

## VIRTUALI SKAITMENINIO PAVELDO SISTEMA - ATMINTIES INSTITUCIJŲ SĄVEIKOS PAGRINDAS

DR. REGINA VARNIENĖ

Lietuvos nacionalinė Martyno Mažvydo biblioteka, Vilnius, Lietuva,

el. paštas: [r.varniene@lnb.lt](mailto:r.varniene@lnb.lt)

GIEDRIUS DAUGIALA

UAB „Sintagma“, Vilnius, Lietuva, el. paštas: [giedrius.daugiala@sintagma.lt](mailto:giedrius.daugiala@sintagma.lt)

*Europos Komisijos komunikatais „Europos informacinė visuomenė augimui ir užimtumui skatinti“, „2010: Europos skaitmeninė biblioteka“, „Dėl mokslinės informacijos skaitmeniniame amžiuje prieinamumo, sklaidos ir išsaugojimo“ Europos Komisija skelbia naują integruotą požiūrį į ilgalaikį kultūros ir mokslo paveldo išsaugojimą. Komunikatuose numatomos kultūrinio ir mokslinio paveldo skaitmeninio, virtualios prieigos bei ilgalaikio išsaugojimo perspektyvos pabrėžiant skaitmeninių bibliotekų iniciatyvą ir vaidmenį. Paminėti dokumentai skatina atminties institucijas atsižvelgti į naują „atminties institucijų sąveikos“ paradigmą, kurios mokslinei analizei ir glaudesnio bendradarbiavimo perspektyvai kuriant skaitmeninį kultūros paveldą buvo skirti svarbiausi šios srities tarptautiniai projektai: Minerva, Minerva Plus, Digicult, Michael ir kiti. Skaitmeninė konvergencija reikalauja daryti atitinkamus politinius sprendimus. Kalbant apie Lietuvos atminties institucijų bendradarbiavimą kuriant skaitmeninio turinio kultūros paveldą, galima būtų tvirtinti, kad 2004–2005 m. LNB skaitmeninio kultūros paveldo turinio iniciatyvos integravo kultūros politiką, metodologiją ir praktiką. Nors projektas dar nėra baigtas, tačiau jau dabar galima kalbėti apie inovacijas, darančias didelę įtaką:*

1) Kultūros politikai - nes to reikia pagal projektą, kad būtų parengta kultūros paveldo skaitmeninio strategija, numatanti atminties institucijų bendradarbiavimo būdą, jų atsakomybę bei tolesnės sąveikos perspektyvą.

2) Metodologijai ir technologijai - nes projekte dalyvaujančios atminties institucijos susitarė dėl bendrų metaduomenų, failų valdymo, klasifikavimo standartų, kurie garantuoja ilgalaikį skaitmeninės informacijos išsaugojimą ir efektyvią prieigą. Virtualios sistemos ir bendro [„epaveldas.lt“](http://epaveldas.lt) portalo sukūrimas patvirtino Kultūros paveldo skaitmeninio strategijoje ir projekte užfiksuoto metodologinio sprendimo teisingumą.

3) Lietuvos integracijai į skaitmeninį Europos turinį, nes pabaigus projektą bus skaitmeninta per 3 mln. 605 tūkst. puslapių originalių rankraščių, laikraščių senų knygų, metrikų, metraščių, taip pat kitų itin vertingų, didelę istorinę vertę turinčių dokumentų. Šioje publikacijoje apžvelgiami pagrindiniai metodologiniai sprendimai, užfiksuoti Virtualios paveldo sistemos lingvistiniame aprūpinime arba suderintoje standartų visuomenėje, kurie leido įgyvendinti daugelį siekiamų tikslų formuojant sistemos architektūrą, organizuojant skaitmeninių objektų archyvavimą, jų erdvės tvarkymą, prieigą bei suderinamumą su projekto partnerių sistemomis: LIBIS, LAFS, RIS.

**Reikšminiai žodžiai:** atminties institucijos, kultūros paveldas, informacijos sistemos, Lietuva

### Įvadas

Pasaulį apimant visuotinumui, tautos ir tautiškumo, tautinio tapatumo išsaugojimo problemos tampa itin aktualios. Kultūros paveldo, kaip gyvosios kultūros vertybių sistemos, padedančios išlaikyti tautinį tapatumą ir kartu atveriančios galimybes visavertei kultūros raiškai pasaulio tautų bendrijoje, puoselėjimas ir šios sistemos vertybių sklaida tapo vienu svarbiausių Europos Sąjungos kultūros politikos prioritetų. Tai patvirtina pastarųjų metų Europos Komisijos komunikatai [1-4] ir programos, skatinančios kultūros ir mokslo paveldo skaitmeninimą ir jo integravimą į Europos skaitmeninį turinį. Tai naujas žingsnis stiprinant besiplečiantį Europos šalių sambūrį, remiantis natūralia kiekvienos tautos egzistencine teise, išreiškia kalba, papročiais, tradicijomis, jos menais ir pagaliau religija. Tai kiekvienos tautos teisė, turinti *lex naturale* statusą. Tik griežtai ją puoselėjant ir gerbiant yra galima darni tautų sandrauga - civilizuotų tautų ir jų kultūrų sambūvio norma. Kaip tik šia kryptimi formuojama naujoji ES kultūros politika, pripažįstanti kiekvienos sandraugos šalies unikalumą, išskirtinumą bei reikšmę. Įvairovė kuria ne tik trintis ir įtampas, ji kuria dar ir žavesį, žmogaus išmintį. Tuo tarpu kultūrinė unifikacija bei niveliacija skatina pasaulio bendruomenės sąstingį. Svarbiausiose ES politiniuose dokumentuose ne tik akcentuojama kultūros paveldo - kaip tautos identiteto - reikšmė, bet taip pat nurodoma, kaip šis paveldas ir kultūros žinios turi būti pateikiami viešai prieigai be jokių laiko ir vietos apribojimų. Drauge išvardijama, kad kultūros paveldo bei žinių pervedimas į skaitmeninę dimensiją yra logiškas žingsnis, sudarantis sąlygas be kliūčių jais naudotis ir pažinti kitas tautas.

Kalbant apie Lietuvos atminties institucijų pastangas kuo plačiau atverti jose saugomus kultūros lobius juos skaitmeninant, galima tvirtinti, kad, iki pradėdant Lietuvos nacionalinei Martyno Mažvydo bibliotekai drauge su partneriais - Lietuvos dailės muziejumi ir Lietuvos archyvų departamentu - vykdyti projektą, atminties institucijų veikla šalyje nebuvo koordinuojama, tokio pobūdžio darbai buvo atsitiktinio pobūdžio.

Publikacijoje pristatomas projektas yra pirmasis tokio plataus masto ir tokio pobūdžio kultūros projektas, suvienijęs skirtingų žinybų kultūros įstaigas - archyvus, bibliotekas ir muziejus vienam bendram tikslui - kuo plačiau atverti kultūros paveldą kiekvienam mūsų šalies piliečiui ir pasauliui. Nors projektas dar ir nėra baigtas, jau šiandien galima kalbėti apie tris pagrindines inovacijas, turinčias didelės įtakos kultūros politikai, nes skaitmeninė konvergencija pareikalavo daryti politinius sprendimus. Pradedant tokio masto projektą turėjo būti parengta šalies kultūros paveldo skaitmeninimo koncepcija, numatanti atminties institucijų bendradarbiavimo būdus, jų atsakomybę bei tolesnės sąveikos perspektyvą. 2005 m. buvo parengta ir Lietuvos Vyriausybės patvirtinta Lietuvos kultūros paveldo skaitmeninimo koncepcija\*.

Kita taip pat svarbi inovacija, apie kurią jau šiandien galima kalbėti - projekto pagrindu susiformavo atminties institucijų sąveikos metodologija, nes buvo susitarta

<sup>1</sup> Dėl Lietuvos kultūros paveldo koncepcijos patvirtinimo: patvirtinta Lietuvos Respublikos Vyriausybės 2005 m. rugpjūčio 25 d. nutarimu Nr. 933

dėl skaitmeninius objektus aprašančių metaduomenų skaitmeninių objektų archyvavimo ir jų saugojimo bei prieigos standartų. Kartu sukurta skaitmeninio paveldo sistema ir bendras portalas patvirtino Lietuvos kultūros paveldo skaitmeninio koncepcijoje užfiksuotų metodologinių sprendimų tikslumą.

Ir pagaliau trečioji inovacija - turtingas ir įvairus kultūros paveldo skaitmeninis turinys, kuris bus sukurtas užbaigus projektą, sudarys galimybes visavertei mūsų kultūros integracijai į Europos skaitmeninę erdvę, nes bus skaitmeninta daug itin didelę vertę turinčių senųjų knygų, laikraščių, dailės kūrinių rankraščių, bažnyčių metrikų. Skaitmeninis turinys ir suformuota efektyvi virtuali skaitmeninio paveldo sistema su bendru portalu - epaveldas.lt sudarys galimybę teikti aukštos kokybės informacines paslaugas - nebereikės atokiau gyvenantiems mūsų šalies piliečiams gaišti daug laiko kelionėse norint gauti reikiamos informacijos iš archyvų ar bibliotekų. Pagaliau plačiau atsivers ir muziejų fondai ir juose esantys eksponatai, kurie dėl vienokių ar kitokių priežasčių buvo sunkiai prieinami lankytojams.

Lietuvos kultūros paveldas, integruotas į Europos skaitmeninį turinį, prisidės prie visavertės mūsų šalies kultūros raiškos Europos kultūrų bendrijoje bei mūsų šalies įvaizdžio gerinimo.

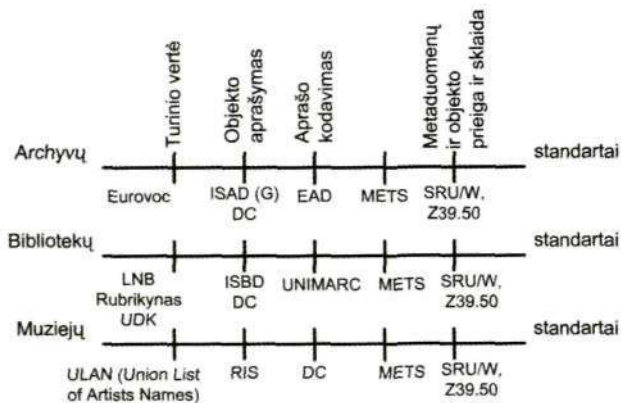
Be visų išvardytų skaitmeninio turinio sukūrimo pranašumų, dar tektų pridurti tai, kad jis prisidės prie pagrindinės atminties institucijų funkcijos gerinimo - kultūros paveldas bus patikimiau saugomas ateities kartoms, nes plačiau visuomenei jis bus prieinamas internetu - epaveldas.lt

Šioje publikacijoje daugiausia dėmesio skiriama metodologiniams ir technologiniams sprendiniams, užfiksuotiems Virtualios skaitmeninio paveldo sistemos lingvistiniame aprūpinime, kuris leido sukurti efektyvią architektūrą, užtikrinti patikimą skaitmeninių objektų archyvavimą, efektyvią sklaidą ir suderinamumą su veikiančiomis partnerių sistemomis: LIBIS, LAFS, RIS.

### **Virtualios paveldo sistemos lingvistinio aprūpinimo ypatumai ir funkcinė struktūra**

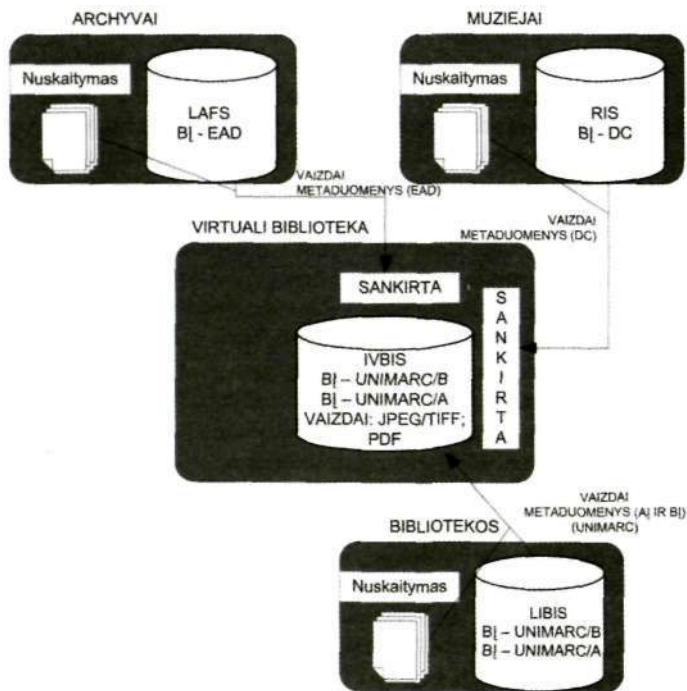
Sukurtos sistemos lingvistinis aprūpinimas buvo suformuotas remiantis Lietuvos ir pasaulio šalių elektroninių archyvų, katalogų bei juos apibūdinančių dokumentų analize. 1 paveiksle pateikiama Virtualios paveldo sistemos lingvistinis aprūpinimas.

Kaip matome iš schemas, Virtualios skaitmeninio paveldo sistemos lingvistinį aprūpinimą sudaro skaitmeninių objektų turinio vertę apibūdinantys standartai (archyvuose - Eurovoc [5], LNB rubrikinas ir UDK [6-7] - bibliotekose, muziejuose - Union List of Artist Names [8]), tarptautiniai bibliografinio ir archyvinio aprašo standartai, skirti objektų aprašymui archyvuose ir bibliotekose bei Lietuvos muziejų aprašo taisyklės (RIS), aprašo kodavimo formatai - archyvuose - koduotasis archyvinis aprašymas (EAD) [9], bibliotekose - UNIMARC [10], metaduomenų ir objekto archyvavimo standartas METS [11] bei prieigos protokolai - SRU/W [12] ir Z39.50 [13].



1 pav. Virtualios skaitmeninio paveldo sistemos lingvistiniai standartai

Suformuoto lingvistinio aprūpinimo pagrindu buvo formuojama virtualios skaitmeninio paveldo sistemos architektūra, kuri pateikiama 2 paveiksle.



2 pav. IVBIS architektūra

Suformuota architektūra užtikrina:

spartų duomenų perdavimą atliekant atskiras užklausas bei operuojant dideliais duomenų kiekiais;

galimybę palaikyti daugialypės terpės dokumentus, sudarytus iš teksto ir atvaizdų;

paskirstyto pradinio duomenų apdorojimo galimybę, jei tokia yra, projekto partneriai gali patys parengti skaitmeninius fondus ir talpinti juos į bendrąją duomenų bazę.

- archyve saugomiems formaliems dokumentams suteikti nuolatinis vardus.

Lingvistinis aprūpinimas taip pat užtikrina Virtualios sistemos archyvavimą ir objektų erdvės tvarkymą:

integruojant skaitmeninius objektus šiais formatais: TIFF, JPEG, PDF.

formuojant skaitmeninių objektų erdvę taip, kad ją galima kuo efektyviau pateikti paieškai ir skaitmeninių objektų pateikimui.

- užtikrinant skirtingų metaduomenų ir formatų suderinamumą, parengus metaduomenų kodavimo sistemų sankirtas.

- IVBIS Lingvistinis aprūpinimas užtikrina monitoringo realizavimą.

• IVBIS Lingvistinis aprūpinimas užtikrina prieigą prie skaitmeninių objektų ir sklaidą bei suderinimą su Lietuvos ir Europos skaitmeniniu turiniu;

• sparčią prieigą prie skaitmeninių objektų, saugomų LNB pagrindiniame skaitmeninės informacijos archyve bei išsamių bibliografinių duomenų projekto partnerių sistemose;

- informacijos pateikimą internete per interneto sąrašą;

patogią prieigą prie dokumentų pagal įvairius jų požymius;

• paieškos rezultatų išdavimą pagal sąrašą, kuriame turi būti nurodomi užklausa atitinkantys dokumentai ar jų dalys.

Virtualios sistemos lingvistinis aprūpinimas užtikrina šių funkcijų vykdymą:

- komplektavimą, apimančią visus veiksmus, susijusius su skaitmeninių objektų surinkimu ir įrašymu į bendrą serverį;

registravimą: naujai gautas skaitmeninis objektas yra užregistruojamas sistemoje ir bibliografinė informacija pateikiama į elektroninius katalogus;

patikrinimą: kurio metu atliekamas failų formato bei skaitmeninio objekto loginio integralumo patikrinimas;

katalogavimą: skaitmeninio objekto kompiuterinio įrašo sudarymo procedūras siekiant užtikrinti, kad jis būtų surastas šalies ir pasaulio bibliotekų, archyvų ir muziejų kataloguose. Šis procesas apima tokias procedūras kaip pirminių autorius aprašomųjų metaduomenų panaudojimą bibliografiniam aprašui sudaryti bei jų papildymą, pavyzdžiui, autoritetinio įrašo sudarymą, skyrimą pagal dalyko sritį ir t.t. Katalogavimo procesai atliekami Nacionalinėje bibliotekoje bei projekto partnerių institucijose;

saugojimo tvarkymą: čia atliekamos tokios procedūros kaip išsaugojimas skaitmeninėse laikmenose, nuolatinis atsarginių kopijų darymas bei kopijavimas įvairiems tikslams. Čia ypač svarbūs techniniai metaduomenys, fiksuojantys informaciją apie skaitmeninio objekto migravimą archyve;

- pakavimą ir pateikimą: užtikrina skaitmeninio objekto prieinamumą, šios

funkcijos naudojamos statistikai, monitoringui ir kitiems tikslams. Jų paskirtis yra skaitmeninio objekto paieška, reikalingos skaitmeninio objekto kopijos paieška ir perdavimas į kreipties modulį. Šis procesas gali apimti tokias procedūras kaip reikalingos skaitmeninio objekto dalies radimas, kopijavimas arba viso teksto pridėjimas, tam tikras skaitmeninio objekto dalies reformatavimas peržiūrai, atspausdinimas;

- kreiptį: užtikrina prieigą prie archyvo skaitmeninių objektų. Šis procesas apima suradimo galimybių suformavimą, vartotojo identifikavimą, vartotojo teisių bei profilio nustatymą ir kt.;
- monitoringą: visi išvardyti procesai turi būti susieti į bendrą visumą ir vykdoma jų kokybės kontrolė;
- integravimą: į bendrą Lietuvos ir Europos elektroninio turinio erdvę.

### **Aprašomųjų metaduomenų panaudojimo sprendimai**

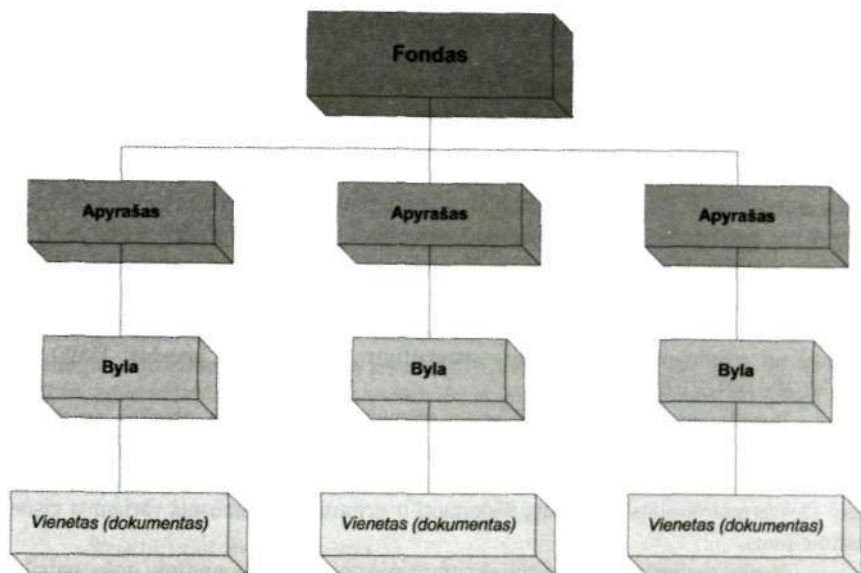
Metaduomenys yra struktūriškai apibrėžta informacija, kuri apibūdina, paaiškina informacinį išteklių, nurodo jo buvimo vietą arba kitokiu būdu palengvina jo suradimą naudojimą arba valdymą. Jie paprastai apima pagrindinę bibliografinę informaciją tokią kaip kūrėjas, sukūrimo data, santrauka, kataloginę, teminę informaciją pavyzdžiui, reikšminiai žodžiai.

Kurdami virtualią skaitmeninio paveldo sistemą pirmiausiai turėjome apsispręsti dėl skaitmeninių objektų aprašo išsamumo lygio: ar skaitmeninio objekto aprašymui naudoti Dublin Core [14] metaduomenų standartą ar rengti išsamius aprašus. Pasirinkus antrąjį būdą vienoje sistemoje reikėjo suderinti objektų aprašymo standartus ir jų kodavimo sistemas.

Iliustracijai pateiksime archyvinio aprašo standartų analizę ir priimtus sprendimus. Bendrame tarptautiniame archyviniame aprašyme geriau žinomą santrumpa - ISAD (G) [15] - yra nurodoma, kaip formuluoti informaciją kiekvienam iš 26 elementų, kuriuos galima derinti taip, kad jie sudarytų archyvinio objekto aprašą. ISAD (G) garantuoja išsamių, tinkamų ir aiškių aprašų sudarymą palengvina informacijos apie archyvinę medžiagą sudarymą mainus ir paiešką leidžia sujungti aprašus iš įvairių saugyklų į bendrą informacijos sistemą.

Tuo atveju, kai fizinis archyvo objektas yra skaitmenintas, būtina parengti prieigą prie jo turinio. Todėl iškyla būtinybė aprašyti skaitmeninio objekto vaizdą arba jo turinį. Jei objektas yra fondas, reikia sudaryti viso fondo archyvinį aprašą su jo hierarchine struktūra, atskleidžiančia jo struktūrinę dalis.

Kuriant virtualią skaitmeninio paveldo sistemą buvo laikomasi nuomonės, kad skaitmeninių objektų aprašymas privalo turėti tą pačią hierarchinę struktūrą kaip ir tradicinių, tokią kuri pateikiama 3 paveiksle.



3 pav. Archyvinio aprašo hierarchinė struktūra

Įrašai turi būti sudaromi pagal Bendrąjį tarptautinį archyvinį aprašymą - ISAD (G), kuris sudaro galimybę parengti archyvinio aprašo hierarchinę struktūrą. Kitaip tariant, išsamesnį ir patogesnį archyvų objektų bei duomenų bazių turinio aprašymą elektroninėje aplinkoje. Šis metodologinis sprendimas buvo priimtas kuriant virtualią skaitmeninio paveldo sistemą. Kitaip tariant, nebuvo atsisakyta išsamaus aprašo kūrimo skaitmeniniams objektams. Už išsamių bibliografinių, archyvinių, muziejų skaitmeninių objektų įrašų parengimą yra atsakingos projekte dalyvaujančios institucijos.

Priimto sprendimo teisingumą patvirtina dar ir tai, kad skaitmeninių objektų metaduomenų schemas yra tik bibliografiniai modeliai, tinkantys aprašomųjų metaduomenų prieigai internete. Skaitmeninių objektų saugojimui ir archyvavimui jie yra pernelyg abstraktūs. Štai, pavyzdžiui, Dublin Core formato, kuris Lietuvoje yra pripažintas šalies standartu, kai kurie elementai gali sukelti netgi tam tikrą sumaištį, pavyzdžiui, „informacija apie kūrėją“. Archyvuose kiekvieno archyvinio vieneto aprašas yra reikšmingas tik tada, kai yra aiškiai aprašytas fondo, kuriam jis priklauso, kūrėjas. Tik tai kūrėjo istorija ir struktūra gali padėti vartotojui suprasti sąlygas, kuriomis dokumentas buvo sukurtas arba surinktas, kritiškai įvertinti dokumentų turinį. Šio tipo duomenys yra labai svarbūs aprašui. Būtent tokių duomenų ir trūksta Dublin Core. Dublin Core elementas vadinamas „kūrėju“ yra apibūdinamas kaip už entitetų turinį pirminės atsakomybės sritis. Archyvams toks apibūdinimas nėra tinkamas. ISAD (G) kūrėją apibūdina, kaip: „Kolektyvas, šeima arba asmuo sukūręs, surinkęs arba turintis dokumentus savo personalinės arba kolektyvinės veiklos metu.“ Asmenys, atsakingi už dokumentų perkėlimą į fondus, dažniausiai skiriasi nuo kūrėjo. Todėl, atsižvelgiant į „kūrėjo“ apibūdinimą, pateikiamą DC, yra tam tikras prieštaravimas. Jeigu naudojamas

DC apibrėžimas, neišmanoma pažymėti arba nurodyti, kas surinko ir yra šio dokumento / archyvo savininkas.

Nors virtualioje sistemoje buvo pasirinktos turtingesnės DC versijos, tačiau, kaip jau minėta, nebuvo atsisakyta ir išsamių aprašų rengimo siekiant kuo tiksliau apibūdinti skaitmeninį objektą. Tokios pat pozicijos yra laikomasi ir dėl Nacionalinės bibliotekos bei Dailės muziejaus skaitmeninių objektų įrašų rengimo. Muziejinėje praktikoje ir tradicinio, ir skaitmeninio objektų aprašymui naudojama 40 elementų duomenų instrukcija, patvirtinta 1997 m. Lietuvos Respublikos kultūros ministerijos įsakymu. Joje nurodoma, kokie duomenys, susiję su eksponatu, turi būti pildomi privalomai ir kokius muziejus gali pildyti pasirinktinai.

Nacionalinės bibliotekos skaitmeninių objektų išsamių bibliografinių įrašų parengimui naudojamas Tarptautinis standartinis bibliografinis aprašas: ISBD (M), ISBD(CR), ISBD(ER) [16-18].

Portale epaveldas.lt pateikiamas trumpas skaitmeninio objekto aprašas su pagrindiniais bibliografinio aprašo elementais (pvz., dokumento pavadinimas, kalba, skelbimo / leidimo duomenys, apimtis, data ir kt., taip pat skaitmeninio dokumento vaizdas). Dėl terminologijos teko daug diskutuoti ir priimti tinkamiausius variantus visoms institucijoms.

Norint susipažinti su išsamiu skaitmeninio objekto bibliografiniu įrašu, pateikiama nuoroda į pirminę sistemą - LNB elektroninį katalogą arba Nacionalinės bibliografijos duomenų banką. Čia įrašą galima peržiūrėti OPAC/WWW formatu - trumpą arba išsamų, taip pat UNIMARC formatu. UNIMARC įrašo 856 lauke pateikiamas pastovus skaitmeninio dokumento elektroninis adresas, kuris leidžia grįžti į portalą.

Taip realizuojamas abipusis ryšys - iš portalo į pirminę sistemą ir iš pirminės sistemos - atgal į portalą.

Siekiant parengtus įrašus pateikti į bendrą portalą, šie duomenys turi būti koduojami. { virtualią paveldo sistemą pateikiamų archyvų skaitmeninių objektų aprašomųjų metaduomenų kodavimui naudojamas koduotasis archyvinis aprašymas - EAD (Encoded Archival Description), skirtas muziejams, bibliotekoms ar saugykloms, kurių fonduose yra archyvinių dokumentų. Lietuvos nacionalinės bibliotekos skaitmeninių objektų bibliografinių įrašų kodavimui naudojamos UNIMARC/B, kuris Lietuvos bibliotekose kultūros ministro įsakymu yra įteisintas bibliografinių įrašų kodavimo standartu.

Perkeliant Dailės muziejaus eksponatų skaitmeninius vaizdus į virtualią paveldo sistemą, naudojamas DC metaduomenų formatas. Skirtingų metaduomenų aprašymo schemų bei šių duomenų kodavimo sistemų suderinamumo problema virtualioje paveldo sistemoje sprendžiama naudojant sankirtas, kurios leidžia derinti skirtingai aprašytus objektus, skirtingomis schemomis ir skirtingais kodavimo formatais.

Pavyzdyje pateikiama aprašomųjų metaduomenų sankirta, nurodanti skaitmeninių objektų, esančių originaliose sistemose, aprašų turinį bei kodavimą (1 lentelė).



I lentelė. Sankirtų panaudojimas

ISAD (G)	ISBD	RIS
3.2.1. Sukūrėjo vardas	1.5. Atsakomybės duomenys	Autorius / gamintojas
EAD	UNIMARC	Dublin COre
<origination>	7--Atsakomybės blokas	<dc:creator>

Sukurtoje virtualioje skaitmeninio paveldo sistemoje yra naudojamos keletas parengtą sankirtų, leidžiančių vienareikšmiškai atvaizduoti aprašomuosius ir kitus reikalingus elementus perduodant iš partnerių sistemų į bendrą duomenų archyvą. Tokių sankirtų, labai svarbių metaduomenų valdymo elementų, tobulinimui ir priežiūrai reikia daug pastangų.

## Metaduomenų ir skaitmeninių objektų archyvavimo sprendimai

### *Struktūriniai metaduomenys*

Ekspresyvūs formatai, kurie leidžia pateikti skaitmeninius objektus iš skirtingų bendruomenių yra vadinamieji kompleksiniai objektų formatai arba mūsų dokumente - struktūriniai. Šiems formatams yra būdinga:

skaitmeninių objektų pateikimas XML išraiška;

galimybė pateikti paprastus bei sudėtingus - kompleksinius skaitmeninius objektus;

galimybė išreikšti atskirų skaitmeninių objektų arba grupės skaitmeninių objektų identifikatorius;

galimybė grupuoti duomenis;

jungti skaitmeninius objektus su aprašomaisiais, techniniais arba kitais metaduomenimis.

Projekto vykdymo metu buvo sukurta darbo sekų programinė įranga, kuri leidžia sujungti skaitmenizuotus kūrinius į bendrus paketus, kur galima organizuoti kokybės kontrolę, gražinimą pakartotiniam skaitmenizavimui netinkamų / nekokybiškų vaizdų, susiejimą su atpažintu visateksčiu failu ir susiejimą su bibliografiniu objektu iš atitinkamo katalogo duomenų bazės. Sujungus aprašomuosius metaduomenis ir objektus, šie duomenys gali būti perduodami įkelti į virtualios paveldo sistemos centrinę duomenų bazę. Duomenų perdavimui yra naudojamas Metaduomenų kodavimo ir perdavimo standartas (METS): <<http://www.loc.gov/standards/mets>>. Eksportuojant duomenis į METS standarto paketą, yra įtraukiama visa su tuo kūriniu susijusi informacija ir nuorodos į išorinius objektus - skaitmeninių vaizdų failus. Skaitmeninio objekto aprašas METS formatu yra atskirų dokumento dalių ir formų jungiamoji grandis. Sistemoje yra naudojamos pirmos 5 METS standarto sekcijos:

1. Bendroji dalis (METS Header).
2. Aprašomoji dalis (Descriptive Metadata).

3. Informacinė dalis (Administrative Metadata).
4. Turinio dalis (File Section).
5. Struktūrinė dalis (Structural Map).

Programinės įrangos pranašumas yra tai, kad ja galima būtų naudotis tiek tada, kai yra tiesioginis ryšys su virtualia sistema, tiek neturint ryšio, o duomenis perduoti įvairiomis laikmenomis. Šiuo metu gali būti naudojamas ftp protokolas duomenų persiuntimui, taip pat ir nešiojami diskai. Siekiant apsaugoti nuo trikdžių kopijuojant failus ir išvengti nepilnų, netikslų skaitmeninių vaizdų failų, yra naudojamas MD5 standarto (<http://tools.ietf.org/html/rfc1321>) kontrolinių sumų algoritmas. Prieš sudarant METS standarto paketą, yra išskaičiuojamos visų susietų su kūrinių vaizdais MD5 sumos ir įrašomos į paketą, o prieš perkeliant duomenis į sistemą - patikrinama, ar MD5 sumos atitinka, šitaip apsaugoma nuo duomenų persiuntimo klaidų.

Vykdyti procesus sunkiau ir dėl to, kad virtualioje sistemoje būtina sujungti skirtingus aprašomuosius metaduomenis skirtingose kodavimo sistemose. Šiai problemai spręsti, kaip jau buvo minėta, naudojamos sankirtos: UNIMARC/DC/EAD. Šiuo metu yra realizuotas duomenų, supakuotų į METS paketą, pateikimas šiais standartais: UNIMARC ir Dublin Core. Prireikus sistema gali būti lengvai praplečiama papildomais bibliografinio aprašo kodavimo standartais, tačiau galima laikyti ją atvirą pagal duomenų standartus.

### *Archyvinio saugojimo metaduomenų organizavimas*

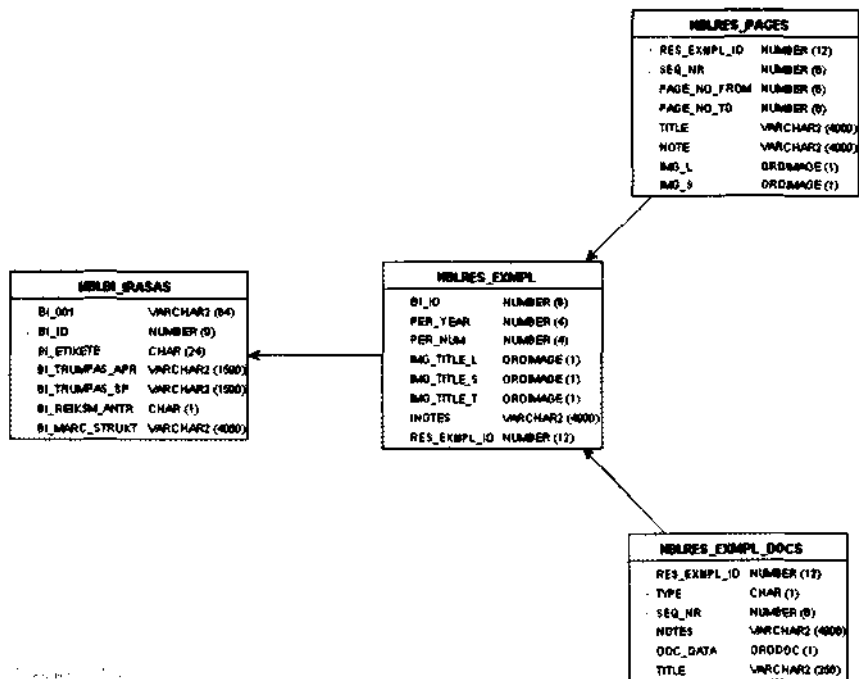
Kitaip negu aprašomųjų metaduomenų schemas (pavyzdžiui, MARC, UNIMARC, Dublin Core), kurios naudojamos skaitmeninių objektų suradimui ir identifikavimui, saugojimo metaduomenys kartais laikomi valdymo metaduomenų porūšiu, padedančiais valdyti informaciją arba techniniais metaduomenimis, palengvinančiais einamąją prieigą prie skaitmeninio objekto. Dabartiniu metu saugojimo metaduomenų kūrimas ir tobulinimas tebėra labiau eksperimentinio pobūdžio: siūlomi išsamūs metaduomenų elementų bei aukšto lygio rekomendacijų rinkiniai. Kitas vystymosi etapas būtų šių metaduomenų elementų praktiškas taikymas vis gausėjančiose skaitmeninių objektų valdymo sistemose.

Metaduomenų saugojimas kartu su jų apibūdinamu objektu užtikrina, kad jie nebus prarandami, panaikina duomenų ir metaduomenų susiejimo problemas bei garantuoja, kad kartu su objektu bus atnaujinami ir metaduomenys. Deja, į kai kurių rūšių objektus metaduomenų įterpti neįmanoma (pavyzdžiui, į meno objektus). Be to, metaduomenų saugojimas atskirai nuo objekto gali palengvinti pačių metaduomenų valdymą, taip pat paiešką ir suradimą. Todėl paprastai metaduomenys saugomi duomenų sistemoje ir susiejami su jų apibūdinamu objektu.

Vienas iš virtualios skaitmeninio paveldo sistemos tikslų yra pateikti visą sistemoje saugomą informaciją kaip vientisą katalogą, bendrą paiešką, vienodą paieškos rezultatų pateikimą bei kūrinių pavaizdavimą vartotojui. Tam tikslui yra naudojama viena duomenų bazė, kurioje visi saugomi įrašai yra vienodo formato, susiejant juos su skaitmeninių vaizdų failais. Kadangi saugomi metaduomenys skirtingomis kodavimo sistemomis (UNIMARC, EAD, DC), buvo nuspręsta pasirinkti vieną iš jų, o kitus formatus naudoti sankirtoms konvertuoti. Šiuo metu naudojama duomenų bazės struktūra

paikio UNIMARC (plačiausio iš naudojamų) aprašo kodavimo standartą, ir duomenys paieškai iš EAD ir DublinCore formato aprašų yra transformuojami į duomenų bazės struktūrą.

Toliau pateikiama duomenų struktūra, skirta vaizdams saugoti (4 pav.).



4 pav. Duomenų struktūra, skirta vaizdų saugojimui

Siekiant užtikrinti duomenų išsaugojimą, visi metaduomenys ir skaitmeniniai objektai yra dubliuojami, t. y. kuriamos atsarginės jų kopijos juostiniame įrenginyje siekiant apsaugoti nuo saugojimo laikmenų fizinių gedimų ir informacijos praradimo.

#### *Duomenų valdymas*

Duomenų valdymui ir tvarkymui buvo sukurta skaitmeninio turinio valdymo sistema, kuri užtikrina pateikiamų duomenų konvertavimą, jų susiejimą su vaizdų bylomis, visos meta-informacijos saugojimą, paiešką juose ir valdymą, vartotojų valdymą, prieigos teisių prie informacijos nustatymą ir informacijos iš duomenų saugyklos pateikimą.

### *Prieigos organizavimas*

Prieigai prie virtualios paveldo sistemos yra sukurta internete svetainė, pasiekiamą adresu: <<http://www.epaveldas.lt>>. Kadangi virtuali skaitmeninio paveldo sistema yra skirta vizualios informacijos pateikimui, prie joje saugomų skaitmeninių objektų yra pateikiami ir aprašomieji metaduomenys, o tikslesniam ir detalesniam nagrinėjimui vartotojas gali pereiti į detalų aprašo puslapį, esantį LIBIS, LAFS ar IRIS sistemose. Šiam tikslui užtikrinti yra realizuota abipusė nuorodų sistema, t. y. kiekvienas kūrinyms gali būti susietas su jį atitinkančiu aprašu originalioje sistemoje, ir atvirkščiai. Nuorodų sistema į virtualioje sistemoje saugomus objektus yra sukurta siekiant išlaikyti nuolatinės HTTP protokolo nuorodas ir net pakeitus virtualios paveldo sistemos realizaciją tos nuorodos nepasikeistų. Jų formatas yra: <<http://www.epaveldas.lt/vbspi/biDetails.do?libisId=C1R0000046342>>.

Čia: <<http://www.epaveldas.lt/vbspi/biDetails.do>> yra pastovioji dalis, o <[libisId=C1R0000046342](http://www.epaveldas.lt/vbspi/biDetails.do?libisId=C1R0000046342)> yra kintanti dalis, kuri kinta priklausomai nuo įrašo, jį atitinka UNIMARC aprašymo 001 lauko reikšmę. Žinant šią taisyklę, galima visuomet suformuoti dominančio kūrinio užklausą Virtualioje paveldo sistemoje.

### **Išvados**

Pirmą kartą Lietuvoje buvo suformuotas bendras požiūris į skaitmeninius kultūros paveldo objektus nepriklausomai nuo jų žinybinės priklausomybės, objekto kilmės, turinio ar struktūros. Šis požiūris buvo išreikštas parengtame virtualios sistemos lingvistiniame aprūpinime, kuris grindžiamas tarptautiniu mastu pripažintais standartais ir rekomendacijomis.

Priimti sprendimai dėl skaitmeninių objektų aprašymo išsamumo ir šių aprašų naudojimas atitinka atminties institucijoms keliamus reikalavimus dėl objektų atspindėjimo elektroniniuose kataloguose. Kartu atitinka vartotojų poreikius, kai reikia gauti išsamesnės informacijos nei jos yra portale kartu su skaitmeniniu objektu.

Sukauptos patirties panaudojimas kuriant ir plėtojant LAFS, LIBIS ir RIS, sudarė palankias sąlygas žengti kokybiškai kitą žingsnį - sukurti virtualią skaitmeninio paveldo sistemą sujungiant šias sistemas bendram tikslui - teikti kuo įvairesnes ir efektyvesnes paslaugas šalies vartotojams ir pasauliui.

Priimtas sprendimas, siekiant kuo daugiau pasinaudoti sukauptomis duomenų bazėmis, bibliografinių įrašų katalogais projekto partnerių sistemose, leido išvengti funkcijų dubliavimo, sumažino duomenų pateikimo į virtualią sistemą sąnaudas bei apsaugojo nuo klaidų.

### **NUORODOS**

1. *Komisijos komunikatas Tarybai, Europos Parlamentui, Europos ekonomikos ir socialinių reikalų komitetui ir regionų komitetui: i2010 Europos informacinė visuomenė augimui ir užimtumui skatinti* [SEC(2005) 717] [interaktyvus]. Europos Bendrijų Komisija. Briuselis, 1.6.2005. KOM(2005) 229 galutinis [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetu: <<http://eur-lex.europa.eu/LexUriServ/site/lt/>>

- com/2005/com2005\_02291t01.pdf>.
2. *Komisijos komunikatas Europos Parlamentui, Tarybai, Europos ekonomikos ir socialinių reikalų komitetui ir regionų komitetui: i2010: skaitmeninės bibliotekos* [SEC(2005) 1194]; [SEC(2005)1195] [interaktyvus]. Europos Bendrijų Komisija, Briuselis, 30.9.2005. KOM(2005) 465 galutinis [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <[http://eur-lex.europa.eu/LexUriServ/site/lt/com/2005/com2005\\_04651t01.pdf](http://eur-lex.europa.eu/LexUriServ/site/lt/com/2005/com2005_04651t01.pdf)>.
  3. *Komisijos komunikatas Europos Parlamentui, Tarybai, Europos ekonomikos ir socialinių reikalų komitetui: dėl mokslinės informacijos skaitmeniniame amžiuje: prieinamumas, sklaida ir išsaugojimas* [SEC(2007) 181] [interaktyvus]. Europos Bendrijų Komisija, Briuselis, 14.2.2007. KOM(2007) 56 galutinis [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <[http://eur-lex.europa.eu/LexUriServ/site/lt/com/2007/com2007\\_00561t01.pdf](http://eur-lex.europa.eu/LexUriServ/site/lt/com/2007/com2007_00561t01.pdf)>.
  4. *Europos Parlamento ir Tarybos direktyva, iš dalies keičianti Tarybos direktyvą 89/552/EEC: dėl valstybių narių įstatymuose ir kituose aktuose išdėstytų nuostatų, susijusių su televizijos programų transliavimu, derinimo* („Garso ir vaizdo žiniasklaidos paslaugos be sienų“): pateikta Komisijos [interaktyvus]. Europos Bendrijų Komisija, Briuselis, 29.03.2007. KOM(2007) 170 galutinis. 2005/0260 (COD) [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <[http://eur-lex.europa.eu/LexUriServ/site/lt/com/2007/com2007\\_01701t01.pdf](http://eur-lex.europa.eu/LexUriServ/site/lt/com/2007/com2007_01701t01.pdf)>.
  5. *Europos žodynas tezaurus EUROVOC* [interaktyvus]. 2007 [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <<http://www3.lrs.lt/pls/ev/ev.main>>.
  6. *Lietuvos nacionalinės Martyno Mažvydo bibliotekos rubrikynas*. Vilnius: LNB, 1994, 646 p.
  7. *Universal Decimal Classification (UDC)*. Vol. 12. Complete ed. London: BSI, 2005.
  8. *ULAN: The Union List of Artist Names* [interaktyvus]. Getty Center, Los Angeles, 2000 [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <[http://www.getty.edu/research/conducting\\_research/vocabularies/ulan/](http://www.getty.edu/research/conducting_research/vocabularies/ulan/)>.
  9. *EAD: Encoded Archival Description*, version 2002 [interaktyvus]. The Network Development and MARC Standards Office. 2002 [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <<http://www.loc.gov/ead/index.html>>.
  10. *UNIMARC Manual: Bibliographic Format* [interaktyvus]. IFLA. 2000, balandžio 6 d. [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <<http://www.ifla.org/VI/3/p1996-1/sec-uni.htm>>.
  11. *METS: Meta Encoding and Transmission Standard* [interaktyvus]. Version 1.6. METS editorial board. 2007 [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <<http://www.loc.gov/standards/mets/mets-home.html>>.
  12. *SRU Search/Retrieval via URL* [interaktyvus]. Version 1.2. The Library of Congress. 2007 rugsėjo 6 d. [žiūrėta 2007 m. rugsėjo 24 d.]. Prieiga per internetą: <<http://www.loc.gov/standards/sru/index.html>>
  13. *ISO 23950: 1998 Information and documentation - Information retrieval (Z39.50) - Application service definition and protocol specification*.
  14. *LST ISO 15836:2007 Informacija ir dokumentai. Dublin Core metaduomenų elementų grupė* (tapatus ISO 15836:2003).
  15. *ISAD(G): General International Standard Archival Description* [interaktyvus]. 2<sup>nd</sup>

- ed. ICA, Ottava, 2000. [Žiūrēta 2007 m. rugsējo 24 d.]. Prieiga per internetā: <[http://www.ica.org/sites/default/files/isad\\_g\\_2e.pdf](http://www.ica.org/sites/default/files/isad_g_2e.pdf)>.
16. *ISBD(CR): International Standard Bibliographic Description for Serials and Other Continuing Resources*. Cataloguing and the IFLA Section on Serial Publications, 2002. 112 p.
  17. *ISBD(ER): International Standard Bibliographic Description for Electronic Resources*. Recommended by the ISBD(CF) Review Group. 1997. 109 p.
  18. *ISBD(M): International Standard Bibliographic Description for Monographic Publications* [interaktyvus]. 2002 Revision. Recommended by the ISBD Review Group Approved by the Standing Committee of the IFLA Section on Cataloguing [Žiūrēta 2007 m. rugsējo 24 d.]. Prieiga per internetā: <[http://www.ifla.org/VII/sl3/pubs/isbd\\_m0602.pdf](http://www.ifla.org/VII/sl3/pubs/isbd_m0602.pdf)>.

## **VIRTUAL SYSTEM OF DIGITAL HERITAGE - THE BASIS FOR COLLABORATION BETWEEN MEMORY INSTITUTIONS**

DR REGINA VARNIENE

M. Mažvydas National Library of Lithuania, Vilnius, Lithuania,  
e-mail: r.varniene@lnb.lt

GIEDRIUS DAUGIALA

UAB „Sintagma“, Vilnius, Lithuania,  
e-mail: giedrius.daugiala@sintagma.lt

### Summary

The communications of the European Commission "i2010: A European Information Society for growth and employment", "i2010: Digital libraries" and "Access to scientific information in digital age" represent a new integrated approach to long-term preservation of cultural and scientific heritage. These documents promote memory institutions to consider the new collaboration paradigm. Major international projects such as Minerva, Minerva Plus, Digicult, Michael and others are aimed at exploration of closer collaboration perspectives. Digital convergence requires to make appropriate policy decisions. Considering collaboration of Lithuanian memory institutions in creation of digital heritage content, it should be argued that initiatives of M. Mažvydas national library of Lithuania in 2004-2005 have integrated principles of cultural policy, relevant methodology and practices. Although the project is not accomplished yet, it is already possible to speak about innovations that have crucial impact to:

1. Cultural policy because the project encouraged to adopt a strategy for digitization of cultural heritage that defined the mechanism for collaboration of memory institutions and their responsibility for the future development of this partnership.
2. Methodology and technology because institutions that participate in the project have agreed on the common metadata, file management, and classification standards that guarantee the long-term preservation of and effective access to digital information. The development of virtual system and common portal "epaveldas.lt" confirmed the appropriateness of methodological solutions reflected in the "Strategy for cultural heritage digitization" and project documentation.
3. Integration of Lithuania into the European space of digital content because by the end of the project 3 millions and 605 thousand of original manuscripts, newspapers, old books, metrics, chronicles, and other material of significant historical value will be digitized.

In this paper the main methodological solutions, recorded in the standards, considering linguistic provisions for the Virtual heritage system are surveyed. These standards allowed implementing the majority of objectives posed when shaping the architecture of the system, organizing of archiving and processing digital objects, access and interoperability of partner systems: LIBIS, LAFS and RIS.