl paprastai pavyksta išvengti tik įrašų su tam tikram konsorciui būdingais arba tarptautiniais identifikatoriais, tokiais kaip tarptautinis standartinis knygos numeris, dubliavimosi. Taip pat sunku ir sinchronizuoti pataisų atlikimą, kad pataisymai, padaryti tam tikros sanaukos įrašuose, nebūtų perkeltami į jų vietines kopijas. Įrašų vietinės ir konsorcių versijos nebesutampa, nors jos apibūdina tą patį išteklių, ir beveik visi jų metaduomenys dubliuojasi. Kai dėmesys sutelkiamas į visą įrašą, neišvengiamai dubliuojami ir skirtingiems vienodomis savybėmis pasižymintiems ištekliams pridėti metaduomenys. Naujos knygos laidos autorius, antraštė, tema ir leidėjas paprastai nesiskiria nuo ankstesnės laidos. Skaitmeninto vaizdo turinys nesiskiria nuo originalo ir t. t. Dubliavimosi mastai yra tokie, kad naujam ištekliui metaduomenys paprastai pridedami nukopijuojant ankstesnio ištekliaus įrašą ir jį pakoreguojant, o ne sudarant naują įrašą.

Besidubliuojantys duomenys kliudo efektyviam saugojimui, prieigai, sklaidai ir apdorojimui. Jie gali sumažinti ir paslaugų, kurių pagrindą sudaro šie duomenys, efektyvumą. FRBR vartotojų tyrimas rodo, kad jų reikmės labiau patenkinamos grupuojant metaduomenis pagal jų atitiktį keturioms bibliografinių entitetų rūšims – kūriniui, išraiškai, apraiškai ir vienetai. Toks grupavimas leidžia išvengti vartotojui pateikiamų metaduomenų dubliavimosi, kaip yra taikant OCLC *FictionFinder* prototipą⁴⁵. Dėl šios priežasties FRBR modelio pagrindu veikiančiose informacijos paieškos sistemose nebereikia tvarkyti besidubliuojančių metaduomenų, ir daugiausia dėmesio jose skiriama įrašams, apibūdinantiems ne visą išteklių, o atskirus kūrinius, išraiškas, apraiškas ir vienetus. Naujai laidai nebereikia sudaryti naujo įrašo kūriniui, skaitmeniniam vaizdui neberekalingas naujas įrašas kūriniui ar išraiškai ir pan. Palyginus su bibliografiniu įrašu, katalogo įrašo struktūriškumo lygis išauga nuo vieno iki keturių kartų.

Semantinis saitynas labai padidina metaduomenų struktūriškumą. Kadangi kiekvieną katalogo apibūdinamą atributą ar santykį atitinka triada, bibliografinį įrašą gali sudaryti nuo kelių dešimčių iki kelių tūkstančių triadų. Todėl galima manyti, kad traktuojant triadą kaip „katalogo įrašą“ ir taip iki loginio minimumo sumažinant dubliavimąsi,

būtų galima efektyviau profesionaliai tvarkyti bibliografinius metaduomenis. Susietų duomenų aplinkoje esančių elektroninių katalogų bibliografiniam aprašui parengti gali būti panaudojami triadų, paimtų iš visų prieinamų šaltinių, duomenys. Tarkime, jei biblioteka gali pasinaudoti leidėjų sudaryta triada, jai nebereikia sudarinėti savos. Archyvų, bibliotekų ir muziejų bendruomenės visas pastangas gali skirti metaduomenų, kurie būtini tik jų reikmėms (kilmės, gaunamumo, tinkamumo, konteksto) tvarkymui, ir naudotis kitų bendruomenių sukurtais bendresniais metaduomenimis (žymėjimo, formato, susijusių vietų ir įvykių) kaip susietais duomenimis.

Susieti duomenys gali būti iš įvairių šaltinių ir vardų erdvių⁴⁶. Dėl skirtingų arba to paties šaltinio triadų dubliavimosi neišvengiamai pasitaikys semantiškai identišku triadų su tokiu pačiu URI. Pateikiant vartotojui išsamią, nuoseklią ir suprantamą informaciją, svarbiausiu uždaviniu taps tikslios duomenų tapatybės nustatymas. Labai svarbu ir susietų duomenų kilmė. Triados tapatumą tiksliai nurodyti gali tik kita triada. Susietuose duomenyse gali būti atsitikinių arba sąmoningai įveltų netikslumų, jų pagrindu gali būti padaroma prieštarinčių loginių išvadų – tai dažnai gali būti iš profesionalių šaltinių paveldėtų netikslų ir neišsamių metaduomenų buvimo pasekmė. Todėl svarbu žinoti triados šaltinį, jos sukūrimo datą ir kontekstą.

Naujoje aplinkoje plėtojamos tik jai skirtos metaduomenų valdymo sistemos. Archyvų, bibliotekų ir muziejų bendruomenių pardavimo sistemos seka pokyčius, tačiau joms tenka įveikti optimalaus kainos ir naudingumo santykio siekiančių klientų inertiškumą. Vis aktyviau diegiamos tradicinio katalogavimo infrastruktūros alternatyvos: atviroji programinė įranga, bibliografinės informacijos socialinių tinklų tarnybos, skatinančios vartotojus patiems pridėti ir koreguoti metaduomenis, pavyzdžiui, *LibraryThing*⁴⁷. Būtent tokios sistemos jau artimiausiu metu galės pasinaudoti susietų duomenų ir semantinio saityno pasiekimų teikiamomis galimybėmis.

Išvada

IFLA ir susijusios institucijos, užsiimdamos standartų plėtra ir juos taikydamos bei skleisdamos profesinius įgūdžius, atliks svarbų, netgi gyvybišką, vaidmenį semantinio saityno plėtroje. Dalyvavimas semantiniame saityne IFLA ir kitoms bendruomenėms leis tobulinti vartotojams teikiamas paslaugas ir produktyviau išnaudoti metaduomenų galimybes.

Iš anglų kalbos vertė T. Auškalis

Straipsnis parengtas pagal pranešimą, skaitytą 2010 m. Geteborge (Šveicarija) vykusioje 76-ojoje IFLA konferencijoje.

Visos nuorodos patikrintos 2010 m. gegužės 31 d.

- 1 Joint Steering Group for Development of RDA. RDA: resource description and access. Available at: <http://www.rda-jsc.org/rda.html>
- 2 Dublin Core Metadata Initiative. Available at: <http://dublincore.org/>
- 3 W3C. SKOS Simple knowledge organization system. Available at: <http://www.w3.org/2004/02/skos/>
- 4 British Library. Bibliographic Standards. Data Model Meeting. Available at: <http://www.bl.uk/bibliographic/meeting.html>
- 5 DCMIRDA Task Group wiki. Available at: <http://dublincore.org/dcmirdataskgroup/>
- 6 W3C. Resource Description Framework (RDF). Available at: <http://www.w3.org/RDF/>
- 7 IFLA Study Group on the Functional Requirements for Bibliographic Records. Functional requirements for bibliographic records. 1998, amended 2009. Available at: <http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records>
- 8 IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR). Functional requirements for authority data. 2009. Available at: <http://www.ifla.org/publications/functional-requirements-for-authority-data>
- 9 IFLA. FRBR Review Group. Meeting Report Durban, August 21, 2007. Available at: http://www.ifla.org/files/cataloguing/frbrrg/meeting_2007.pdf
- 10 IFLA. ISBD/XML Study Group. Available at: <http://www.ifla.org/en/node/1795>
- 11 NSDL Registry. Available at: <http://metadataregistry.org/>
- 12 Dunsire, G. Declaring FRBR entities and relationships in RDF. July 2008. Available at: <http://www.ifla.org/files/cataloguing/frbrrg/namespace-report.pdf>
- 13 NSDL Registry. Element sets: show detail for FRBRer model. Available at: <http://metadataregistry.org/schema/show/id/5.html>
- 14 Element set: FRBRer model. Available at: <http://metadataregistry.org/schema/show/id/5.rdf>
- 15 W3C. OWL Web Ontology Language: overview. Available at: <http://www.w3.org/TR/owl-features/>
- 16 International Working Group on FRBR and CIDOC CRM Harmonization. FRBR: object-oriented definition and mapping to FRBRer. Version 1.0. 2009. Available at: http://cidoc.ics.forth.gr/docs/frbr_oo/frbr_docs/FRBR00_V2.0_2009_june.pdf
- 17 International Council of Museums. The CIDOC conceptual reference model. Available at: <http://cidoc.ics.forth.gr/>
- 18 ICS-FORTH (ISL-ICS). CIDOC CRM v5.0.1 encoded in RDFS. 2009. Available at: http://cidoc.ics.forth.gr/rdfs/cidoc_crm_v5.0.1.rdfs
- 19 IFLA. ISBD Review Group. Worldwide review of ISBD. Available at: <http://www.ifla.org/en/news/worldwide-review-of-isbd>
- 20 International standard bibliographic description (ISBD). Consolidated edition. Draft as of 2010-05-10. IFLA. ISBD Review Group. International standard bibliographic description (ISBD). Consolidated edition. Draft as of 2010-05-10. Available at: http://www.ifla.org/files/cataloguing/isbd/isbd_wwr_20100510_clean.pdf, pp. VI-VII.
- 21 Escolano Rodríguez, Elena; Lynne Howarth; Mirna Willer; Boris Bosančić. News of ISBD. Project development of ISBD/XML schema: goals and objectives. Presented at World Library and Information Congress: 75th IFLA General Conference and Assembly, 23-27 August 2009, Milan, Italy. Available at: <http://www.ifla.org/files/hq/papers/ifla75/107-escolano-en.pdf>
- 22 Dunsire, Gordon. UNIMARC, RDA and the Semantic Web. Presented at World Library and Information Congress: 75th IFLA General Conference and Assembly, 23-27 August 2009, Milan, Italy. Available at: <http://www.ifla.org/files/hq/papers/ifla75/135-dunsire-en.pdf>
- 23 RDA (resource description and access) vocabularies. Available at: <http://metadataregistry.org/rdabrowse.htm>
- 24 Hillmann, Diane; Karen Coyle; Jon Phipps, Gordon Dunsire. RDA vocabularies: process, outcome, use. In: D-Lib magazine, vol. 16, no. 1/2 (January/February 2010). Available at: <http://www.dlib.org/dlib/january10/hillmann/01hillmann.html>
- 25 Dublin Core Metadata Initiative. Libraries Working Group. Library application profile. 2004. Available at: <http://dublincore.org/documents/library-application-profile/>
- 26 Malmsten, Martin. Exposing library data as linked data. Presented at the IFLA satellite preconference sponsored by the Information Technology Section "Emerging trends in technology: libraries between Web 2.0, semantic web and search technology", Florence, 19-20 August 2009. Available at: <http://www.ifla2009satelliteflorence.it/meeting3/program/assets/MartinMalmsten.pdf>
- 27 Styles, Rob; Danny Ayers; Nadeem Shabir. Semantic MARC, MARC21 and the Semantic Web. Presented at Linked Data on the Web (LDOW2008). 2008. Available at: <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-369/paper02.pdf>
- 28 Regole italiane di catalogazione REICAT / a cura della Commissione permanente per la revisione delle regole italiane di catalogazione; [la redazione del testo è stata curata da Alberto Petrucciani]. Roma: Istituto centrale per il catalogo unico delle biblioteche italiane e per le informazioni bibliografiche, 2009. Bozza complessiva, Gennaio 2009. Available at: http://www.iccu.sbn.it/upload/documenti/REICA_bozza_complessiva_genn2009.pdf
- 29 Tillett, Barbara B. and Ana Lupe Cristán. IFLA cataloguing principles: the statement of international cataloguing principles (ICP) and its glossary. München: K.G. Saur, 2009. Also available at: <http://www.ifla.org/en/publications/statement-of-international-cataloguing-principles>
- 30 Petrucciani, Alberto. Every reader his work, every work its title (& author): the new Italian cataloguing code REICAT. Presented at World Library and Information Congress: 75th IFLA General Conference and Assembly, 23-27 August 2009, Milan, Italy. 107 Cataloguing. Available at: <http://www.ifla.org/files/hq/papers/ifla75/107-petrucciani-en.pdf>
- 31 Commissione RICA. L'applicazione del modello FRBR ai cataloghi: problemi generali e di impiego normative. Available at: <http://www.iccu.sbn.it/upload/documenti/rica-frbr.pdf>
- 32 Vocabulary Mapping Framework. Available at: <http://cdlr.strath.ac.uk/vmf/index.htm>
- 33 RDA/ONIX framework for resource categorization. Version 1. 2006. Available at: <http://www.rdaajsc.org/docs/5chair10.pdf>
- 34 Dunsire, Gordon. Distinguishing content from carrier: the RDA/ONIX framework for resource categorization. In: D-Lib magazine, vol. 13, no. 1/2 (January/February 2007). Available at: <http://www.dlib.org/dlib/january07/dunsire/01dunsire.html>
- 35 OCLC. Classify: an experimental classification web service. Available at: <http://classify.oclc.org/classify2/>
- 36 Library of Congress. Authorities & vocabularies. Available at: <http://id.loc.gov/authorities/search/>
- 37 OCLC. Terminology services: experimental services for controlled vocabularies. Available at: <http://spilot.oclc.org/resources/index.html>
- 38 RAMEAU subject headings as SKOS linked data. Available at: <http://www.cs.vu.nl/STITCH/rameau/>
- 39 MACS Project. Available at: <https://macs.hoppie.nl/pub/>
- 40 Dewey Decimal Classification: summaries. Available at: <http://dewey.info/>
- 41 UDC summary. Available at: <http://www.udcc.org/udccsummary/php/index.php>
- 42 IFLA Classification and Indexing Section. Newsletter, no.40 (December 2009). Available at: http://www.ifla.org/files/classification-and-indexing/newsletters/ifla-newsletter-classification-40_rev.pdf
- 43 Hickey, Thom. VIAF as linked data. 2009. Available at: <http://outgoing.typepad.com/outgoing/2009/09/viaf-as-linked-data.html>
- 44 IFLA. ISBD Review Group. International standard bibliographic description (IS edition). Draft as of 2010-05-10. Available at: http://www.ifla.org/files/cataloguing/isbd/isbd_wwr_20100510_clean.pdf, pp. 0.1-1-0.3-2
- 45 OCLC FictionFinder. Available at: <http://fictionfinder.oclc.org/>
- 46 W3C. SWEO Community Project: Linking Open Data on the Semantic Web. Statistics on data sets. Available at: <http://esw.w3.org/TaskForces/CommunityProjects/LinkingOpenData/DataSets/Statistics>
- 47 LibraryThing. Available at: <http://www.librarything.com/>