

Elektroninių išteklių metaduomenų bazės kūrimas Rusijoje: problemos ir perspektyvos

Natalija KASPAROVA,
Michail ŠVARCMAN

Rusijos valstybinė biblioteka, 3/5 Vozdvizhenka street, Moscow, 101000 Russia, el. paštas:
kaspar@rsl.ru
shvar@rsl.ru

Elektroninių išteklių katalogavimo problema yra labai aktuali Rusijai. Čia ji yra ne mažiau sudėtinga nei bet kurioje kitoje pasaulio dalyje. Elektroninių išteklių bibliotekose nuolat gausėja, o juos kataloguojančių bibliotekų skaičius išlieka labai mažas. Ir tos, kurios kataloguoja elektroninius išteklius, naudoja tam įvairius bibliografinius formatus, įvairias taisykles ir skirtingas technologijas. Vienoje bibliotekoje mes netgi matėme interneto išteklių kortelių katalogą.

1998 m. Rusijos valstybinės bibliotekos specialistai parengė metodinį vadovą „Elektroninių išteklių bibliografinis aprašas“ ir atitinkamo nacionalinio standarto pirmąją laidą pagal tarptautinio standarto - ISBD(ER) - rekomendacijas.

Elektroninių išteklių bibliografinio aprašo taisyklių standartų sukūrimas negarantuoja, kad bus išspręstos elektroninių išteklių katalogavimo, bibliografinio įrašo duomenų unifikavimo problemos. Pagrindinė problema yra ta, kad aprašomame objekte būtinų duomenų nėra arba jie yra sunkiai prieinami. Be to, informacija, pateikta elektroniniame ištekliuje, skiriasi nuo informacijos, pateiktos reklamoje ar išankstiniuose pranešimuose. Sunkumų kyla ir dėl to, kad bibliografinę informaciją apie elektroninius išteklius skirtingai kataloguojant gali skirtingai interpretuoti.

Reikšmingiausias faktorius, padėjęs išspręsti šią problemą, yra interneto išteklių bibliografinio aprašo tarptautinis standartas „Dublin Core“, kuris įgalina pateikti minimalius duomenis ir sukurti internete informacinę terpę, kuri profesinėje literatūroje vadinama metaduomenimis. Graikų kalboje žodis „meta“ reiškia „tarpinis“ arba „tarpininkas“. Tai yra preliminarūs duomenys, sukurti kaip minimali informacija apie informacijos išteklių, šiuo atveju - elektroninių išteklių, atsirandantį jį sukūrus, bet prieš įtraukiant į katalogą.

1998-1999 metais Rusijos valstybinė biblioteka pradėjo vykdyti eksperimentus siekdama sukurti Rusijos išteklių internete metaduomenų bazę (adresas www.rsl.ru/dc). Šios bibliotekos serveryje veikia interaktyvi forma,

į kurią kiekvienas gali įvesti interneto išteklių aprašą ir gauti šio išteklių aprašą „Dublin Core Metadata Set“ formatu. Ta informacija patenka į duomenų bazę, kurioje sukaupta per 2000 elektroninių leidinių „Dublin Core Russian State Library“ (toliau - DC RSL) programinė įranga jau eksploatuojama, ji prieinama nemokamai.

Į DC RSL programinės įrangos paketą buvo įvestas „Dublin Core“ metaduomenų elementų paketas, įgalinantis kataloguojant automatiškai gauti metaduomenis.

Paketą sudaro šie komponentai:

⇒ kompiuterinės grafikos interfeiso programų rinkinys, veikiantis serveryje bei palaikantis duomenų įvedimą, paiešką ir rodymą ekrane;

⇒ papildomos paslaugų programos peržiūrai, verifikacijai ir duomenų bazės integralumo valdymui.

Programinė įranga Rusijos valstybinėje bibliotekoje buvo parengta naudojant „Inprise Delphi 5.0“ programą.

Metaduomenų vartojimas elektroniniams ištekliams aprašyti Rusijoje dar yra pradinės stadijos. Tam tikslui naudojami įvairūs formatai. Rusijos valstybinėje bibliotekoje elektroniniai ištekliai, kurie yra bibliotekos fondų dalis, kataloguojami USMARC formatu, tuo tarpu interneto išteklių katalogavimui taikomas „Dublin Core Metadata Set“.

Nevyriausybinių organizacijų (NGO) elektroninėje bibliotekoje (<http://ngo.org.ru/ngoss/>) knygų ir interneto išteklių aprašas sudaromas naudojant nežymiai papildytą „Dublin Core Metadata Set“ standartą. Pavyzdžiui, kūrėjų, leidėjų ir platintojų atsakomybės duomenims mes naudojame „Elektroninių verslo kortelių standartą“ (vCard). Paieškos rezultatai gali būti pateikti RTF ar XML formatais.

Šiuo metu Rusijos mokslų akademijos Gamtos mokslų bibliotekoje vykdomas didžiulis darbas, kurio tikslas - sukaupti duomenis apie Rusijos mokslininkus. Ši duomenų bazė apims informaciją apie leidinius, mokslinio tyrimo įstaigas, taip pat asmenų bibliografijas. Projekto grupė pagrindiniu formatu pasirinko „Dublin Core Metadata Set“ ir planuoja jį išplėsti į kai kurias

kitas sritis.

Rusijos kultūros paveldo tinklo programa, pagal kurią bus sukurtas muziejų tinklapis internete, apimantis muziejų eksponatų aprašus, taip pat numato kaip pagrindinį formatą naudoti „Dublin Core Metadata Set“.

Galima teigti, kad visuotinė tendencija tapo „Dublin Core Metadata Set“ formato naudojimas pagrindu aprašui rengti ir jį praplėsti specifinėmis užduotimis konkrečiose srityse. Todėl federalinė programa „Elektroninės bibliotekos“ numato vartoti keletą aprašo kalbų skirtinguose lygmenyse. Žemutiniuose lygmenyse naudojami labiau detalizuoti formatai, tokie kaip MARC, GILS arba kiti, priklausomai nuo kataloguojamų objektų.

Dabartiniu metu daugiausia elektroninių išteklių metaduomenų rengia bibliografijos tarnybos. Rusijos interneto išteklių tyrimas parodė, kad leidėjai etiketėje labai retai vartoja žymę „meta“ ir beveik niekada nesinaudoja „Dublin Core Metadata Set“ formatu. Informacija CD-ROM viršeliuose ir aprašų failuose dažniausiai taip pat labai menka. Todėl norint surasti visus bibliografiniam aprašui sudaryti reikalingus elementus reikia daug laiko ir pastangų.

Atsižvelgiant į spartų elektroninių išteklių gausėjimą ir itin ribotas kataloguotojų galimybes, atrodytų, kad tinkamiausia išeitis - pačių elektroninių išteklių kūrėjų dalyvavimas formuojant savo leidinių metaduomenis. Panaši praktika egzistuoja dabar, t.y. spausdintos produkcijos išankstinis katalogavimas. Tačiau šis uždavinys yra sudėtingas, o kai kurių specialistų nuomone - net utopinis.

Rusijoje pagrindinė institucija, atsakinga už elektroninių išteklių registraciją, siekiant apsaugoti autoriaus teises, yra „Informregistas“. Dabartiniu metu „Informregistro“ darbuotojai, atsakingi už elektroninių išteklių duomenų bazės tvarkymą, susiduria su visomis jau minėtomis problemomis.

Remdamiesi Rusijos valstybinės bibliotekos interneto išteklių katalogavimo patirtimi ir turėdami tam reikalui būtiną programinę įrangą, mes planuojame sukurti išankstinę elektroninių išteklių registracijos sistemą, vykdomą pačių kūrėjų. Bet kuris autorius gali pasinaudoti mūsų serveriu ir užpildyti siūlomos formos laukus. Kai jis įveda savo duomenis, gauna registracijos numerį ir elektroninio leidinio aprašą standartiniu formatu. Po to privalo užpildyti savo paraiškos „Informregistru“ formą, kurioje nurodo Rusijos valstybinės bibliotekos suteiktą elektroninio išteklių registracijos numerį. Šis numeris įgalins „Informregistro“ darbuotojus surasti elektroninio išteklių aprašą Rusijos valstybinės bibliotekos metaduomenų bazėje ir, jei reikia, pataisyti esamus arba pridėti papildomus laukus.

Mūsų ryšys su „Informregistru“ peraus į formalią bendradarbiavimo sutartį, kuri bus sudaryta kuriant elektroninių išteklių metaduomenų bazę. Turime išspręsti

daug organizacinių klausimų, kad pajėgtume įvykdyti šį planą. Svarbiausia problema yra priversti elektroninių išteklių kūrėjus: pirma - registruoti leidinius „Informregistre“, antra - naudoti išankstinę registracijos sistemą. Tam reikia sukurti patikimą teisinį mechanizmą. Formaliai „Informregistas“ saugo kiekvieno elektroninio išteklių autoriaus teises, ar jis būtų skelbiamas internete, ar CD-ROM, ar dar kitoje laikmenoje. Tačiau iš tikrųjų iki šiol „Informregistas“ registruodavo tik kompaktinius diskus. Produkcija internete dėl jos nuolatinės kaitos „Informregistru“ buvo nepasiekiamas. Šią problemą turėtų padėti spręsti Rusijos įstatymai, reguliuojantys interneto išteklių autorių teisių apsaugos klausimus.

Rusijos elektroninių išteklių metaduomenų bazė gali tapti pirmuoju elektroninių išteklių registracijos sistemos komponentu. Todėl ji negali būti pelno siekiantis objektas. Tačiau Rusijos valstybinei bibliotekai reikės papildomo finansavimo metaduomenų bazei išlaikyti ir plėtoti. Svarbiausi sistemos privalumai:

⇒ elektroninių išteklių bibliografinio aprašo unifikavimas ir katalogavimo išlaidų Rusijos katalogavimo tarnyboms ir bibliotekoms sumažinimas;

⇒ elektroninių išteklių informacinės aplinkos, atitinkančios tarptautinius standartus, sukūrimas, įgalinantis naudotis šia informacija ne tik Rusijos, bet ir užsienio vartotojus.

Rusijos valstybinė biblioteka pateikė nacionalinės metaduomenų bazės sukūrimo projektą, taikant „Dublin Core“ formatą. Šio projekto tikslas - unifikuoti elektroninių išteklių duomenis ir sudaryti galimybę katalogavimo tarnyboms, dalyvaujančioms tolesniame katalogavimo procese, pasinaudoti jais. Projektas apima ne tik interneto išteklius, bet ir tuos, prie kurių galima prieiti tikta vietoje.

Projektui įgyvendinti buvo imtasi šių priemonių:

1. Parengtas rusiškos „Dublin Core“ formato versijos projektas, pritaikytas Rusijos katalogavimo taisyklėms ir praktikai;
2. 8 lauke „Išteklių rūšys“ įvestas papildomas polaukis, skirtas vietinių elektroninių išteklių identifikavimui, jame pateikiami pagrindiniai šios sistemos reikalavimai ištekliui: kompiuterio pavadinimas, tipas ir/arba jo numeris, reikalingas atminties kiekis, operacinė sistema, programinė įranga ir periferiniai įrenginiai, taip pat techninės priemonės;
3. Išnagrinėtos problemos, iškylančios kuriant papildomus informacinius išteklius interneto metaduomenų terpėje, kurie padėtų elektroninių išteklių kūrėjams įvesti duomenis „Dublin Core“ formatu.

Šiandien kompiuterinėje laikmenoje galima naudotis Rusijos bibliotekinės ir bibliografinės klasifikacijos

duomenimis, dalykinių rubrikų rubrikynu ir autoritetinių įrašų bazė. Tęsdami elektroninių išteklių metaduomenų kūrimo procesą, turėsime sudaryti sutartis su abiem autoritetinių įrašų kūrėjais - Rusijos nacionaline biblioteka ir Rusijos knygų rūmais - dėl naudojimosi jų duomenimis sąlygų. Iki šiol ši produkcija internete dar nėra laisvai prieinama.

Sudėtingiausia šio projekto dalis - teisinio mechanizmo sukūrimas. Jis turėtų apibrėžti Rusijos elektroninių išteklių metaduomenų bazės statusą kaip naują technologinį mechanizmą valstybinės elektroninių išteklių registracijos sistemoje. Mums svarbi tarptautinė šios srities patirtis, kurią dabar analizuojame ir rengiamės taikyti.

Kaip jau minėta, yra ir kitų bibliografinių išteklių unifikavimo būdų, kurie nesiremia „Dublin Core“ formatu. Mūsų nuomone, pirmenybė turėtų būti teikiama būtent jam, nes tai palyginti nesudėtingas ir atitinkantis Rusijos bei tarptautinius bibliografinio aprašo reikalavimus formatus.

Mes tikimės IFLA ir kitų pasaulio bibliotekų specialistų paramos ir suvienytų pastangų toliau plėtoti bendrą informacinę erdvę ir naudoti naujas elektroninių išteklių galimybes.

Iš anglų k. vertė A. Staskevičius

UDK 004.915(474.5)
655.4/5(474.5)

Elektroninių išteklių leidyba ir ilgalaikis išsaugojimas

Regina VARNIENĖ

Bibliografijos ir knygotyros centras, Gedimino pr. 51, 2600 Vilnius, el. paštas: varniene@lnb.lt

Aušra VAŠKEVIČIENĖ

Lietuvos nacionalinė Martyno Mažvydo biblioteka, Gedimino pr. 51, 2600 Vilnius, el. paštas: auvask@lrs.lt

Įvadas

Šiandien daugelis didžiųjų pasaulio leidyklų savo periodinius leidinius, knygas, ypač mokslo darbus publikuoja ne tik spausdinta, bet ir elektronine forma. Nuo jų neatsilieka valstybės institucijos, įstaigos bei organizacijos, įvairius dokumentus pateikiančios savo interneto svetainėse arba kompaktiniuose diskuose. Elektroninė leidyba apibūdinama kaip medžiagos elektronine forma pateikimas, siekiant padaryti šią informaciją prieinamą visuomenei.

Šiuo metu nėra vieningos terminologijos elektroniniams dokumentams apibūdinti. Vartojami terminai: „elektroninis išteklius“, „elektroninė publikacija“, „elektroninis dokumentas“. Tačiau atsižvelgdami į ISBD(ER)¹ mes vadovausimės šio svarbaus dokumento, reglamentuojančio bibliografinio aprašo struktūrą, terminologija.

NEDLIB (Networked European Deposit Library) programos dokumentai skiria šiuos elektroninių išteklių tipus:

- autonominiai (*off-line*), t.y. fizinės elektroninės informacijos laikmenos, pavyzdžiui, CD-ROM, DVD, diskeliai, magnetinės juostos;
- hibridiniai, t.y. fizinės elektroninės informacijos laikmenos, tačiau tuo pat metu turinčios ryšį su internetu;
- neautonominiai (*on-line*), t.y. nuotoliniai elektroniniai ištekliai, esantys pasauliniame kompiuterių tinkle (internet) bei prieinami interaktyviu režimu.

Skiriama keletas neautonominių elektroninių išteklių rūšių:

- statiniai ištekliai, kurių turinys kinta per jų gyvavimo laikotarpį, tačiau jie išlaiko tam tikrą apibrėžtą išorinę struktūrą. Šie elektroniniai ištekliai yra artimiausi serialiniams leidiniams;
- dinaminiai ištekliai, kurių forma ir/arba turinys nuolat kinta per visą gyvavimo laikotarpį, pavyzdžiui, interneto duomenų bazės.

Leidinių kompaktiniuose diskuose bei diskeliuose įsiliejimas į bibliotekų fondus buvo gana neskausmingas.

Darbas su elektroniniais ištekliais, turinčiais fizinę formą, mažai kuo skyrėsi nuo darbo su įprastais leidiniais. Buvo pradėta jų bibliografinė apskaita. Internete esančios informacijos negalime paprasčiausiai „padėti“ į savo lentyną, todėl turime labai aiškiai apibrėžti darbo su tokia informacija principus.

Kuriant elektroninių išteklių archyvus, kuriems prigijo elektroninių bibliotekų pavadinimas, kyla nemažai organizacinių, ekonominių bei techninių problemų, kurias iš esmės emėsi spręsti daugelis tarptautinių susivienijimų bei programų: NEDLIB, Nacionalinių bibliotekų direktorių konsorciumas (CDNL), Europos nacionalinių bibliotekų konsorciumas (CENL), Kompiuterinių bibliografinių įrašų programa CoBRA ir kt. Daugelis metodinio ir organizacinio pobūdžio problemų buvo išspręsta, parengus atitinkamas rekomendacijas bei standartus vykdančias programas².

Kai kurios šalys, pirmosios pradėjusios spręsti elektroninių išteklių išsaugojimo bei bibliografinės apskaitos klausimus, daug prisidėjo prie svarbių projektų vykdymo (pavyzdžiui, Australijos nacionalinės bibliotekos projektas „Pandora Archive“³, Didžiosios Britanijos elektroninių bibliotekų programa „eLib“⁴, Kanados elektroninių bibliotekų iniciatyva⁵, Elektroninių bibliotekų federacija⁶). Šie pavyzdžiai ir objektyvi būtinybė išsaugoti elektroninius išteklis kaip svarbią Lietuvos kultūros paveldo dalį paskatino išanalizuoti šios dokumentų rūšies padėtį mūsų šalyje ir parengti analogišką projektą elektroniniams ištekliais išsaugoti.

Elektroninės leidybos raida ir dabartinė būklė Lietuvoje

S.E. Arnoldas ir E.S. Arnoldas išskiria tris elektroninių išteklių raidos etapus⁷. 1977-1987 metai laikomi elektroninės informacijos eros pradžia, tuomet pirmaisiais elektroniniais leidiniais pradėjo naudotis akademinė visuomenė. Tuo laikotarpiu elektroninės informacijos vartojimas augo lėtai ir buvo labai ribotas, ja galėjo naudotis tik mokslininkai ir inžinieriai.