

DOERR, Martin. Ontologies for Cultural Heritage. Iš: S. Staab and R. Studer (eds.), *Handbook on Ontologies*. Springer-Verlag Berlin Heidelberg. 2009, p. 463–486.

DOI 10.1007/978-3-540-92673-3

<https://pdfs.semanticscholar.org/9cff/f0ff6652e0e81dce434bbe572e391d2f57bf.pdf>

Vieno iš CIDOC CRM kūrėjų įžvalgos apie kultūros paveldo ontologijas

Kultūros paveldo srityje vis labiau įsigali informacijos technologijos, milžiniškais kiekiais gaminamos fizinių objektų skaitmeninės kopijos, atminties institucijos jaučia stiprų politinį spaudimą, verčiantį jas teikti viešą skaitmeninę savo fondų prieigą <...>

Vis didesne būtinybe tampa įvairių labai specializuotų sistemų funkcinis suderinamumas, integruota informacinė prieiga ir informacijos integracija kaip sėkmingų mokslinių tyrimų, profesinio paveldo valdymo, išsaugojimo, viešo domėjimosi ir ugdymo garantas. Todėl [kultūros paveldo] sričiai yra iškilusi kompleksinė schemų integravimo problema – kaip susieti sąryšingą įvairių specializuotų sistemų informaciją; ši problema gali būti veiksmingai sprendžiama naudojant formaliąsias ontologijas <...>

Neprofesionalui kultūros paveldas gali asocijuotis pirmiausia su vaizduojamojo meno rinkiniais; jam gali atrodyti, kad paveldo objektų aprašymas ir dalykinimas yra palyginti paprastas veiksmas, labiau susijęs su moksline refleksija apie šių objektų vertę, o ne su logika. Iš tikrųjų kultūros paveldas yra platesnis reiškiny. Jis apima platų spektrą funkcijų, susijusių su įvairialypės žmonijos praeities veiklos fizinių įrodymų tyrinėjimu ir išsaugojimu <...> Siauresne prasme kultūros paveldas gali būti suvokiamas kaip *atminties institucijose* (t. y. muziejuose, archeologinių tyrinėjimų vietose <...>, archyvuose, bibliotekose) saugojami objektai <...>

Bibliotekos tam tikra prasme irgi saugoja kultūros paveldą, jeigu jose yra unikalių knygų; vis dėlto pagrindinė bibliotekų funkcija yra teikti neunikalių informacijos šaltinių prieigą <...> Dauguma kultūros paveldo objektų yra

gana nebylūs praeities įvykių liudininkai; jie svarbūs tiek, kiek yra suprantamas jų kilmės ir istorijos kontekstas. Tam tikras objektas gali būti ne tiek informacijos šaltinis, kiek praeities „ilustracija“. Ši skirtis yra svarbi suvokiant bibliotekų teikiamą informaciją ir informaciją, kurią apima kultūros paveldas, bei nepaprastą pastarosios sudėtingumą <...>

Apdorojant [kultūros paveldo] informaciją, vartojama labai turtinga terminologija – daugiakalbė ir dažnai būdinga tik tam tikroms bendruomenėms arba net atskiriems mokslininkams. Būna sunku pasiekti sutarimą dėl vienodo terminų vartojimo, o ir dažnai nebūna terminų atitikmenų kitomis kalbomis. Visa tai kelia akivaizdžių iššūkių taikant formaliąsias ontologijas: tai ir techninio pobūdžio sunkumai, ir intelektualinės problemos siekiant suderinti intuityvias arba tradicines sąvokas teikiant logines apibrėžtis, pavyzdžiui, to paties termino galimas siauresnes ar platesnes reikšmes <...>

Kultūros paveldo informacijos sistemas galima klasifikuoti pagal joms būdingas pagrindines funkcijas. Šios sistemos gali būti skirtos:

- Fondų valdymui <...>
- Išsaugojimui <...>
- Moksliniams tyrimams <...>
- Pateikimui <...>

Fondų valdymo sistemos <...> dažniausiai remiasi sąryšinėmis arba hierarchinėmis duomenų bazėmis <...> Jos užtikrina fondų, archeologinių tyrimų vietų ir paminklų techninį valdymą bei administravimą <...>

<...> *Išsaugojimo informacija* gali būti sudedamoji fondų valdymo sistemos dalis arba atskira sritis. Ji susijusi su moksline objektų analize, profilaktinėmis ir intervencinėmis priemonėmis. Ši informacija gali būti reikalinga taip pat ir fondų išduoties valdymui bei istoriniams tyrinėjimams <...> Ontologijų taikymas šiose sistemose suteikia joms privalumų, nes jos susijusios su *kategorinėmis* (bendrosiomis) žiniomis <...>

Mokslinių tyrimų informacijos sistemos yra labai specializuotos ir dažniausiai kuriamos pagal konkrečių projektų poreikį. Yra informacinių sistemų, apimančių visus žinomus tam tikro tipo elementus, pavyzdžiui „Lotyniški įrašai“ <...>, „Suvestinis menininkų vardų sąrašas“ <...> Deja, dėl unikalios dizaino ir nepakankamo išteklių nuorodų valdymo projektui pasibaigus nebeįmanoma panaudoti sukauptos informacijos <...> Kol kas nėra sukurta veiksmingų duomenų konvertavimo ir perkėlimo priemonių. Svarų indėlis į tokių priemonių kūrimą galėtų įnešti ontologijų taikymas <...>

Pateikimo sistemos teikia kultūros paveldo informacijos prieigą visuomenei arba prenumeratoriams, ypač mokytojams ir dėstytojams <...> Mūsų turimais duomenimis, daugiau kaip 95% muziejuose saugomų objektų nėra eksponuojama; archyvai irgi paprastai visuomenei neprieinami. Todėl jaučiamas didelis politinis spaudimas teikti bent jau viešą objektų aprašų, kurie rengiami naudojant fondų valdymo sistemas, prieigą <...> Pateikimo sistemos – tai ir virtualios parodos, fizinę objektų parodą papildanti informacija, tam tikrai temai skirta

elektroninė publikacija. Ontologijos yra svarbi priemonė teikiant struktūrinius kreipties elementus tokioms sistemoms ir sisteminant jų teikiamą informaciją <...>

Kadangi bibliotekose saugomi objektai ir skaitmeninėse bibliotekose teikiami objektai apima duomenis, bibliotekos šių objektų aprašus vadina „metaduomenimis“, t. y. duomenimis apie duomenis. Muziejai, nors jų objektai ir neapima duomenų, taip pat vartoja šį terminą. Labai dažnai metaduomenys sisteminami kaip vienalygės aibės savybių, agreguojamų vadinamuosiuose „taikomuosiuose profiliuose“ <...>, o įvairių metaduomenų formatų atitiktys ir jų konvertavimo pateiktys kartais vadinamos „metaduomenų sankirtomis“ <...> Metaduomenų savybių žymos (pvz., „kūrėjas“, „data“ ir kt.) primena sąvokas. Todėl kai kurie autoriai metaduomenų schemas traktuoja kaip „žodynus“ arba tam tikra prasme ontologijas. Mūsų nuomone, taip painiojami informaciniai modeliai ir ontologijos <...> Kompiuterijoje formaliąsias ontologijas imta naudoti siekiant apibūdinti įvairialypių schemų konceptualų sąryšį <...>; formalioji ontologija nebuvo suvokiama kaip schemas sinonimas. Be to, sudėtingą objekto istoriją teikiant kaip vienalygę savybių aibę, šios savybės tampa semantiškai pernelyg daugiareikšmiškos, o tai visiškai neatitinka ontologijos apibrėžimo <...>

Integruojant informaciją objektų radimo priemonių pagrindu, į visumą nejungiamo informacija apie šių objektų liudijamą istorinį kontekstą, kuris jiems suteikia vertę. Siekiant tinkamai integruoti kultūros paveldo informacinį turinį, būtina taikyti sudėtingesnius modelius. Štai Kalifornijos Mokslinių bibliotekų konsorciumas, įgyvendindamas Kultūros medžiagos iniciatyvą, sėkmingai ir iš esmės nenuostolingai integravo duomenis iš maždaug tūkstančio kultūros institucijų į vieną daug sudėtingesnę sistemą; duomenys buvo koduoti naudojant maždaug šimtą įvairių schemų. Ši sistema remiasi CIDOC CRM ontologija, dabar jau turinčia tarptautinio standarto statusą (ISO 21127), kuri šiuo metu yra labiausiai išplėtotą kultūros paveldo informacijos integravimui skirta ontologija.

CIDOC CRM yra formalioji ontologija <...>, skirta palengvinti įvairialypės kultūros paveldo informacijos integravimą, sklaidą ir mainus. Ją kūrė ekspertai, atstovaujantys įvairias mokslo šakas (kompiuteriją, archeologiją, muziejų informacijos dokumentavimą, meno istoriją, gamtotyrą, bibliotekininkystę, fiziką, filosofiją), kuriems vadovavo Tarptautinės muziejų tarybos (ICOM) tarptautinis dokumentavimo komitetas (CIDOC). Ontologija buvo kuriama <...> perprojektuojant ir integruojant <...> įvairių su muziejais ir archyvais susijusių duomenų bazių schemų semantinį turinį ir dokumentavimo struktūras; pastaruoju metu šis procesas apėmė ir bibliotekas <...>

Dabartinė modelio versija apima 80 klasių ir 132 savybes, kurios apibūdina keletą šimtų schemų semantiką <...>

CRM taikymo srities mastą patvirtina daugelio kultūros paveldo duomenų struktūrų ir CRM atitiktys. Galima netgi nustatyti tikslią CRM ir MARC <...> atitiktį <...> *Dublin Core* ir EAD bei CRM <...> atitiktys rodo, kaip tiksliai apibrėžta bendroji semantika gali derėti su tipiniais metaduomenų formatais. Parengta netgi MPEG7 ir CRM atitiktis <...> CRM modelis vis plačiau naudojamas integruotoje informacinėje aplinkoje veikiančiose kultūros paveldo sistemose <...> Dėl CRM būdingo empirinio pagrindo (t. y. dėl to, kad jos duomenų struktūros skirtos

fondų rinkiniams koduoti), šis modelis nelabai tinka giminės santykiams, teisėms ir intelektualiniams procesams aprašyti. Tokiems aprašams neseniai sukurtas FRBRoo modelis <...>

FRBR („Funkcinių reikalavimų bibliografiniams įrašams“) modelis buvo parengtas kaip į esybes ir jų santykius orientuotas modelis <...> 1991–1997 m. Jis buvo paskelbtas 1998 m. Modelis skiria keturis abstrakcijos lygmenis pradedant koncepcija ir baigiant turimu fiziniu objektu: „kūrinį“, „išraišką“, „apraišką“ ir „vienetą“. Jis nėra skirtas kokiai nors konkrečiai sričiai ir gali būti laikomas pažangiausia bibliotekų informacijos konceptualizavimo priemone <...>

2003 m. buvo sutelkta Tarptautinė FRBR ir CIDOC CRM darninimo darbo grupė, kurios tikslas –išreikšti IFLA FRBR modelį naudojant CIDOC CRM teikiamas sąvokas, ontologinę metodologiją ir kodavimo nuostatas bei susieti šiuos du į objektą orientuotus modelius. Nors abi sritys [FRBR ir CIDOC CRM] susijusios su kultūros paveldo objektais, kiekvienos iš šių sričių objektai labai skiriasi. Didžiąją dalį bibliotekų fondų sudaro neunikalūs objektai, t. y. industrinio proceso produktai. Todėl FRBR pirmiausia orientuotas į „abstrakčias“ savybes, kurios būdingos visiems tam tikros publikacijos egzemplioriams ir pagal kurias šie egzemplioriai identifikuojami kaip atitinkamos publikacijos egzemplioriai. Atskirų egzempliorių ir nematerialaus turinio istorija bibliotekų kataloguose nelaikoma ypatingai svarbia, todėl FRBR į ją beveik neatsižvelgiama. Bibliotekose irgi saugoma unikali medžiaga (pvz., rankraščiai), tačiau nėra jai skirtų tarptautinių standartų; FRBR apie tokią medžiagą užsimenama, bet ji šiame modelyje išsamiai neaptariama.

Muziejų veikla dažniausiai yra susijusi su unikaliais objektais, tačiau ji orientuota taip pat ir į kultūrinį šių objektų sukūrimo kontekstą <...> Todėl CIDOC CRM akcentuojami fiziniai objektų požymiai, šių objektų sukūrimo kontekstas bei įvairūs jų sieties su kitais objektais, objektų kategorijomis ir netgi kultūros tendencijų ideologinėmis sistemomis būdai <...>

FRBRoo, t. y. CIDOC CRM besiremianti į objektą orientuota FRBR versija, <...> skirta pamatinei bibliografinės informacijos semantikai fiksuoti ir aprašyti bei palengvinti bibliografinės ir muziejų informacijos integravimą, sklaidą ir mainus.

Pagrindinis inovatyvus FRBRoo aspektas yra tikroviškas formalizuotas intelektualinio kūrybinio proceso modelis, kuris dar turi būti tobulinamas ir pritaikomas bibliotekoms bei įvairioms muziejų mokslo disciplinoms. FRBRoo nustato esminį kūrybinio proto („kūrinys“), žmogaus interpretuojamų ženklų („išraiška“) ir fizinės informacijos laikmenos (apraiška“) skirtumą <...>

Maždaug 1997 metais multimedijos ekspertų konsorciumas *IndeCs* parengė pamatinį modelį, skirtą multimedijos produktų turinio kilmei ir susijusioms nuosavybės teisėms fiksuoti bei su juo susijusią informacinę sistemą. Šio modelio pagrindu buvo sukurta ABC ontologija <...> ABC ontologijos kūrimo veiklos pirminis tikslas ir ilgalaikė paskata yra susijusi su poreikiu skaitmeninėse bibliotekose integruoti įvairių multimedijos žanrų informaciją <...> ABC remiasi FRBR modelio koncepcijomis <...> „kūrinys“, „išraiška“, „apraiška“ ir „vienetas“ <...>

ABC ir CRM panašumas buvo toks ryškus, kad nuo 2001 iki 2003 m. abiejų modelių kūrėjai kartu įgyvendino šių modelių darninimo projektą, kurio metu abi ontologijos pritaikė sau viena kitos koncepcijas <...>

Esama dar vienos įdomios pamatinės ontologijos – DOLCE <...> Tai *WordNet* (lingvistinio išteklių, besiremiančio žodynais ir praturtinto <...> pamatiniais santykiais, pvz., „dalyvavimas“, dalis ir visuma“, „struktūriniai ryšiai“ ir kt.) pagrindinių koncepcijų kruopščios pertvarkos rezultatas. Ši ontologija remiasi labai griežta logika, todėl gana sunkiai suvokiama ir pritaikoma <...>

Ontologijos, besiremiančios duomenų struktūroms tapačia semantika (t. y. mūsų apibūdintos ontologijos), apima mažai klasių, bet daug santykių <...>; tuo jos skiriasi nuo terminologijos ontologijų, skirtų objektams sisteminti. Duomenų struktūros gali būti interpretuojamos kaip teiginių apie tam tikrą sritį ekvivalentai <...>. Todėl jų semantika atspindi tam tikrai sričiai būdingo diskurso elementus <...>

Kultūros paveldas gali būti traktuojamas kaip socialiai reikšmingos praeityje atliktos žmogaus veiklos materialus įrodymas. Todėl:

- Informacija yra vidutinės apimties, t. y. <...> ne astronominio masto ir ne mikroskopinė, išskyrus mikroskopinę pėdsakų ir medžiagų analizę <...> Procesai apibūdinami arba tampa akivaizdūs kaip pavieniai įvykiai, apimantys pavienius dalykus – kitaip negu atspindint geologinius ar meteorologinius reiškinius.
- Informacija yra orientuota į įvykį. Dalykai, žmonės ir idėjos siejasi per įvykius.
- Ji [informacija] aprašoma retrospektyviai – kitaip negu aprašant perspektyvą, pvz., gamybos srityje.

Tam tikru mastu informacija yra neišsami. Ji gali būti papildoma, tačiau negali būti baigtinė. Jos aprašymas primena paiešką siekiant rekonstruoti *galimą praeitį*. Todėl, remiantis tariama praeitimi, informacijos neįmanoma standartizuoti ir integruoti <...> Siekiant teisingai interpretuoti stebimus faktus, būtina dokumentuoti stebėjimo procesą. Netgi tai, kad mokslininkas klasifikuoja tam tikrą objektą vartodamas kokį nors terminą, dokumentuojama kaip istorinis ir intelektualinis procesas (tai būdinga ir biologinės įvairovės sričiai). Stebimi pavieniai faktai naudojami kaip pagrindas teikti kategoriškas išvadas apie tam tikrą elgseną, pvz., „visi faraonai buvo mumifikuojami“.

Visa tai būdinga taip pat ir kitiems aprašomiesiems mokslams: geografijai, biologinei įvairovei, paleontologijai, klinikiniam stebėjimui, epidemijų tyrimui ir kt., lygiai kaip ir dokumentuojant eksperimentus bei faktus gamtos moksluose. Šiose srityse stebėjimo išvados formuluojamos vartojant kategorines išraiškas („ $F=m*a$ “ arba „bet kuris nepritvirtintas Žemės atmosferoje esantis objektas krenta“); tačiau kultūros paveldo interpretuotojai paprastai vengia formaliai reikšti kategorines išvadas ar hipotezes <...> Interpretacija dažniausiai pateikiama tekstu, atspindinčiu <...> nuomonių gausą. Todėl tokios ontologijos yra stebėtinai universalios <...>

Įgyvendinant bendrus projektus ir diskutuojant su muziejų fondų saugotojais bei archeologais, susidurta su neigiama nuostata kontroliuojamų žodynų ir netgi formaliųjų ontologijų naudojimo atžvilgiu. Ekspertai yra linkę nepritari kolegų vartojamai terminologijai <...> Visiškai kitaip yra bibliotekų srityje, kuri tiesiog negali apsieiti be standartizuotos terminologijos. Mes tai aiškiname tuo, jog kultūros paveldo srityje terminologija yra ne tiek ekspertų bendro sutarimo priemonė, kiek intelektinė hipotezių kūrimo skiriamą reiškinį priemonė. Todėl archeologijoje automatiškai atliekamas klasifikavimas yra ilgą istoriją turinti disciplina, tačiau plačiau naudojamos tik kelios terminologinės sistemos. Jos skirtos tik specifiniams mokslinių tyrimų aspektams <...>

Bibliotekinių tikslas yra ne kurti hipotezes, o teikti informacijos prieigą. Todėl natūralu, kad jiems jau seniai įprasta remtis bendrais terminais kaip kreipties elementais. Kultūros paveldo ekspertams nėra taip paprasta pripažinti bendrų paieškai skirtų terminų būtinybę <...> Mes manome, kad galėtų būti skiriami du kultūros paveldo terminologijos lygmenys: aukštesnysis, stabilus, lygmuo, tinkantis paieškai, ir žemesnis, kintamas, lygmuo, skirtas hipotezių kūrimui. Tokią nuomonę suformavo patirtis, įgyta rengiant ISO 21127 ir kuriant įvairias informacines sistemas. Didžiausias ir stabilus [kultūros paveldo] sisteminis žodynas (*Art & Architecture Thesaurus*, AAT) <...>, apimantis daugiau negu 30 000 sąvokų, iš tikrųjų remiasi bibliotekų patirtimi <...>

Materialių objektų klasifikavimą sunkina tai, kad klasifikuoti galima remiantis įvairiais aspektais. Tuo tikslu dažniausiai naudojami aspektai yra objekto funkcija, jo forma ir išvaizda bei jo struktūros principai arba elementai <...> Tai išspręsti bandoma taikant vadinamąją aspektinę analizę <...>: kiekvienam aspektui sistemingai atrenkamos sąvokos. Atliekant „aspektinę klasifikaciją“, kurios pradininkas yra Indijos bibliotekininkas Ranganathanas <...>, kiekvienam atitinkamam aspektui naudojama sisteminė klasifikacijos terminų aibė <...>

Visų terminologinių sistemų hierarchijų pamatiniai elementai yra labai bendri terminai. Jie gali būti labai įvairialypiai ir apsunkinti ontologijų suderinamumą, nes šių ontologijų paskirtis nėra spręsti filosofinius dalykus, su kuriais šie bendri terminai yra susiję. Tarkime, AAT ontologijoje *vizualiniai kūriniai* apima materialius ir nematerialius dalykus (*tapybos kūrinius* ir *elektroninius vaizdus*). CIDOC CRM ontologijoje materialūs ir nematerialūs dalykai yra dvi skirtingos koncepcijos <...>