



**BŪTINI POKYČIAI IR INVESTICIJŲ
KRYPTYS SIEKIANČIĄ LIETUVOS
ĮSIJUNGIMO Į EUROPOS
ATVIROJO MOKSLO DEBESĮ:
GALIMYBIŲ STUDIJA**

Vilnius, 2026



**Bendrai finansuoja
Europos Sąjunga**

Projektas įgyvendinamas pagal Lietuvos Respublikos švietimo, mokslo ir sporto ministerijos mokslo plėtros programos pažangos priemonę Nr. 12-001-01-01-01 „Gerinti mokslo ir studijų aplinką“ (veikla „Įsijungti į Europos atvirojo mokslo debesį“), projekto kodas: 10-073-P-0001. Veikla, įgyvendinama 2021–2027 metų Europos Sąjungos fondų ir Lietuvos Respublikos valstybės biudžeto lėšomis. Projekto „Įsijungti į Europos atvirojo mokslo debesį – galimybių studija dėl būtinų pokyčių ir investicijų krypčių“ vadovė – Gerinta Raguckė.

Elektroninis leidinys

Bibliografinė informacija pateikiama Lietuvos integralios bibliotekų informacinės sistemos (LIBIS) portale ibiblioteka.lt.

ISBN 978-609-95905-7-8 (pdf)

© Lietuvos mokslo taryba, 2026

Šis kūrinys licencijuojamas remiantis CC BY 4.0 (© ⓘ).

Galimybių studiją parengusių ekspertų grupė

Rūta Petrauskaitė – Vytauto Didžiojo universiteto profesorė (darbo grupės vadovė)

Darius Amilevičius – Vytauto Didžiojo universiteto docentas, vyriausiasis mokslo darbuotojas

Valdonė Darškuvienė – ISM vadybos ir ekonomikos universiteto profesorė

Saulius Gražulis – Vilniaus universiteto vyriausiasis mokslo darbuotojas

Aušra Gribauskienė – Švietimo, mokslo ir sporto ministerijos Mokslo skyriaus vyriausioji specialistė

Gintė Medzvieckaitė – Vilniaus universiteto Mokslinės informacijos ir duomenų skyriaus duomenų vadybininkė

Saulė Milčiuvienė – Vytauto Didžiojo universiteto Teisės fakulteto docentė, Skaitmeninių išteklių ir tarpdisciplininių tyrimų instituto vyriausioji mokslo darbuotoja

Vaidas Morkevičius – Kauno technologijos universiteto Socialinių, humanitarinių mokslų ir menų fakulteto profesorius

Gerinta Raguckė – Lietuvos mokslo tarybos Mokslo ir studijų politikos analizės skyriaus vyriausioji politikos (mokslo) analitikė

Mindaugas Šarpis – Vilniaus universiteto mokslinių tyrimų grupės vadovas

Jolanta Šinkūnienė – Vilniaus universiteto Filologijos fakulteto profesorė

Gintarė Tautkevičienė – Kauno technologijos universiteto bibliotekos direktorė

TURINYS

Santrauka	6
Santrumpos	8
Įvadas	10
1. Galimybių studijos rengimo metodika ir šaltinių atranka	14
1.1. Bendras metodologinis požiūris ir taikytų metodų visuma	14
1.1.1. Bendras metodologinis požiūris	14
1.1.2. Teorinė-normatyvinė analizė ir šaltinių parinkimo metodika	14
1.1.3. Esamos situacijos analizė	15
1.1.4. Analitinė sintezė ir strateginis modeliavimas	16
1.1.5. Kokybės užtikrinimas ir etika	17
1.2. Apklausa ir fokusuotos grupinės diskusijos	17
1.2.1. Apklausa metodika	17
1.2.2. Fokusuotų grupinių diskusijų metodika	19
1.3. Mokslinių tyrimų duomenų saugyklų identifikavimo ir aprėpties vertinimo metodika	20
2. Mokslinių tyrimų duomenų panaudojimo teisinis reglamentavimas, principai, standartai, sąveikumas, valdysena ir tvarkysena	23
2.1. Mokslinių tyrimų duomenų naudojimo strateginiai tikslai ir teisinė bazė	23
2.1.1. Mokslinių tyrimų duomenų apibrėžimas	23
2.1.2. ES mokslinių tyrimų duomenų valdymo strateginės nuostatos ir teisinis reguliavimas	24
2.1.3. LR teisės aktai ir Lietuvos mokslą finansuojančių institucijų reikalavimai	30
2.1.4. Apibendrinimas	32
2.2. Mokslinių tyrimų duomenų FAIR, CARE, TRUST principai	33
2.2.1. FAIR principai	33
2.2.2. CARE principai	36
2.2.3. TRUST principai	37
2.2.4. Apibendrinimas	39
2.3. Mokslinių tyrimų duomenų tipai ir standartai, sąveikumas, valdysena ir tvarkysena, duomenų saugyklos	40
2.3.1. Mokslinių tyrimų duomenų tipai ir standartai	40
2.3.2. Mokslinių tyrimų duomenų ir sistemų semantinis, organizacinis, teisinis ir techninis sąveikumas	45
2.3.3. Mokslinių tyrimų duomenų saugyklos – patikimumas ir sąveikumas	53
2.3.4. Apibendrinimas	55
2.4. Įvairių Europos šalių mokslinių tyrimų duomenų infrastruktūros ir valdysenos modeliai	56
2.4.1. Airija	56
2.4.2. Austrija	57
2.4.3. Čekija	57
2.4.4. Danija	58
2.4.5. Estija	59
2.4.6. Graikija	60
2.4.7. Ispanija	60
2.4.8. Latvija	61
2.4.9. Lenkija	62
2.4.10. Nyderlandai	63
2.4.11. Apibendrinimas	63

3. Mokslinių tyrimų duomenų atvėrimui, tvarkymui ir priežiūrai reikalingi žmogiškieji ištekliai ir kompetencijos bei jų ugdymas	65
3.1. EOSC ekosistemos dalyvių vaidmenys ir reikiamos kompetencijos	65
3.2. Duomenų valdytojų bei kitų su mokslo duomenų valdysena ir tvarkysena susijusių specialistų kompetencijų ugdymas .	70
3.3. Apibendrinimas.....	74
4. Mokslinių tyrimų duomenų ekosistemos Lietuvoje analizė	75
4.1. Lietuvos mokslo ir studijų institucijų apklausos analizė	75
4.2. Mokslininkų apklausos analizė	85
4.3. Fokusuotosios grupinės diskusijos	90
4.4. Apibendrinimas.....	91
5. SSGG analizė ir Lietuvos įsijungimo į EOSC alternatyvos ir jų scenarijai	93
5.1. SSGG analizė	93
5.2. Pirmoji alternatyva: status quo – decentralizuotas modelis	94
5.3. Antroji alternatyva: nacionalinis EOSC mazgas	95
5.3.1. Scenarijus 1: Nacionalinio EOSC-LT mazgo kaip valstybės informacinės sistemos scenarijus	98
5.3.2. Scenarijus 2: Konsorciumo ir akademinio tinklo pagrindu organizuojamo EOSC-LT mazgo scenarijus (sektorinė mokslo duomenų erdvė)	102
5.3.3. Scenarijus 3: Konsorciumo savivaldos pagrindu organizuojamo EOSC-LT mazgo, kaip autonominės sektorinės mokslo duomenų ervės, scenarijus	104
5.3.4. Galutinė išvada dėl antros alternatyvos scenarijų	105
5.4. Trečioji alternatyva: regioninis EOSC mazgas	106
5.5. Alternatyvų palyginimo išvados	107
6. Finansinė MSI mokslinių tyrimų duomenų valdymo situacija	109
6.1. Dabartinis finansavimas	109
6.2. Būtinasis finansavimas ir jo paskirstymas integruojantis į EOSC	110
7. Literatūra	124
8. Priedai	129

Santrauka

Šioje galimybių studijoje analizuojami būtini organizaciniai, teisiniai, technologiniai ir kompetencijų pokyčiai, reikalingi Lietuvai įsijungti į Europos atvirojo mokslo debesį (EOSC), pagrindžiamos investicijų kryptys ir galimi įgyvendinimo modeliai. Studija remiasi sekvenciniu mišriųjų metodų dizainu, jungiančiu teorinę-normatyvinę analizę, empirinį esamos situacijos tyrimą ir strateginį alternatyvų modeliavimą. Rengiant studiją išanalizuota daugiau kaip 100 strateginių, teisinių ir metodinių šaltinių, atliktos institucijų ir mokslininkų apklausos, organizuotos fokusuotosios grupinės diskusijos, sudarytas Lietuvos mokslinių tyrimų duomenų (MTD) ekosistemos analitinis žemėlapis. Tokia metodologinė prieiga leido ne tik aprašyti dabartinę situaciją, bet ir sistemaiškai įvertinti Lietuvos pasirengimą integruotis į besiformuojančią EOSC federaciją.

Studija grindžiama prielaida, kad atvirasis mokslas ir EOSC šiandien tampa ne vien mokslo politikos kryptimi, bet ir svarbia Europos mokslinių tyrimų, skaitmeninio suvereniteto bei duomenimis grįstos inovacijų ekosistemos dalimi. EOSC plėtra siejama su poreikiu užtikrinti, kad MTD būtų ne tik prieinami, bet ir surandami, sąveikūs, pakartotinai naudojami bei tinkami mašininiam apdorojimui. Dėl to studijoje daug dėmesio skiriama FAIR, CARE ir TRUST principams, taip pat teisiniam, semantiniam, organizaciniam ir techniniam sąveikumui. Pabrėžiama, kad šiuolaikinė MTD valdysena evoliucionuoja nuo pasyvaus duomenų saugojimo prie aktyvaus jų įveiklinimo, o duomenų kokybė ir standartizacija tampa strategine sąlyga tiek moksliniam bendradarbiavimui, tiek DI taikymams. Lietuvos pasirengimas dalyvauti EOSC turi būti vertinamas ne vien pagal saugyklų ar rinkinių kiekį, bet pagal tai, kiek nuosekliai institucijos identifikuoja duomenų tipus, taiko standartus ir geba integruoti savo išteklius į federuotą Europos aplinką.

Tarptautinė analizė, apėmusi dešimties Europos valstybių patirtį, parodė, kad vieno universalus MTD valdysenos modelio nėra, tačiau sėkmingą integraciją į EOSC paprastai lemia aiškus nacionalinis koordinavimas, institucinė atsakomybė, suderinta infrastruktūra ir nuoseklus kompetencijų ugdymas. Ši išvada ypač svarbi Lietuvos kontekste, nes empiriniai tyrimai atskleidė netolygų institucijų pasirengimą. MTD valdymo praktikos, technologiniai sprendimai, duomenų atvėrimo lygis ir kompetencijos skirtingose mokslo ir studijų institucijose labai skiriasi. Lietuvos MTD ekosistemos žemėlapis parodė fragmentuotą, daugiasluoksnę struktūrą, kurioje veikia 81 objektas ir 189 ryšiai, tačiau ši struktūra vis dar stokoja bendro nacionalinio katalogavimo, aiškių valdymo sąsajų ir centralizuoto koordinavimo. Be to, apklausos rodo, kad gamtos, technologijų, medicinos ir sveikatos moksluose vyrauja didelės apimties eksperimentiniai bei modeliavimo duomenys, o humanitariniuose ir socialiniuose moksluose – tekstiniai, audiovizualiniai ir archyviniai išteklių. Tai reiškia, kad Lietuvos integracija į EOSC negali būti grindžiama vienu techniniu modeliu ir turi remtis diferencijuotu, tačiau horizontaliai suderintu požiūriu.

Atskira studijos išvada susijusi su žmogiškaisiais išteklių. Nustatyta, kad vienas esminių Lietuvos pasirengimo ribojimų yra nepakankamai išplėtotas duomenų valdytojų¹ ir kitų su MTD valdysena susijusių specialistų tinklas. Kompetencijų ugdymas negali būti laikomas antrine ar tik papildoma veikla, nes jis turi vykti tuo pat metu su infrastruktūros ir valdysenos sprendimais. Studijoje argumentuojama, kad Lietuvos institucijoms nepakanka pavienių specialistų dalyvavimo užsienio mokymuose. Būtina kurti nacionalinei situacijai pritaikytą duomenų valdytojų rengimo programą, paremtą vietos infrastruktūros, teisinio reguliavimo ir institucinių poreikių pažinimu. Tai reikštų perėjimą nuo fragmentiško kompetencijų kaupimo prie sistemiško nacionalinio gebėjimų stiprinimo modelio.

Strateginio modeliavimo dalyje studijoje lyginamos trys pagrindinės Lietuvos įsijungimo į EOSC alternatyvos: decentralizuotas *status quo* modelis, nacionalinio mazgo kūrimas ir regioninio mazgo alternatyva. Atlikta analizė leidžia daryti aiškią išvadą, kad didžiausią pridėtinę vertę Lietuvai sukurtų nacionalinio mazgo alternatyva, nes ji geriausiai suderina nacionalinį interesą, sąveikumą su EOSC, institucinį tvarumą ir praktinį įgyvendinamumą. *Status quo* modelis palieka fragmentuotą MTD erdvę, neišsprendžia bendro nacionalinio matomumo, katalogavimo ir ilgalaikio koordinavimo problemos, o regioninis mazgas, nors teoriškai įmanomas, Lietuvos atveju būtų mažiau palankus dėl ribotesnės tiesioginės nacionalinės naudos, sudėtingesnės valdysenos ir didesnės priklausomybės nuo išorinio koordinavimo. Todėl studijoje tinkamiausiu pripažįstamas pirmasis nacionalinio mazgo

¹ Duomenų valdytojas šiame dokumente apibūdinamas kaip duomenų rengimo ir tvarkymo ekspertas, įskaitant duomenų atranką, saugojimą, išsaugojimą, kilmės ir kitų metaduomenų rengimą, tvarkymą bei sklaidą.

scenarijus – nacionalinio EOSC-LT mazgo organizavimas kaip valstybės informacinės sistemos. Šiame modelyje EOSC-LT-Core veiktų kaip nacionalinis centrinis MTD metaduomenų katalogas ir pagrindinis Lietuvos integracijos į ES EOSC taškas, o strateginis valdymas būtų aiškiai atskirtas nuo techninio įgyvendinimo. Toks sprendinys sudarytų prielaidas mazgą institucionalizuoti kaip ilgalaikę nacionalinės svarbos skaitmeninę infrastruktūrą, o ne laikiną projektinę iniciatyvą.

Finansinės analizės dalyje konstatuojama, kad dabartinis MTD finansavimas Lietuvoje yra mišrus ir daugiakanalis, tačiau iš esmės fragmentuotas. Jis dažniausiai susijęs su platesnėmis MTEP, infrastruktūros ar skaitmenizacijos veiklomis, o ne su nuoseklia nacionaline MTD gyvavimo ciklą apimančia finansavimo logika. Todėl, siekiant realios integracijos į EOSC, būtini kertiniai pokyčiai, apimantys ne tik techninę infrastruktūrą, bet ir valdyseną, kompetencijas, katalogavimo sluoksnį, atvirumo ir ilgalaikio išsaugojimo mechanizmus. Studija rodo, kad investicijos turi būti planuojamos kaip tarpusavyje susijęs paketas: infrastruktūros modernizavimas, nacionalinio mazgo branduolio kūrimas, duomenų valdytojų rengimas, metaduomenų ir sąveikumo standartų diegimas, institucinės politikos stiprinimas bei tęstinis finansavimo modelis po projektinio etapo. Taigi studija pagrindžia, kad Lietuvos įsijungimas į EOSC neturėtų būti suvokiamas kaip siauras technologinis projektas. Tai nacionalinės MTD erdvės formavimo uždavinys, reikalaujantis suderintos teisėkūros, valdymo, infrastruktūros, kompetencijų ir investicijų politikos. Tik toks nuoseklus modelis leistų Lietuvai ne tik techniškai prisijungti prie EOSC, bet ir tapti visaverte, tvaria ir konkurencinga europinės atvirojo mokslo ekosistemos dalimi.

Santrumpos

ADD	Atvirųjų duomenų direktyva
AGA	Anotuotas dotacijos susitarimas (angl. <i>Annotated Grant Agreement</i>)
AM	atvirasis mokslas
AP	atviroji prieiga
API	taikomųjų programų sąsaja (angl. <i>Application Programming Interface</i>)
BDAR	Bendrasis duomenų apsaugos reglamentas
DA	duomenų aktas
DCC	duomenų citavimo korpusas (angl. <i>Data Citation Corpus</i>)
DI	dirbtinis intelektas
DOI	skaitmeninis objekto identifikatorius (angl. <i>Digital Object Identifier</i>)
DRI	Airijos skaitmeninė saugykla (angl. <i>Digital Repository of Ireland</i>)
DVA	Duomenų valdymo aktas
DVP	duomenų valdymo planas
eLABa	Lietuvos akademinės elektroninės bibliotekos informacinė sistema (eLABa)
EIF	Europos sąveikumo pagrindai (angl. <i>European Interoperability Framework</i>)
EK	Europos Komisija
EMTE	Europos mokslinių tyrimų erdvė
EOSC	Europos atvirojo mokslo debesis (angl. <i>European Open Science Cloud</i>)
EOSC-IF	EOSC sąveikumo sistema (angl. <i>EOSC Interoperability Framework</i>)
ERIC	Europos mokslinių tyrimų infrastruktūros konsorciumas (angl. <i>European Research Infrastructure Consortium</i>)
ES	Europos Sąjunga
FDO	FAIR skaitmeniniai objektai (angl. <i>FAIR Digital Objects</i>)
GTC	Gamtos tyrimų centras
GTM	gamtos, technologijos, medicinos ir sveikatos, žemės ūkio mokslai
HPC	aukšto našumo skaičiavimų infrastruktūra (angl. <i>High-performance computing</i>)
HSM	humanitariniai ir socialiniai mokslai
IEA	Sąveikios Europos aktas (angl. <i>Interoperable Europe Act</i>)
IKT	informacinės ir komunikacinės technologijos
IMC	Valstybinio mokslinių tyrimų instituto Inovatyvios medicinos centras
IT	informacinės technologijos
KTU	Kauno technologijos universitetas

KU	Klaipėdos universitetas
LAMMC	Lietuvos agrarinių ir miškų mokslų centras
LKA	Generolo Jono Žemaičio Lietuvos karo akademija
LLTI	Lietuvių literatūros ir tautosakos institutas
LMT	Lietuvos mokslo taryba
LMTA	Lietuvos muzikos ir teatro akademija
LR	Lietuvos Respublika
LSMU	Lietuvos sveikatos mokslų universitetas
LVK	Lietuvos verslo kolegija
MRU	Mykolo Romerio universitetas
MSI	mokslo ir studijų institucijos
MTD	mokslinių tyrimų duomenys
MTEPI	moksliniai tyrimai, eksperimentinė plėtra ir inovacijos
MTI	mokslinių tyrimų infrastruktūros
MVDDA	metinis vidutinis mokslininkų visos darbo dienos atitikmuo
NORF	Nacionalinis atvirojo mokslo forumas (angl. <i>National Open Research Forum</i>)
NVI	Nacionalinis vėžio institutas
ODRL	atviroji skaitmeninių teisių kalba (angl. <i>Open Digital Rights Language</i>)
PID	nekintantis identifikatorius (angl. <i>Persistent identifier</i>)
RAISE	Europos dirbtinio intelekto mokslo išteklių centras (angl. <i>Resource for AI Science in Europe</i>)
RML	teisių žymėjimo kalba (angl. <i>Rights Markup Language</i>)
SLA	paslaugų lygio susitarimas (angl. <i>Service Level Agreement</i>)
SSGG	stiprybių, silpnybių, galimybių ir grėsmių analizė (angl. <i>SWOT</i>)
ŠMSM	Lietuvos Respublikos švietimo, mokslo ir sporto ministerija
UML	unifikuota modeliavimo kalba (angl. <i>Unified Modeling Language</i>)
URL	unifikuotas interneto išteklių adresas (angl. <i>Uniform Resource Locator</i>)
VDA	Vilniaus dailės akademija
VDU	Vytauto Didžiojo universitetas
VILNIUS TECH	Vilniaus Gedimino technikos universitetas
VIKO	Vilniaus kolegija
VU	Vilniaus universitetas

Įvadas

Atvirasis mokslas (toliau – AM) – vienas svarbiausių šiuolaikinės mokslinių tyrimų ekosistemos veikimo elementų, skatinančių tvarų keitimąsi mokslo rezultatais ir pakartotinį jų naudojimą, mokslininkų bendradarbiavimą, piliečių įsitraukimą į mokslą. AM rekomendacijos, gairės ir strategijos Europos Sąjungoje (toliau – ES) apima ne tik atvirosios prieigos (toliau – AP) prie mokslinių publikacijų, duomenų, metodų, infrastruktūrų aspektus, bet ir glaudžiai siejasi su gerokai platesnėmis mokslo politikai ir strategijai svarbiomis sritimis – mokslinių tyrimų kultūra, mokslo komunikacija, visuomenės pasitikėjimu mokslu, mokslo vertinimo reforma, mokslinių tyrimų etika ir sąžiningumu, mokslininkų įgūdžių stiprinimu ir mokslinės karjeros plėtra. Kaip pažymi tarptautinė asociacija „Mokslo Europa“, vienijanti daugiau nei 40 pagrindinių mokslinius tyrimus finansuojančių ir vykdančių Europos institucijų, AM iniciatyvų tikslai – stiprinti mokslo bendruomenės bendradarbiavimą ir supažindinti visuomenę su mokslo žinių kūrimo procesais, mokslo vertinimu ir sklaida. Siekiant šių tikslų, ne tik didinamas mokslinių tyrimų pajėgumas spręsti šiuolaikinius iššūkius visuomenei, bet ir suteikiama papildoma nauda visuomenei tuo metu, kai tam tikroje šalyse auga visuomenės nepasitikėjimas mokslu ir politinis spaudimas, ribojantis mokslinių tyrimų laisvę².

Nepaisant akivaizdžios AM principų svarbos ir atnešamos naudos, jų taikymas sukelia ir nemažai iššūkių, susijusių su ekonominiais, finansiniais, teisiniais aspektais, šiuo metu paplitusiais mokslo vertinimo ir publikavimo modeliais, publikacijų ir duomenų rinkinių atvėrimo rizikomis ir grėsmėmis, intelektinės nuosavybės teisių bei etikos klausimais. Dažnai minimos kliūtys AM iniciatyvoms įgyvendinti, ypač mokslinių tyrimų duomenims (toliau – MTD) atverti, yra susijusios su tam pritaikytos mokslo duomenų e. infrastruktūros modernizavimo poreikiu ir su kvalifikuotų duomenų valdytojų (angl. *data stewards*) trūkumu, jų rengimo programų būtinybe (plačiau apie tai žr. 3 skyrių), mokslo bendruomenės narių suinteresuotumu bei mokslo sričių specifika (plačiau apie AM iššūkius žr. Europos universitetų asociacijos AM veiksmų planą (European University Association, 2022).

Europos atvirojo mokslo debesis (angl. *European Open Science Cloud*³) (toliau – EOSC) yra Europos Komisijos (toliau – EK) inicijuojama ir plėtojama federuota, atvira ir patikima daugiadalykė skaitmeninė aplinka, kurios tikslas – sukurti virtualiąją erdvę mokslininkams saugoti, valdyti, analizuoti ir pakartotinai naudoti MTD. Tokia aplinka ne tik skatina mokslinių tyrimų pažangą bei inovacijas, didesnę investicijų į mokslą grąžą bei bendradarbiavimą tarp mokslininkų, mokslo institucijų ir mokslo sričių, bet ir reikšmingai prisideda siekiant stiprinti visuomenės pasitikėjimą mokslu bei Europos mokslinių tyrimų ekosistemos konkurencingumą ir vieningumą.

EOSC taip pat siekiama stiprinti sąveiką ir bendradarbiavimą tarp mokslinių tyrimų infrastruktūrų (toliau – MTI) visoje Europoje, užtikrinti prieigą prie MTD ir su jais susijusių paslaugų bei skatinti AM principų įgyvendinimą tarp tyrėjų, mokslo ir studijų institucijų (toliau – MSI) ir verslo. Vienas iš esminių EOSC strateginių siekių – plėtoti federuotų duomenų ir paslaugų ekosistemą, kurioje skirtingi išteklių būtų tarpusavyje sąveikūs ir grindžiami bendrais standartais. Integruodamas įvairias duomenų saugyklas ir gerindamas duomenų dalijimosi praktiką, EOSC siekia sudaryti prielaidas nuosekliam mokslinių tyrimų ciklo palaikymui – nuo duomenų generavimo iki jų publikavimo ir pakartotinio naudojimo. EOSC, kaip bendroji Europos tyrimų ir inovacijų duomenų erdvė, taip pat siejamas su kitomis sektorių duomenų erdvėmis pagal ES duomenų strategiją.

EOSC veikimo pagrindą sudaro įsipareigojimas laikytis surandamumo, prieinamumo, sąveikumo ir pakartotinio naudojimo principų (angl. *Findable, Accessible, Interoperable, Reusable*) (toliau – FAIR). Tuo užtikrinama, kad mokslinių tyrimų rezultatai būtų lengvai randami ir naudojami įvairių suinteresuotųjų šalių (plačiau apie FAIR principus žr. 2.2.1 poskyrį). EOSC siekia standartizuotų duomenų ir metaduomenų formatų, būtinų skirtingų šalių ir mokslo sričių atstovams bendradarbiauti. Sukurdamas struktūruotą aplinką tvarkyti MTD, EOSC siekia stiprinti mokslinį bendradarbiavimą, mažinti mokslinių tyrimų dubliavimąsi ir galiausiai sutrumpinti kelią į mokslinius atradimus.

Be to, EOSC infrastruktūra numatyta kaip priemonė, sudaranti sąlygas kurti ir plėtoti naujas MTD valdymo, prieigos ir naudojimo paslaugas. Sudarydamas prielaidas mokslininkams dalytis MTD ir jais naudotis, EOSC gali prisidėti prie didesnio Europos mokslinių tyrimų matomumo, jų rezultatų sklaidos ir AM principų įgyvendinimo.

² <https://scienceeurope.org/our-priorities/open-science/>

³ <https://eosc.eu/eosc-about>

EOSC taip pat siejamas su didesniu mokslinių tyrimų skaidrumu ir rezultatų kartojamumu. Kartu ši iniciatyva orientuota į kai kurių aktualių mokslinių tyrimų sistemos iššūkių sprendimą, įskaitant MTD fragmentaciją, ribotą prieigą prie rezultatų ir poreikį stiprinti bendradarbiavimą tarp mokslininkų Europoje ir už jos ribų.

Sukurdamas bendrą skaitmeninę platformą, EOSC taip pat siekia užpildyti spragas tarp įvairių mokslinių tyrimų sričių ir nacionalinių infrastruktūrų, formuodamas labiau integruotą požiūrį į mokslinius tyrimus. Didėjant MTD apimtims, tokios iniciatyvos svarba siejama su poreikiu užtikrinti veiksmingesnį MTD valdymą, prieigą, pakartotinį naudojimą ir vienodais standartais grindžiamą sąveikumą. Pažymėtina, kad dirbtinio intelekto (toliau – DI) ir mašininio mokymosi integravimas į EOSC sistemą dar labiau padidina jos potencialą, leidžia atlikti pažangią MTD analizę. Tikimasi, kad tolesnė EOSC plėtra prisidės prie mokslinių tyrimų ekosistemos stiprinimo, skatins veiksmingą ir savalaikį MTD naudojimą ir dalijimąsi bei įvairiakryptį bendradarbiavimą.

EOSC istorija ir perspektyva

EOSC iniciatyva buvo oficialiai pradėta 2016 m. kaip dalis platesnės EK strategijos, skirtos AM skatinti ir mokslinių tyrimų bendradarbiavimui stiprinti visoje Europoje. EOSC sukūrimą paskatino supratimas, kad daugėjant MTD reikalinga nuosekli jų dalijimosi ir valdymo sistema. Pirminėse diskusijose buvo pabrėžta, kad MTD yra išsibarstę po įvairias saugyklas ir platformas, o tai trukdo prieinamumui ir sąveikai. Todėl EK inicijavo konsultacijų ir seminarų su akademinės bendruomenės, pramonės ir vyriausybinių institucijų atstovais seriją, siekdamą apibrėžti EOSC viziją ir tikslus. EOSC buvo kuriamas kaip turimų MTD (drauge su publikacijomis ir programine įranga) infrastruktūrų federacija, siekianti sukurti FAIR duomenų ir su jais susijusių paslaugų tinklą (European Commission, 2024). Pradinėje EOSC kūrimo fazėje 2016–2018 m. buvo paremti 35 projektai, 2018 m. buvo sukurta trišalė EOSC valdymo struktūra, susidedanti iš ES, atstovaujamos EK, EOSC asociacijos bei valstybių narių ir asocijuotųjų šalių, atstovaujamų dalyvaujant EOSC valdymo taryboje. Vėliau iš šių grupių buvo suformuotos valdymo ir įgyvendinimo tarybos bei socialinių partnerių forumas, atskiro projekto teisėmis įsteigtas sekretoriatas. Šio pereinamojo laikotarpio EOSC valdymo organai įsteigė EOSC asociaciją, palaikančią partnerystės idėją.

EK parengė EOSC kelrodį⁴, kuris apėmė pirmąjį įgyvendinimo etapą 2018–2020 m., dar galiojant programai „Horizontas 2020“, kuris galioja taip pat ir naujosios programos „Europos horizontas“ 2021–2027 metų laikotarpiui – antrajai įgyvendinimo fazei. Jame EOSC federacijos įkūrimas buvo numatytas kaip esminis struktūrinis principas.

2020 m. buvo oficialiai įsteigta EOSC vykdomoji valdyba ir sukurta EOSC asociacija⁵ kaip federacinė ekosistema, integruojanti įvairias duomenų infrastruktūras ir paslaugas. Ji tapo viena iš trijų trišalio valdymo⁶ (angl. *tripartite governance*) narių – EK, iš valstybių narių ir asocijuotųjų šalių atstovų sudarytos valdymo tarybos (angl. *EOSC Steering Board*) – ir mokslinės bendruomenės (MSI) tikrųjų narių ar stebėtojų. Trišalė valdymo schema buvo parinkta siekiant užtikrinti ES ir nacionalinių politikų darną.

Iniciatyvai vystantis, finansavimas buvo skiriamas vykdant įvairias Europos programas, įskaitant „Horizontas 2020“, siekiant remti EOSC portalo ir susijusių paslaugų plėtrą. Iki 2021 m. EOSC išplėtė savo veiklos sritį ir įtraukė daug suinteresuotųjų šalių. Tai leido įgyvendinti bandomuosius projektus ir užmegzti partnerių ryšius. Juos plėtojant metu buvo išbandyti įvairūs EOSC infrastruktūros aspektai, įskaitant MTD sąveiką, paslaugų integraciją ir vartotojų įtrauktį. Šie projektai buvo labai svarbūs siekiant įrodyti, kad įmanoma sukurti federacinę MTD aplinką, pritaikytą moksliniams tyrimams.

Trečiajame XXI a. dešimtmetyje EOSC tapo labiau struktūrizuotu ir veikiančiu subjektu, o EOSC asociacijai teko pagrindinis vaidmuo valdymo ir suinteresuotųjų šalių įtraukimo srityje. Ši asociacija apima nacionalinių mokslinių tyrimų organizacijų, universitetų ir privataus sektoriaus subjektų atstovus, užtikrinančius, kad EOSC atspindėtų įvairius mokslinių tyrimų bendruomenės poreikius. EOSC portalas buvo sukurtas kaip centrinis prieigos taškas, kur mokslininkai gali rasti įvairių paslaugų, duomenų rinkinių, įrankių ir jais naudotis, taip dar labiau didinant bendradarbiavimo galimybes debesyje. EOSC taip pat pabrėžė mokymo ir paramos mokslininkams svarbą,

⁴ <https://eosc.eu/eosc-about/eosc-macro-roadmap>

⁵ <https://eosc.eu/>

⁶ <https://eosc.eu/tripartite-collaboration>

siekiant palengvinti įgūdžių ugdymą MTD valdymo praktikoje ir AM metodikos vystymo srityje. 2025 metų pabaigoje per EOSC simpoziumą buvo įkurta EOSC federacija. Jos organizacinė struktūra, veikimo principai ir techniniai reikalavimai pristatyti EOSC federacijos vadove (EOSC Association, 2025).

EOSC federacijos vadovas apima nuosekliai susietas temas, kurios padeda suprasti federacijos veikimo principus ir reikalavimus: pirmame skyriuje pristatomi pagrindai – federacijos tikslas, pagrindiniai naudotojai, pranašumai ir sėkmės kriterijai; antrame aptariama valdymo ir teisinė struktūra, komunikacija su kitomis iniciatyvomis bei rizikų valdymas, taip pat nustatomos federacijos ir galimų mazgų valdymo funkcijos. Trečiame skyriuje apibrėžiama operacinė struktūra ir reikalavimai, kuriuos turi atitikti EOSC mazgai, ketvirtame nuodugnai nagrinėjama jų architektūra, paslaugų integravimas ir sąveika su ES EOSC mazgu. Penktame skyriuje pateikiama FAIR duomenų ir mokslinių išteklių, galimų teikti per federaciją, klasifikacija, šeštame skyriuje aprašomi mazgų įtraukimo į federaciją ir naujų paslaugų įtraukimo procesai.

2026 m. vyksta intensyvios konsultacijos su EOSC asociacijos nariais ir stebėtojais dėl naujojo ES finansinio laikotarpio perspektyvų: EOSC valdymo struktūros ir finansavimo modelių, EOSC vietos bendrojoje ES duomenų erdvėje. EOSC siekia išlaikyti partneryste grindžiamą struktūros modelį, paremtą trišale valdysena, apimančia EK, valdymo tarybą ir mokslinių tyrimų bendruomenę per EOSC asociaciją. Kartu matyti ir tam tikras EOSC fokuso poslinkis: greta pirminių tikslų didinti MTD ir paslaugų pasiekiamumą, sąveikumą, FAIR principų įgyvendinimą, kompetencijų stiprinimą ir sektorių bendradarbiavimą, vis labiau akcentuojami mokslinių tyrimų saugumo, duomenų ir skaitmeninio suvereniteto, ilgalaikio finansinio tvarumo, federuotos prieigos prie duomenų, įrankių ir paslaugų, taip pat DI reikmėms tinkamų, patikimų ir atsakingai pakartotinai naudojamų duomenų klausimai.

EOSC struktūra ir plėtra

EOSC įgyvendinama kaip federacinė infrastruktūra – sistemų sistema, jungianti kelis autonominius infrastruktūros vienetus, vadinamuosius mazgus (angl. *nodes*), kurie suteikia prieigą prie FAIR principus atitinkančių MTD, skaičiavimo paslaugų, įrankių ir susijusių išteklių visoje Europoje.

Kiekvienas EOSC mazgas veikia kaip patekimo į EOSC federaciją garantas, leidžiantis vartotojams, ypač tyrėjams, išsaugoti, naudotis, apdoroti, analizuoti ir pakartotinai naudoti mokslinių tyrimų rezultatus (pvz., duomenis, leidinius ir programinę įrangą), laikantis FAIR principų. EOSC konstrukciniai blokai yra paremti naujausiomis žiniomis, prisitaikant prie įsitvirtinusių ir gerai veikiančių struktūrų. Šiuo metu EOSC perėjo iš parengiamojo etapo į veiklos režimą: EK užsakytos paslaugos EOSC platformai palaipsniui pradeda veikti. Europos lėšomis remiamos įvairių mokslo sričių bendruomenės ir nacionalinės iniciatyvos brandina savo infrastruktūras ir paslaugas, kad pasiektų EOSC reikiamą lygį.

Lietuva taip pat dalyvauja EOSC trišalės valdysenos procese: EOSC valdymo tarybos lygmeniu šaliai atstovauja du deleguoti atstovai iš Lietuvos Respublikos švietimo, mokslo ir sporto ministerijos (toliau – ŠMSM) ir Lietuvos mokslo tarybos (toliau – LMT), o instituciniu lygmeniu Vytauto Didžiojo universitetas (toliau – VDU) ir Vilniaus universitetas (toliau – VU) EOSC asociacijoje dalyvauja stebėtojų teisėmis.

Lietuvos įsijungimo į EOSC galimybių į(si)vertinimo svarba

Lietuvos įsijungimo į EOSC klausimas yra glaudžiai susijęs su AM įgyvendinimo iššūkiais, su kuriais susiduria ir kitos Europos šalys. Siekdama didesnio suderinamumo su tarptautinėmis AM praktikomis, Lietuva taip pat susiduria su poreikiu stiprinti MTD valdymo, prieigos, sąveikumo ir pakartotinio naudojimo prielaidas. 2022 m. LMT užsakymu parengtoje AP prie mokslo publikacijų ir duomenų gairių taikymo praktikos ir pasiektos pažangos vertinimo analizėje (Tautkevičienė et al., 2022) įvardijami sunkumai, su kuriais susiduria MSI, bei siūlomos rekomendacijos leido susidaryti vaizdą apie AM plėtros nacionaliniu lygmeniu poreikius, finansavimo, mokslo publikacijų ir duomenų atvėrimo iššūkius, vertinimo sistemos keitimo bei AM principų įgyvendinimo stebėsenos būtinybę. 2021–2030 metų nacionalinio pažangos plano 1 strateginiu tikslu – „[p]ereiti prie mokslo žiniomis, pažangiosiomis technologijomis, inovacijomis grįsto darnaus ekonomikos vystymosi ir didinti šalies tarptautinį konkurencingumą“ (Lietuvos Respublikos Vyriausybė, 2020) – pabrėžiama visos valstybės skaitmeninimo ir technologinio atsinaujinimo svarba aukštos pridėtinės vertės produktų kūrimui. Lietuvos MTD ekosistemos vystymo ir integracijos į

EOSC aspektai minimi ir XVIII Lietuvos Respublikos Vyriausybės programos nuostatų įgyvendinimo veiksmų plane, įskaitant 1.5.4 veiksmą – „[p]arengti priemonių rinkinį, siekiant sudaryti sąlygas tyrėjams dalytis Lietuvos mokslo rezultatais per Europos atvirojo mokslo debesį ir naudotis tarptautiniais mokslo ištekliais (tarptautinėmis mokslo duomenų bazėmis ir infrastruktūromis), ir pradėti šias priemones įgyvendinti“ (Lietuvos Respublikos Vyriausybė, 2021, p. 12).

Lietuvos atveju įsijungimas į EOSC sudarytų prielaidas tyrėjams plačiau naudotis Europos MTD ištekliais, paslaugomis ir skaičiavimo pajėgumais, kurių išlaikymas nacionaliniu lygmeniu būtų sudėtingas ir brangus. Formuojant nacionalinį mazgą tikslinga apžvelgti ir įvertinti MTD išteklius MSI, jų atitikimą FAIR principams, naudojamumą ir kitus svarbius AM aspektus. Toks procesas taip pat sudarytų prielaidas nacionaliniu mastu nuosekliau koordinuoti ir standartizuoti MTD valdyti skirtus išteklius, stiprintų sąveikumą tarp atskirų MSI kuriamų duomenų išteklių. Įsijungimas į EOSC galėtų prisidėti prie didesnio Lietuvos mokslo atvirumo, mokslinių tyrimų rezultatų pakartotinio naudojimo, tarptautinio matomumo ir poveikio, taip pat prie MSI duomenų valdymo gebėjimų ir infrastruktūrų stiprinimo. Didesnis MTD atvėrimas ir prieinamumas taip pat galėtų prisidėti prie skaidresnių mokslinių tyrimų praktikų ir sudarytų palankesnes sąlygas mokslinių tyrimų rezultatų patikrai bei pakartojamumui.

Lietuvai toliau įgyvendinant Nacionalinio pažangos plano nuostatas ir planuojant ES investicijas į AM, kilo poreikis parengti galimybių studiją (toliau – GS), kurioje išsamiai įvertinama esama Lietuvos MTD ekosistema ir galimybės integruotis į EOSC.

Galimybių studijos tikslas ir uždaviniai

Tikslas – taikant sekvencinį mišrių metodų dizainą (teorinę-normatyvinę analizę, empirinį esamos situacijos tyrimą, strateginį scenarijų modeliavimą) identifikuoti būtinus organizacinius ir technologinius pokyčius, reikalingus Lietuvai įsijungti į EOSC, sukurti Lietuvos pasirengimo ir įsijungimo į EOSC alternatyvas, jų scenarijus, jų įgyvendinimo modelius ir architektūrinius modelius, suformuoti įrodymais grįstas investicijų kryptis.

Šiam tikslui pasiekti buvo numatyti šie uždaviniai:

- **Atlikti sisteminę nacionalinių ir tarptautinių strateginių dokumentų, teisės aktų, rekomendacijų ir gerosios praktikos pavyzdžių analizę**, apimančią AP, MTD valdymo, duomenų valdytojų rengimo, MTD standartizavimo, MTD archyvavimo ir tarptautinės integracijos procesų sritis.
- **Išanalizuoti ir apibendrinti EOSC dokumentus, EK gaires, ŠMSM ir kitų susijusių institucijų dokumentus**, reglamentuojančius minėtas sritis, ir atlikti jų lyginamąją analizę, siekiant nustatyti esamas reguliacines, organizacines ir institucines prielaidas bei spragas.
- **Parengti ir pagrįsti tyrimo metodiką**, apibrėžiant Lietuvos MTD ir su jais susijusios infrastruktūros analizės principus, tyrimo instrumentus, vertinimo kriterijus, duomenų rinkimo ir analizės būdus bei pagrindžiant pasirinktų metodų tinkamumą ir validumą.
- **Išsamiai ištirti ir įvertinti esamą Lietuvos MTD valdymo infrastruktūrą**, analizuojant MTD tipus, valdymo lygius, infrastruktūros sprendinius, funkcinis parametrus ir kitus sistemai reikšmingus aspektus.
- **Ištirti ir įvertinti Lietuvos MSI turimas kompetencijas ir organizacinius pajėgumus**, reikalingus MTD tvarkyti, saugoti, atverti, kuruoti ir kitoms susijusioms veikloms, taip pat nustatyti kitus veiksnius, darančius įtaką esamoms MTD valdymo praktikoms.
- **Sukurti esamos MTD valdymo ekosistemos Lietuvoje žemėlapi**, identifikuojant pagrindinius veikėjus, infrastruktūros komponentus, funkcinis sąsajas, duomenų srautus ir sisteminius ryšius.
- **Remiantis atliktos analizės rezultatais nustatyti Lietuvos MSI pasirengimą, poreikį ir potencialą integruotis į EOSC**, identifikuojant pagrindinius barjerus, prielaidas ir plėtros galimybes.
- **Parengti ir pagrįsti Lietuvos įsijungimo į EOSC alternatyvas ir jų scenarijus, jų įgyvendinimo modelius**, atsižvelgiant į identifikuotas sistemines, organizacines, technologines ir kompetencines prielaidas bei ribojimus.
- **Sukurti ir aprašyti Lietuvai įsijungti į EOSC reikalingus architektūrinius modelius**, atspindinčius galimus MTD valdymo, sąveikumo, paslaugų organizavimo, atsakomybių pasiskirstymo ir institucinės sąrangos sprendinius.

- **Nustatyti ir pagrįsti į EOSC įsijungti reikalingas investicijų kryptis**, įvertinant jų įgyvendinimo prielaidas, reikalingus išteklius, prioritetiškumą, galimą poveikį ir tvarumo sąlygas.

Galimybių studijos rengimo etapai

GS buvo rengiama keliais etapais. Pirmojo etapo metu buvo atlikta išsami sisteminė lyginamoji šaltinių analizė, t. y. peržiūrėti, surinkti ir išanalizuoti nacionaliniai bei tarptautiniai strateginiai dokumentai, teisės aktai, rekomendacijos ir gerosios praktikos pavyzdžiai, susiję su AP, MTD valdymo, duomenų valdytojų rengimo, MTD standartų, MTD archyvavimo ir tarptautinės integracijos procesais.

Antrojo etapo metu buvo sukurta ir nuodugnai aprašyta metodika, kuria remiantis atlikta Lietuvos MTD ir jų infrastruktūros analizė, pagrįstas pasirinktų metodų validumas.

Trečiojo etapo metu, siekiant surinkti išsamią informaciją iš tikslinių respondentų grupių apie infrastruktūros būklę, kompetencijų pakankumą, FAIR principų laikymąsi, duomenų valdytojų pareigybių poreikį ir kitus aktualius MTD valdymo aspektus, buvo parengtas išsamus klausimynas ir organizuotos fokusuotosios grupinės diskusijos. Trečiojo etapo metu taip pat sukurta esamos MTD valdymo ekosistemos Lietuvoje analitinis žemėlapis. Surinkti duomenys buvo sistemingai dokumentuoti, struktūrizuoti ir išsamiai tiriami, jais remiantis nustatytas Lietuvos MSI pasirengimas ir potencialas integruotis į EOSC, parengtos galimos EOSC-LT organizacinės valdymo schemos alternatyvos, jų scenarijai ir architektūriniai modeliai, identifikuotos ir pagrįstos investicijų kryptys, reikalingos Lietuvai siekiant įsijungti į EOSC.

GS, jos tikslai ir preliminarūs rezultatai buvo aptarti kelių renginių metu. Pirmasis renginys organizuotas kaip diskusijų forumas. Renginio metu buvo pristatyti GS tikslai, aptarti pirminės MSI apklausos rezultatai, diskutuojama apie Lietuvos įsijungimo į EOSC galimybes. Antrajame renginyje pristatyta parengta GS, atlikta analizė, Lietuvos įsijungimo į EOSC alternatyvos ir jų scenarijai.

1. Galimybių studijos rengimo metodika ir šaltinių atranka

1.1. Bendras metodologinis požiūris ir taikytų metodų visuma

1.1.1. Bendras metodologinis požiūris

Šios GS rengimas grindžiamas sekvenciniu mišrių metodų tyrimo dizainu (angl. *sequential mixed-methods design*), kuris sistemiskai jungia teorinę-normatyvinę analizę, esamos situacijos tyrimą, strateginį alternatyvų ir jų scenarijų modeliavimą. Remiantis Creswell ir Plano Clark (2017), pasirinktas požiūris leidžia integruoti skirtingus duomenų šaltinius trianguliacijos principu (Noble & Heale, 2019) užtikrinant, kad techniniai infrastruktūros parametrai būtų vertinami kartu su socialiniais ir vadybiniais veiksniais. Tyrimas orientuotas į Lietuvos MTD ekosistemos parengtumą integracijai į EOSC. Kiekvienas tyrimo etapas informuoja būsimą etapą, pavyzdžiui, teorinėje analizėje identifikuoti EOSC sąveikumo sistemos (angl. *EOSC Interoperability Framework*) (toliau – EOSC-IF) standartai tampa empirinio tyrimo vertinimo kriterijais, o esamos situacijos tyrimo rezultatai nurodo alternatyvų ir jų scenarijų modeliavimo kryptis.

1.1.2. Teorinė-normatyvinė analizė ir šaltinių parinkimo metodika

Pirmasis etapas apima sisteminę daugiau nei 100 strateginių dokumentų, teisės aktų, techninių standartų, mokslinių publikacijų, leidėjui pateiktų nepublikuotų mokslo darbo versijų (angl. *preprint*), AM politikos dokumentų, organizacijų ir projektų parengtų studijų bei rekomendacijų (pilkosios literatūros), tarptautinių projektų ir organizacijų interneto svetainėse sukauptos informacijos analizę. Analizuojant buvo siekiama apimti keturias

pagrindines dimensijas: 1) ES ir Lietuvos politinę ir teisinę aplinką (pvz., EOSC deklaracija, nacionaliniai atvirojo mokslo planai, žr. 2.1 poskyrį); 2) duomenų principų – FAIR, CARE, TRUST – praktinį taikymą (žr. 2.2 poskyrį); 3) EOSC technologinę architektūrą ir sąveikumo standartus (EOSC-IF) (žr. 2.3 poskyrį); 4) gerosios praktikos pavyzdžius iš kitų ES šalių (žr. 2.4 poskyrį ir 3 skyrių).

Mokslinių publikacijų paieškai buvo naudojamos tarptautinės bibliografinės duomenų bazės „Web of Science“, „Scopus“, „Europe PMC“, „IEEE Xplore“, „Google Scholar“, taip pat AP saugyklos (pvz., „Zenodo“) ir eLABa. Publikacijų paieška ir atranka buvo vykdoma taikant sisteminę paieškos strategiją, apimančią publikacijų antraščių ir santraukų analizę pagal raktinius žodžius.

Informacijos apie nacionalines AM politikas ir AM iniciatyvas paieškai buvo naudojami portaluose „OpenAIRE AMKE“⁷ ir „EOSC Open Science Observatory“⁸ pateikti duomenys. Su EOSC susijusios šalių veiklos aprašomos, remiantis EOSC asociacijos interneto svetainėje⁹ pateikiamų šalių profilių informacija.

Pilkosios literatūros šaltinių atranka buvo vykdoma analizuojant tarptautinių projektų, orientuotų į AM ir MTD valdymą, metodines nuorodas.

Šaltinių paieškos rezultatai buvo papildyti taikant vadinamąjį sniego gniūžtės (angl. *snowballing*) metodą, t. y. analizuojant pirminės paieškos metu identifikuotuose šaltiniuose pateikiamas citatas ir nuorodas ir pagal jas atrenkant papildomus GS aktualius šaltinius.

Taikomas lyginamasis politikos vertinimas (angl. *comparative policy analysis*) Lietuvos reguliacinės aplinkos suderinamumui su ES gairėmis nustatyti. Esminis analitinis pagrindas – EOSC-IF, apimanti teisinį, organizacinį, semantinį ir technologinį lygmenis (European Commission, 2021a). FAIR, CARE ir TRUST principai (Carroll et al., 2020; Lin et al., 2020; Wilkinson et al., 2016) taikomi kaip vertinimo kriterijai duomenų kokybės įvertinti vėliau.

1.1.3. Esamos situacijos analizė

Empirinę tyrimo dalį sudaro trys tarpusavyje susiję komponentai: a) duomenų išteklių inventorizacija; b) valdysenos ir kompetencijos vertinimas; c) kiekybinė apklausa, fokusuotosios grupinės diskusijos ir esamos MTD valdymo ekosistemos Lietuvoje analitinio žemėlapiu sudarymas.

Duomenų išteklių inventorizacijai atlikti taikoma duomenų audito (patikros) metodika (angl. *Data Audit Framework*), siekiant identifikuoti Lietuvos MSI turimą duomenų išteklių (angl. *data assets*). Vertinami trys aspektai:

- **Duomenų apimtys ir tipologija.** Kiekybiškai įvertinami generuojamų duomenų kiekiai (nuo GB iki TB skaleje ir atsižvelgiant į institucinę priklausomybę) ir jų pobūdis – tekstynai, garsynai, jutiklių duomenys, skaitmeniniai vaizdų duomenynai, kt. Tai leidžia prognozuoti infrastruktūros poreikius (skaičiavimo ištekliai, saugyklų talpa).
- **Formatai ir standartizacija.** Analizuojamas atvirųjų (pvz., CSV, JSON, HDF5) ir uždarytųjų formatų santykis. Remiantis Mons et al. (2017), duomenų mašininis apdorojamumas (angl. *machine-actionability*), t. y. gebėjimas automatizuotai skaityti ir interpretuoti duomenis be žmogaus intervencijos, yra ypač svarbus EOSC integracijos veiksnys. Todėl vertinama, kiek naudojami formatai atitinka tarptautinius standartus ir ontologijas.
- **Technologinis pasirengimas.** Inventorizuojamos MSI saugyklos (angl. *data repositories*), jų techninės charakteristikos (talpa, prieigos protokolai), taikomųjų programų sąsajos (angl. *Application Programming Interface*) (toliau – API) ir integracija su tarptautiniais tinklais.

Atliekant valdysenos ir kompetencijų vertinimą, ypatingas dėmesys skiriamas vadinamajai minkštajai infrastruktūrai (angl. *soft infrastructure*), t. y. institucinėms taisyklėms ir žmogiškiesiems ištekliams. Tyrimas remiasi mokslo duomenų valdymo brandos modeliu (angl. *Research Data Management Maturity Model*), vertinant:

⁷ <https://www.openaire.eu/about>

⁸ <https://www.eoscobservatory.eu/home>

⁹ <https://eosc.eu/tripartite-collaboration>

- **duomenų valdysena** (angl. *data governance*): institucinių politikų (duomenų valdymo planai (toliau – DVP), AP mandatai), etikos gairių ir praktinio jų taikymo lygis. Vertinama atitiktis OECD (2021) rekomendacijoms dėl prieigos prie viešai finansuojamų MTD;
- **kompetencijų žemėlapis**: identifikuojamas duomenų valdytojų poreikis ir esamas jų skaičius MSI. Remiantis Wildgaard et al. (2020) kompetencijomis, vertinamas atotrūkis tarp tyrėjų gebėjimų (pvz., gebėjimo parengti DVP) ir EOSC keliamų standartų. Tai sudaro pagrindą nacionalinės mokymų programos rekomendacijoms rengti.

Trečioji empirinio tyrimo dalis – taikyti kiekybinės apklausos ir fokusuotųjų grupinių diskusijų principai – aprašomi 1.2 poskyryje.

Šiai GS taip pat buvo parengtas esamos MTD valdymo ekosistemos Lietuvoje analitinis žemėlapis. Jis sudarytas remiantis viešai prieinamais duomenimis apie Lietuvos MTD valdymo ekosistemą. Duomenų šaltiniais pasitelkti LMT, ŠMSM, MSI viešai skelbiami infrastruktūrų, saugyklų ir informacinių sistemų aprašymai, taip pat „OpenDOAR“ ir „Re3data“ tarptautiniai registrai. Žemėlapyje sudaro 81 mazgas (objektas) ir 189 ryšiai tarp jų. Objektai suskirstyti į šešias kategorijas: įstaigos (29), infrastruktūra (24), duomenys (11), saugyklos (9), sistemos (4) ir registrai (3). Kiekvienas mazgas aprašytas atskiru dokumentu su metaduomenimis (pavadinimas, tipas, identifikatorius) ir „Wiki“ stiliaus nuorodomis į susijusius mazgus. Žemėlapis pateikiamas GS priedų dalyje (žr. 8 priedą).

1.1.4. Analitinė sintezė ir strateginis modeliavimas

Visi empirinio tyrimo metu surinkti duomenys (infrastruktūros apimtys, duomenų formatai, kompetencijų lygis ir valdysenos spragos) buvo apdorojami ir interpretuojami keturiais etapais, užtikrinant perėjimą nuo esamos būklės fiksavimo prie strateginių rekomendacijų formavimo:

1. Duomenų analizė ir sisteminimas. Apklauso duomenys buvo išvalyti, patikrinti ir parengti analizei: peržiūrėtas atsakymų nuoseklumas, pašalinti techniniai netikslumai, suvienodinti atsakymų formatai, atliktas duomenų kodavimas ir validavimas. Statistiniai apklausos duomenys apdoroti atliekant aprašomąją statistiką ir lyginamuosius pjūvius pagal institucijų tipą bei mokslo sritis (naudojant SPSS / R programinę įrangą). Kokybinio tyrimo (fokusuotųjų grupinių diskusijų) duomenys integruoti taikant teminę analizę, tai leido interpretuoti tendencijas per institucinės kultūros ir specifinių barjerų prizmę. Rezultatai iliustruojami vaizdinėmis priemonėmis – grafikais, lentelėmis ir ekosistemos ryšių žemėlapiams.

2. Stiprybių, silpnųjų, galimybių ir grėsmių (angl. SWOT) analizė (toliau – SSGG). Remiantis surinktais duomenimis, atliktas sisteminis Lietuvos situacijos palyginimas su EOSC koncepcija, FAIR principais ir EOSC-IF idealiojo modelio reikalavimais. Strateginiam vertinimui taikyta SSGG analizė (Dyson, 2004), kuria remiantis identifiukuotos Lietuvos integracijai į EOSC reikšmingos vidinės stiprybės ir silpnybės bei išorinės galimybės ir grėsmės. Vertinant greito ir realistiško įsijungimo prielaidas, ypatingas dėmesys skirtas toms stiprybėms ir galimybėms, kurias galima išnaudoti artimiausiu laikotarpiu, o identifiukuotos silpnybės ir grėsmės vertintos kaip rizikos ir tolesnio stiprinimo sritys. Vadovaujantis šiuo vertinimu suformuluotos trys integracijos alternatyvos ir jų scenarijai: 1) decentralizuotas vystymasis (*status quo* – išlaikant esamą fragmentišką vystymosi modelį be sisteminės integracijos, pvz., integracija vyksta fragmentiškai per konkrečios nacionalinės MTI integraciją į Europos mokslinių tyrimų infrastruktūros konsorciumo (angl. *European Research Infrastructure Consortium*, toliau – ERIC) tipo MTI; 2) nacionalinis mazgas; 3) regioninis mazgas (plačiau žr. 5 skyrių).

3. Finansinis pagrindimas ir investicijų poreikio vertinimas. Alternatyvos vertintos taikant sąnaudų ir naudos analizės (angl. *Cost-Benefit Analysis*) principus. Metodika numato ne tik tiesioginių investicijų (techninių sprendimų, administravimo, duomenų valdytojų parengimo) poreikio nustatymą, bet ir socioekonominės bei mokslinės naudos šaltinių apibendrinimą (pvz., laiko sutaupymas, pakartotinis duomenų naudojimas, inovacijų potencialas). Investicijų poreikis skirstomas į trumpalaikį (projekto įgyvendinimo) ir tęstinį (tvarumo užtikrinimo), vertinant galimus finansavimo šaltinius (ES fondai, nacionalinis biudžetas).

4. Įgyvendinimo plano ir organizacinės struktūros projektavimas. Strateginis veiksmų planas (angl. *roadmap*) kurtas remiantis loginės sekos principu, išskiriant trumpalaikes, vidutinės trukmės ir ilgalaikes veiklas. Siekiant už-

tikrinti plano įgyvendinamumą, taikytas RACI (angl. *Responsible, Accountable, Consulted, Informed*) atsakomybių matricos modelis, identifikuojantis pagrindinių suinteresuotųjų šalių (ŠMSM, LMT, MSI, kt.) vaidmenis ir funkcijas nacionalinėje MTD valdysenos struktūroje. Rizikos vertintos taikant tikimybės ir poveikio matricą pagal ISO 31000 rizikų valdymo standartą.

1.1.5. Kokybės užtikrinimas ir etika

Siekiant užtikrinti tyrimo validumą, taikyta metodų trianguliacija (kiekybinių ir kokybinių duomenų derinimas bei dokumentų analizė), grupinis ekspertinis vertinimas ir viešosios konsultacijos (du forumai su mokslo bendruomene, politikos formuotojais, plačiąja visuomene GS rengimo procesui ir rezultatams aptarti).

Etiniai aspektai griežtai atitinka Bendrojo duomenų apsaugos reglamento (toliau – BDAR) reikalavimus: visi apklausų duomenys nuasmeninti, fokusuotųjų grupinių diskusijų metu surinkta medžiaga nuasmeninta, surinkti duomenys naudoti tik apibendrinta forma ir išskirtinai šios GS tikslams, taip užtikrinant konfidencialumą ir duomenų apsaugą.

1.2. Apklausos ir fokusuotosios grupinės diskusijos

Rengiant GS buvo atlikti keturi empiriniai tyrimai, siekiant ištirti esamą Lietuvos MTD infrastruktūros situaciją bei egzistuojančias Lietuvos MSI ir mokslininkų MTD valdymo ir tvarkymo praktikas bei kompetencijas. Pirmas tyrimas buvo trumpa žvalgomoji institucijų apklausa, kurios tikslas – išsiaiškinti bendrąją Lietuvos institucijų situaciją ir atrinkti institucijas išsamiajai apklausai. Antrame tyrime – išsamiojoje institucijų apklausoje – buvo nuodugniai tiriama Lietuvos MTD infrastruktūros situacija bei MSI duomenų valdymo ir tvarkymo praktikos bei kompetencijos. Trečiame tyrime, siekiant išsiaiškinti Lietuvos mokslininkų patirtis dirbant su MTD ir naudojantis MTI, buvo atliekama jų apklausa. Ketvirtas tyrimas įgyvendintas kaip dvi fokusuotosios grupinės diskusijos su antrame tyrime dalyvavusioms MSI atstovaujantiems atsakingais asmenimis, kuriuo siekta pristatyti ankstesnių tyrimų rezultatus ir surinkti MSI atgalinį ryšį.

1.2.1. Apklausų metodika

Pirmiausia, siekiant išsiaiškinti esamą Lietuvos MTD infrastruktūros situaciją ir Lietuvos MSI MTD valdymo ir tvarkymo praktikas bei kompetencijas, buvo atliekama žvalgomoji Lietuvos MSI apklausa. Apklausa vykdyta 2025 m. gegužės 6–16 d. per internetinę „Microsoft Forms“ platformą. Pagrindinis šios apklausos tikslas – surinkti bendrąją informaciją apie Lietuvos MSI taikomas MTD valdymo praktikas, institucinius pajėgumus ir poreikius. Kitas svarbus apklausos tikslas – identifikuoti MSI kontaktinius asmenis, į kuriuos bus kreipiamasi dėl išsamiosios MSI apklausos. Apklausos pradžioje el. paštu buvo kreiptasi į visas Lietuvos MSI prašant užpildyti elektroninį klausimyną (pirmojo tyrimo el. kreipimosi laiškas ir klausimynas pateikti 1 ir 2 prieduose, išsamūs apklausos rezultatai ir jų analizė – 4.1 poskyryje).

Antrajame tyrime jau siekta išsamiai išsiaiškinti esamą Lietuvos MTD infrastruktūros situaciją bei Lietuvos MSI duomenų valdymo ir tvarkymo praktikas bei kompetencijas. Taip pat atlikta Lietuvos MSI apklausa, vykdyta nuo 2025 m. liepos 8 d. iki rugsėjo 30 d. per internetinę „Microsoft Forms“ platformą. Pagrindinis šios apklausos tikslas – surinkti nuodugnią informaciją apie Lietuvos MSI taikomas MTD valdymo praktikas, institucinius pajėgumus ir poreikius.

Tyrimo klausimynas suskirstytas į šešias temines sritis, apimančias technologinę infrastruktūrą, MTD politiką ir reglamentavimą, MTD parametrus, žmogiškuosius išteklius, MTD atvėrimo praktiką ir MSI dalyvavimą tarptautinėse iniciatyvose.

1. Technologinė ir programinė infrastruktūra. MSI prašyta įvardinti turimus technologinius išteklius, naudojamus MTD kaupiti, saugoti, apdoroti ir atverti. Klausta, ar ištekliai įtvirtinti MSI norminiuose dokumentuose. Taip pat teirautasi, ar MSI duomenų saugyklos yra sertifikuotos pagal tarptautinius standartus („CoreTrust-Seal“, „nestor Seal“, ISO 16363, ISO 27001), ar planuojamas sertifikavimas, kokiomis kitų organizacijų saugyklomis naudojama ir ar jų naudojimą koordinuoja MSI.

2. Dalyvavimas MTI. MSI prašyta pateikti informaciją apie dalyvavimą tarptautinėse ir nacionalinėse MTI: ar jose generuojami MTD, teikiamos paslaugos, naudojama bendrais ištekliais ir paslaugomis. Taip pat klausta, ar šios infrastruktūros įtrauktos į Lietuvos MTI kelrodį ar kitas tarptautines institucijas.

3. MTD politika, reglamentavimas ir valdymas. MSI prašyta įvardinti ir apibūdinti savo norminius dokumentus: ar jose numatyta MTD valdymo politika, DVP rengimo reikalavimai, archyvavimo ir sunaikinimo tvarka, metaduomenų išsaugojimas. Klausta, ar MTD sunaikinami pagal dokumentuotas procedūras, ar saugomos sunaikinimo ataskaitos, ar reglamentuojama intelektinės nuosavybės politika, susijusi su MTD. Institucijų taip pat teirautasi, kaip vykdoma teisės aktų laikymosi stebėseną, kokie kontrolės mechanizmai taikomi, ar registruojami tyrėjų paviešinti MTD rinkiniai, kaip vykdoma jų stebėseną. Galiausiai institucijų prašyta įvardinti pagrindinius iššūkius, susijusius su MTD valdymu.

4. Sukaupti ir sukuriami MTD. Šioje dalyje institucijų klausta apie MTD saugojimo sąnaudas ir finansavimo šaltinius, turimų MTD apimtis: kiek skirtingų mokslo sričių MTD rinkinių institucijos yra sukopusios, kiek naujų rinkinių kuria per metus, kokia bendra MTD apimtis. Taip pat prašyta įvardyti generuojamų MTD tipus, saugojimo vietas ir įvertinti, kiek MTD ar jų metaduomenų yra atvirai prieinami.

5. Žmogiškieji ištekliai, kompetencijos ir paramos mechanizmai. Institucijų prašyta pateikti informaciją apie darbuotojus, dirbančius su MTD: jų skaičių, pareigas ir etato dalį, skirtą MTD valdymui. Klausta, kokius mokymus institucija rengia tyrėjams, ar MTD valdymo pagrindai integruoti į studijų programas, kokios konsultacijos teikiamos, ir kas jas teikia. Taip pat prašyta apibūdinti teikiamą praktinę pagalbą, pvz., MTD dokumentavimo, saugojimo ar metaduomenų rengimo paslaugas.

6. MTD atvėrimo ir skatinimo praktikos. Institucijų prašyta apibūdinti taikomas skatinimo priemones: ar MTD tvarkymas ir atvėrimas yra aktyviai skatinamas, kokios karjeros, finansinės ar pripažinimo paskatos taikomos. Klausta, ar organizuojami informaciniai renginiai apie MTD valdymą, ar jie yra reguliarūs. Taip pat prašyta įvardyti kitus veiksnius, darančius įtaką institucijų MTD valdymo praktikoms.

Antrosios apklausos pradžioje el. paštu buvo kreiptasi į 33 Lietuvos MSI, kurios pirmajame tyrime nurodė, kad jose kuriami ir valdomi MTD, prašant užpildyti elektroninį klausimyną (antro tyrimo el. kreipimosi laiškas ir klausimynas pateikti 3 ir 4 prieduose, išsamūs apklausos rezultatai ir jų analizė pateikiami 4.1 poskyryje).

Trečiajame tyrime siekiant atskleisti Lietuvos mokslininkų patirtis su MTD ir naudojantis MTI buvo vykdyta mokslininkų apklausa. Apklausa vykdyta 2025 m. spalio 15–29 d. per internetinę „Microsoft Forms“ platformą. Respondentų imtis buvo formuojama remiantis LMT administruojamos priemonės „Mokslininkų grupių projektai“ (MIP) XI, XII ir XIII kvietimų paraiškų duomenimis. Imtis sudaryta ne tik iš finansuotų projektų, bet iš visų pateiktų paraiškų (tiek finansuotų, tiek nefinansuotų), siekiant aprėpti kuo platesnį tyrėjų ratą ir užtikrinti didesnę skirtingų mokslo sričių, krypčių ir institucijų įvairovę. Mokslininkų, kuriems buvo siunčiamas kvietimas dalyvauti apklausoje, sąrašas sudarytas iš visų MIP XI–XIII kvietimų paraiškose nurodytų tyrėjų duomenų. Pradiname etape buvo identifikuotos 2379 paraiškos, iš kurių, sutvarkius įrašus ir pašalinus pasikartojančius asmenis (kai tas pats asmuo paraiškas teikė daugiau nei vieną kartą skirtinguose kvietimuose ar keliuose paraiškų cikluose), suformuotas 15 689 unikalių mokslininkų sąrašas.

Apklausos pradžioje el. paštu buvo kreiptasi į visus atrinktus Lietuvos mokslininkus prašant užpildyti elektroninį klausimyną (trečiojo tyrimo el. kreipimosi laiškas ir klausimynas pateikti 5 priede). Po savaitės (2025 m. spalio 22 d.) buvo išsiųstas priminimas dalyvauti apklausoje (išsamūs apklausos rezultatai ir jų analizė pateikiami 4.2 poskyryje).

Trečiojo tyrimo klausimynas suskirstytas į šešias temines sritis, kurios apima mokslininkų santykį su MTD, jų naudojamą infrastruktūrą, institucines sąlygas, teisinius reikalavimus, motyvaciją, kompetencijų poreikį ir MTD valdymo praktikas.

1.Mokslininkų santykis su MTD, naudojami išteklių ir duomenų tipai. Respondentų prašyta nurodyti, ar jie patys kuria, renka, tvarko, archyvuoja, validuoja MTD, ar naudoja kitų parengtus (mokamus ar nemokamus) MTD. Taip pat klausta, kokia aparatine ir programine įranga jie naudojami, su kokių tipų MTD dirba.

2.Dalijimasis MTD ir naudojamos saugyklos. Ši klausimų grupė buvo skirta išsiaiškinti, ar ir kaip mokslininkai dalijasi MTD, o jei nesidalija, prašyta įvardyti priežastis. Taip pat klausta, kokias saugyklas naudoja saugoti MTD.

3.FAIR reikalavimai ir DVP. Šioje dalyje klausta, kiek mokslininkai žino ir taiko FAIR principus. Taip pat klausta apie DVP rengimą.

4.Institucinės sąlygos, infrastruktūra ir dalyvavimas tarptautinėse iniciatyvose. Ši klausimų grupė padeda nustatyti, ar mokslininkų institucijose sudarytos tinkamos sąlygos tvarkyti MTD: ar yra pakankamai infrastruktūros, specialistų, finansavimo, programinės įrangos ir mokymų. Jei sąlygos netinkamos, prašyta nurodyti priežastis. Klausta, kiek ir kokia apimtimi mokslininkai įsitraukia į tarptautines MTI.

5.Kompetencijų poreikiai, institucijų skatinimas ir motyvacija. Šioje dalyje klausta, kokių kompetencijų, susijusių su MTD, mokslininkams trūksta, su kokiais sunkumais jie susiduria. Taip pat klausta, ar institucijos skatina MTD kaupimą, archyvavimą ir atvėrimą, kokios paskatos taikomos. Prašyta įvertinti, kurie veiksniai labiausiai motyvuoja saugant, tvarkant ir atveriant MTD, ar naudojami duomenų valdytojų paslaugomis.

6.MTD apimtis. Šioje dalyje klausta apie sukurtą MTD apimtį: kiek MTD rinkinių mokslininkai kuria per metus pagal MTD tipą, ir kokios apimties paprastai būna vienas MTD rinkinys.

Baigiamojoje klausimyno dalyje taip pat rinkta kai kuri informacija apie respondentus: tyrėjų karjeros pakopa (pradedantysis, patvirtintas, pripažintas ar pirmaujantis tyrėjas), darbo patirtis metais, mokslo sritis ir kryptys.

1.2.2. Fokusuotųjų grupinių diskusijų metodika

Ketvirtasis tyrimas įgyvendintas surengiant dvi fokusuotąsias grupines diskusijas su antrajame tyrimo dalyvavusių MSI atsakingais atstovais, šiame etape taikant iteracinę duomenų rinkimo ir analizės logiką. Šis metodas pasirinktas siekiant pagilinti ankstesnių trijų tyrimų rezultatus, užpildyti analizėse išlikusias informacines spragas, patikslinti klausimus, į kuriuos ankstesniais tyrimo etapais nepavyko pateikti pakankamai aiškių atsakymų, ir papildomai pagrįsti ar sustiprinti tam tikras preliminarias išvadas, svarbias GS bei tolesniam jungimosi į EOSC planavimui.

Į ketvirtąjį tyrimą buvo kviečiami institucijų, dalyvavusių antrajame (institucijų) tyrimo (iš viso 22) atstovai. Šis tyrimas įgyvendintas kaip dvi nuotolinės (per „MS Teams“ platformą organizuotos) fokusuotosios grupinės diskusijos su šioms institucijoms atstovaujanciais atsakingais asmenimis. Dalyvauti diskusijose sutiko 15 institucijų, tačiau susitikimuose faktiškai dalyvavo 13 institucijų, 2 institucijos į susitikimus neprisijungė.

Diskusijos vyko 2026 m. balandžio 3 d. 10.00 val. ir 13.00 val. Pirmoje grupėje dalyvavo atstovai iš šių institucijų: SMK aukštosios mokyklos, Vilniaus kolegijos (toliau – VIKO), VU, Vilniaus dailės akademijos (toliau – VDA), Kauno technologijos universiteto (toliau – KTU), Lietuvos istorijos instituto ir Generolo Jono Žemaičio Lietuvos karo akademijos (toliau – LKA). Antroje grupėje buvo atstovai iš Vilniaus Gedimino technikos universiteto (toliau – VILNIUS TECH), Lietuvos sveikatos mokslų universiteto (toliau – LSMU), Lietuvos muzikos ir teatro akademijos (toliau – LMTA), Lietuvos agrarinių ir miškų mokslų centro (toliau – LAMMC), Mykolo Romerio universiteto (toliau – MRU) ir Valstybinio mokslinių tyrimų instituto Inovatyvios medicinos centro (toliau – IMC).

Diskusijos vykdytos tokia tvarka: 1) dalyviams pristatyti jau atliktų trijų tyrimų rezultatai; 2) dalyviams pristatyti tolesni ŠMSM ir LMT planai, susiję su rengiama GS ir jungimusi į EOSC; 3) po pristatymų dalyviai buvo kviečiami reflektuoti pateiktas išvadas, įvardyti jų institucijų patirties požįriui svarbius papildymus, atkreipti dėmesį į galimas spragas, pateikti vertinimus dėl siūlomų tolesnių kryptų. Tokiu būdu surinkta kokybinė informacija buvo naudojama ankstesnių tyrimų rezultatams papildyti, patikslinti ir interpretuoti (plačiau apie fokusuotųjų grupinių diskusijų išvadas žr. 4.3 poskyrį).

1.3. Mokslinių tyrimų duomenų saugyklų identifikavimo ir aprępties vertinimo metodika

Sparčiai augant MTD apimtims, plečiasi ir įvairiems MTD tipams saugoti bei valdyti skirtų saugyklų įvairovė. Deja, išsamus tokių saugyklų sąrašas nėra prieinamas vienoje vietoje – nėra resurso, analogiško domenų vardų sistemai (angl. *Domain Name System*), kuriame būtų kaupiama informacija apie visas internete prieinamas duomenų saugyklas. Dėl šios priežasties išsamiam saugyklų kraštovaizdžiui nustatyti taikytas kelių šaltinių ir kelių paieškos strategijų derinys siekiant identifiкуoti visus MTD rinkinius, taip pat ir visus rinkinius, sukurtus Lietuvos mokslininkų. Šiame tyrime MTD saugyklų sąrašas buvo sudaromas remiantis keliais vienas kitą papildančiais šaltiniais: bendraisiais MTD saugyklų katalogais, mokslo žurnalų rekomendacijomis autoriams, specializuotų duomenų bazių žurnalų publikacijomis, universitetų ir koledžų rekomendacijomis, taip pat duomenų rinkinių ir duomenų citavimo katalogais. Toks kelių šaltinių derinimas pasirinktas siekiant sumažinti atskirų šaltinių aprępties ribotumus ir susidaryti kuo išsamesnį dabartinio duomenų saugyklų kraštovaizdžio vaizdą.

Duomenų saugyklų katalogai

Esamos duomenų saugyklos gali būti identifiкуojamos įvairiais būdais. Nė vienas iš jų savaime neužtikrina visapusiško sąrašo, tačiau derinant kelis paieškos metodus galima susidaryti gana išsamų dabartinio duomenų saugyklų kraštovaizdžio vaizdą.

Siekiant palengvinti mokslininkams saugyklų suradimą, buvo sukurti keli duomenų saugyklų katalogai. Didžiausi tokie katalogai šiuo metu yra du: „Re3Data“¹⁰ (pradėjęs veikti 2012 m.) ir „FAIRsharing“¹¹ (sukurtas 2011). Šie du katalogai minimi daugelyje kitų šaltinių, tokių kaip universitetų duomenų valdymo rekomendacijos, mokslinių žurnalų nurodymai autoriams ir pan. Pagal duomenis, gautus naudojant API, 2025 kovo mėn. (nukėlimas pradėtas 2025-03-11, baigtas 2025-03-12) šiuose kataloguose buvo 4784 („FAIRsharing“; tarp jų 2334 duomenų bazių) ir 3338 (Re3Data) įrašų. „FAIRsharing“ turi šiek tiek daugiau įrašų, bet, skirtingai negu „Re3Data“, „FAIRsharing“ kaupia ir informaciją apie kitus išteklius – formatus, žurnalus, terminijos artefaktus ir pan., todėl duomenų bazių ten yra kiek mažiau.

Šie katalogai yra išsamiausi ir didžiausi duomenų bazių ir duomenų saugyklų sąrašai. Be to, juose pateikiama išsamūs metaduomenys, tarp jų ir šalys, kurios dalyvauja kuriant ir palaikant aprašomas duomenų saugyklas, o tai leidžia įvertinti Lietuvos indėlį į pasaulinę duomenų ekosistemą.

Mokslinių žurnalų rekomendacijos

Be minėtų bendrų duomenų saugyklų katalogų, buvo peržiūrėtos pagrindinių mokslinių žurnalų rekomendacijos autoriams, mininčios duomenų tvarkymo reikalavimus, duomenims skirtas saugykla ir duomenų formatus. Buvo orientuotasi tiek į labiausiai cituojamus bendruosius mokslinius žurnalus leidžiančias leidyklas ir jų žurnalus („Nature“ grupės žurnalai, AAAS (žurnalas „Science“), PNAS, „Cell Press“ žurnalai), tiek į didžiausias pasaulines mokslo draugijas: IUCr, ACS, RSC. Pažymėtina, kad didelė dalis žurnalų redakcijų šiuo metu rekomenduoja pasirinkti tinkamas saugyklas duomenims talpinti iš „FAIRSharing“ ir „Re3Data“ katalogų, o tai patvirtina

¹⁰ <https://www.re3data.org/>

¹¹ <https://fairsharing.org/>

svarbų šių katalogų statusą ir išsamumą; prieš maždaug dešimt metų ir seniau patys žurnalai rekomenduodavo konkrečias saugyklas pagrindiniams duomenų tipams (kristalografiniams duomenims, struktūrinės biologijos duomenims, sekoskaitos duomenims, klinikiniais bandymams ir pan.) Šios konkrečios saugyklos buvo įtrauktos į tirtų saugyklų sąrašą, jei jų dar nebuvo „FAIRSharing“ ir „Re3Data“ sąrašuose.

Duomenų bazių žurnalų publikacijos

Ieškant gamtos mokslams skirtų duomenų bazių, buvo peržiūrėti du plačiausiai žinomi specializuoti žurnalai, sukurti mokslinėms duomenų bazėms. Vienas jų, „Oxford“ leidyklos leidžiamas „Database“, iš esmės skirtas straipsniams apie naujas duomenų bazes publikuoti. Kitas žurnalas „Nucleic Acids Research“ (toliau – NAR), nors ir nėra skirtas vien tik duomenų bazėms, kasmet išleidžia specializuotą duomenų bazių apžvalgos numerį (NAR „Database Issue“). Tuo būdu šie žurnalai duoda gana gerą naujų ir veikiančių duomenų bazių apžvalgą. Šių žurnalų pranašumas tas, kad jie abu yra recenzuojami moksliniai žurnalai, tad duomenų bazės, kaip ir jas aprašantys straipsniai, pereina bent minimalias kokybės patikras. Jų trūkumas yra tematinis vienpusiškumas – abu žurnalai, o ypač NAR, skirti su gyvybės mokslais ir bioinformatika susijusioms duomenų bazėms. Nepaisant to, jų analizė leidžia įvertinti didžiųjų pirmiau minėtų katalogų išsamumą surandant, kiek naujų duomenų bazių nebuvo paminėta „FAIRSharing“ ir „Re3Data“ kataloguose ir ekstrapolijuojant šį skaičių visoms gamtos mokslų sritims.

NAR „Database Issue“ tome pateikiamos tuo metu publikuotos duomenų bazės (saugyklos); duomenų bazės atrenkamos pagal jų mokslinę kokybę, stabilumą ir atitiktį mokslo kryptims. 2024 m. leidinyje buvo pateiktos 90 duomenų bazių¹². Iš jų, naudojant unifikotą interneto ištekliaus adresą (angl. *Uniform Resource Locator*) (toliau – URL), nuorodas su protokolu, pakeistu visur į https:, ir pašalinus galinį pasvirąjį brūkšnį, „FAIRSharing“ kolekcijoje randamos 9, o „Re3Data“ – tik 2 duomenų bazės; naudojant DB pavadinimus randamas dar mažesnis šių bendrų DB poaibis. Tenka daryti išvadą, kad dauguma autorių arba publikuoja savo DB aprašymus NAR, arba aprašo jas duomenų saugyklų kataloguose, bet retai daro ir viena, ir kita, o DB pavadinimai skirtinguose DB gali skirtis.

Kiti žurnalai, tokie kaip „Briefings in Bioinformatics“, „BioData Mining“, „Biodiversity Informatics“ ar MDPI žurnalai, publikuoja straipsnius apie duomenų bazes sporadiškai, ir dėl to į analizę įtraukti nebuvo.

Vakarų universitetų ir koledžų rekomendacijos

Papildomas informacijos šaltinis buvo universitetų ir koledžų skelbiamos rekomendacijos dėl duomenų saugyklų pasirinkimo. Šie sąrašai atspindi praktinį akademinį institucijų požiūrį į patikimas ir dažniausiai rekomenduojamas saugyklas. Nors tokie sąrašai paprastai nėra išsamūs ir dažnai kiek dubliuojasi tarpusavyje, jie buvo naudingi kaip papildomas validacijos šaltinis, leidžiantis įtraukti kai kurias praktikoje svarbias, bet kataloguose ne visuomet aiškiai matomas saugyklas. Didžiausias saugyklų sąrašas (220 nuorodų) buvo surastas Simmonso universiteto resursų „Wiki“ puslapyje¹³. Kiti universitetai ir koledžai turi pastebimai mažiau rekomenduojamų duomenų bazių (iki 20), dalis jų cituoja Simmonso universiteto sąrašą. Tad šis sąrašas buvo pasirinktas kaip tipinis universitetų rekomendacijų šaltinis.

Duomenų rinkinių katalogai

Siekiant įvertinti ne tik pačias saugyklas, bet ir jose registruojamų duomenų rinkinių mastą, papildomai naudoti „DataCite“ ir „Crossref“ metaduomenų katalogai. Šie šaltiniai leidžia analizuoti registruotus skaitmeninius objektus, įskaitant duomenų rinkinius, bei jų metaduomenis. „DataCite“ katalogas šiuo požiūriu buvo ypač svarbus dėl platesnės metaduomenų struktūros, įskaitant autorių afiliacijas, kurios sudaro prielaidas identifikuoti su Lietuva susijusius įrašus. Didžiausi ir žinomiausi iš jų yra šiuo metu „DataCite“ (Neumann & Brase, 2014) bei „CrossRef“ (Lammey, 2019). Nors „DataCite“ prasidėjo kaip pelno nesiekianti organizacija, siekianti pagerinti duomenų citavimą ir priskirti stabilius identifikatorius duomenų rinkiniams, o „CrossRef“ – kaip infrastruktūra

¹² <https://academic.oup.com/view-large/434765845>

¹³ https://oad.simmons.edu/oadwiki/Data_repositories

publikacijų metaduomenims registruoti, šiuo metu abi organizacijos registruoja DOI identifikatorius (angl. *Digital Object Identifier*) ir pateikia metaduomenų katalogus, naudodamos savo užregistruotus DOI identifikatorius. Tiek „DataCite“, tiek „CrossRef“ savo duomenis pateikia naudodamos API, o tai leidžia automatiškai parsiųsti palyginti didelius norimų duomenų kiekius mašina skaitomu pavidalu ir juos analizuoti įvairiais pjūviais.

Naudojant šiuos įrankius, iš „DataCite“ buvo parsiųsti 2024 ir 2025 m. užregistruoti DOI įrašai, priskirti atskiriems skaitmeniniams objektams (straipsniams, duomenų rinkiniams ir kt., iš viso 27 skirtingi tipai). Iš viso „DataCite“ minėtais metais buvo surasta virš 26 mln. įrašų, tarp jų per 11 duomenų rinkinių įrašų (galimybe parsiųsti pradėta naudotis 2026-02-13, baigta 2026-02-16). Daugeliui įrašų pateikti išsamūs metaduomenys, tarp kurių yra autorių afiliacijos, leidžiančios įvertinti Lietuvoje sugeneruotų duomenų rinkinių kiekį. Į šiuos duomenis nebuvo įtraukti Japonijos „National Institute for Fusion Science“, per minėtą laiką užregistruoti dar apie 20 mln. labai specifinio tipo duomenų rinkinių, kuriuose autoriai iš Lietuvos neminimi.

Naudojant „CrossRef“ API¹⁴ buvo surasta papildomai 3 mln. duomenų rinkinių. „DataCite“ ir „Crossref“ identifikatorių rinkiniai beveik nesutampa, bendri yra tik 1875 DOI identifikatoriai su prefiksu 10.4211, priklausantys Hidrologijos universitetų konsorciui (angl. *Consortium of Universities for the Advancement of Hydrologic Science, Inc.*). Taigi Lietuvoje sukurti duomenų rinkiniai turėjo būti ieškomi „Crossref“ duomenyse nepriklausomai. Deja, „Crossref“ atsakymuose nebuvo nė vienos autoriaus afiliacijos, o institucijos buvo nurodytos dažniausiai tos, kurios kaupia duomenis, o ne tos, kurios duomenis generuoja. Todėl Lietuvoje sukurtų duomenų „Crossref“ lentelėse teko ieškoti netiesioginiais būdais (žr. toliau).

Duomenų citavimo korpusas

„DataCite“, bendradarbiaudama su *Make Data Count* iniciatyva¹⁵, rengia taip pat ir duomenų citavimo lenteles, vadinamąjį duomenų citavimo korpusą (angl. *Data Citation Corpus*) (toliau – DCC) (DataCite & Make Data Count, 2025). Jose pateikiami moksliniuose straipsniuose paminėti (pacituoti) duomenų rinkiniai; šiuo metu turima v4.1 šių lentelių versija. Mūsų tikslams buvo naudinga patikrinti, kiek dubliuojasi „DataCite“ ir „Crossref“ kataloguose minimi duomenų rinkiniai su cituojamais straipsniuose. Toks sugretinimas buvo naudingas vertinant bendrą duomenų ekosistemos aprėptį, tačiau nebuvo laikomas tiesioginiu saugyklų sąrašo sudarymo įrankiu.

DCC nukeltame archyve (2025-08-15 laida) pateikiama apie 10 mln. citavimo eilučių. Tačiau iš visų stulpelyje „dataset“ paminėtų identifikatorių toli gražu ne visi rodo į duomenų rinkinius; pvz., dažniausiai pasitaikantis identifikatorius LY294002, panašu, identifikuoja dažnai eksperimentuose naudojamą cheminį junginį, bet ne duomenų rinkinį. Patikimai identifikuoti kaip duomenų rinkinių identifikatorius DCC galima tik DOI pavidalo eilutes (atitinkančios „Perl“ suderinamą reguliarią išraišką $10\.[0-9]+/[^\s]+$); tokius identifikatorius turi kiek daugiau nei 1,8 mln. citatų, kuriose yra 1,6 mln. unikalių identifikatorių (vidutinis cituojamumas – apie 1,1 karto, daugiausia cituojamas duomenų rinkinys 10.17882/42182 su pasaulinio vandenyno temperatūros ir druskingumo duomenimis pacituotas daugiau nei 6000 kartų, antras pagal cituojamumą 10.7937/K9/TCIA.2017.KLXWJJ1Q pacituotas 505 kartus).

DCC minimų cituojamų duomenų rinkinių ir „DataCite“ bei „Crossref“ kolekcijų dubliavimosi apimtis nėra didelė: sujungtų „DataCite“ ir „Crossref“ DOI sąrašų bei cituojamų DCC bendrų yra tik kiek daugiau negu 100 tūkst. identifikatorių; likusių 1,5 mln. nei „DataCite“, nei „CrossRef“ nepasitaiko.

Sąrašų sugretinimas ir išsamumo ribotumai

Saugyklų sąrašo išsamumui patikrinti buvo peržiūrėtos saugyklų aprašymų sutaptys įvairiuose kataloguose. Jei skirtingi, pageidautina nepriklausomi, katalogai minėtų vienas ir tas pačias saugyklas, galėtume teigti, kad jose yra populiariausių tarp mokslininkų saugyklų sąrašas. Jis yra beveik galutinis, palyginamas su saugyklų sutaptimis. Jei sutapčių būtų nedaug, galėtume konstatuoti, kad skirtinguose kataloguose aprašomos skirtingos saugyklų aibės dalys yra nebaigtos, neišsamios. Todėl galime tikėtis surasti dar nemažai dabartiniuose kataloguo-

¹⁴ <https://www.crossref.org/documentation/retrieve-metadata/rest-api/>

¹⁵ <https://makedatacount.org/>

se nepaminėtų duomenų bazių ar saugyklų. Norint gauti kuo išsamesnį sąrašą, reikėtų sujungti skirtingų saugyklų sąrašus.

Palyginti keliuose kataloguose pateikiamus saugyklų sąrašus nėra visiškai paprasta – įvairūs šaltiniai tą pačią saugyklą gali pavadinti kiek skirtingai, ir sunku automatiškai nustatyti, kada du aprašymai žymi tą pačią saugyklą, o kada ne. Net ir URL nėra visiškai patikimi, nes tas pats URL gali būti užrašomas skirtingais būdais (pvz., duomenų bazė gali būti pasiekama skirtingais keliais), URL gali būti pakeisti laikui bėgant, tas pats URL gali būti pateiktas kelioms duomenų bazėms. Patikimiausias būdas – sujungti skirtingus saugyklų sąrašus naudojant unikalius stabilius identifikatorius, bet saugykloms, skirtingai negu patiems duomenų rinkiniams ar straipsniams, tokių identifikatorių nėra paplitusių. Patikimiausi, ko gero, yra „FAIRSharing“ ir „Re3Data“ priskiriami identifikatoriai. Procesą palengvina tai, kad abu saugyklų katalogai registruoja vienas kito identifikatorius.

Naudojant „Re3Data“ ir „FAIRsharing“, sąrašai iš dalies dubliuojasi: aptinkami 925 „Re3Data“ identifikatoriai yra bendri abiem katalogams ir 763 – „FAIRSharing“ identifikatoriai. Procesas nėra vienareikšmiškas, nes „Re3Data“ mini 811 „FAIRSharing“ identifikatorių, bet 48 iš jų pačioje „FAIRSharing“ šiuo metu neberandami. Tėka manyti, kad tai neišvengiami duomenų tvarkymo netobulumai, ir pasitenkinti apytikrių saugyklų skaičiumi. Kadangi tiek „FAIRSharing“, tiek „Re3Data“ minima apie 4000 saugyklų, sutaptys sudaro apie 20 %. Tai nėra daug, todėl saugyklų sąrašą reikėtų pildyti; tačiau galima tikėtis, kad svarbiausios Lietuvos MSI naudojamos saugyklos šiame sąrašė bus, jei sujungsime visus turimus saugyklų sąrašus.

Remiantis šiame skyrelyje aprašyta metodika įvertintas Lietuvos mokslininkų indėlis į pasaulinę duomenų erdvę ir Lietuvoje generuojamos duomenų apimtys (rezultatai pristatomi 2.3.3 poskyryje).

2. Mokslinių tyrimų duomenų naudojimo teisinis reglamentavimas, principai, standartai, sąveikumas, valdysena ir tvarkysena

2.1. Mokslinių tyrimų duomenų naudojimo strateginiai tikslai ir teisinė bazė

2.1.1. Mokslinių tyrimų duomenų apibrėžimas

ES suformavo MTD apibrėžimą, kuris privalomai taikomas visose valstybėse narėse įgyvendinant ES teisę. Atvirųjų duomenų direktyvoje pateikiamas MTD apibrėžimas: „Mokslinių tyrimų duomenys – skaitmeniniai dokumentai, išskyrus mokslines publikacijas, kurie yra renkami ar rengiami vykdant mokslinių tyrimų veiklą ir yra naudojami kaip įrodymai mokslinių tyrimų procese arba yra mokslinės bendruomenės visuotinai pripažinti būtinai mokslinių tyrimų išvadoms ir rezultatams patvirtinti“ (Directive (EU) 2019/1024, 2 str. 9 d.). Šis apibrėžimas yra beveik pažodžiui perkeltas į Lietuvos Respublikos teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatymą (Lietuvos Respublikos Seimas, 2000) (3 str. 12 d.) nurodant, kad „mokslinių tyrimų duomenys – duomenys, išskyrus mokslines publikacijas, kurie renkami ar rengiami vykdant mokslinių tyrimų veiklą ir naudojami kaip įrodymai atliekant mokslinius tyrimus arba mokslinės bendruomenės visuotinai pripažinti būtinai mokslinių tyrimų išvadoms ir rezultatams patvirtinti“.

Atkreiptinas dėmesys, kad 2025 m. lapkričio 19 d. EK pateiktame pasiūlyme dėl Skaitmeninio bendrojo reglamento (angl. *Digital Omnibus*) MTD apibrėžimas iš esmės nesikeičia (European Commission, 2025d). Kadangi šiuo metu nėra oficialaus vertimo į lietuvių kalbą, pateikiame minėtame pasiūlyme esantį MTD apibrėžimą anglų kalba: *‘research data’ means data, other than scientific publications, which are collected or produced in the cour-*

se of scientific research activities and are used as evidence in the research process, or are commonly accepted in the research community as necessary to validate research findings and results.

Toliau gilinantis į MTD teisinį reguliavimą verta atkreipti dėmesį, kad skaitmeniniai duomenys nepatenka į nuosavybės teisės reguliavimo lauką bei dažniausiai ir į autorių teisių bei duomenų bazių teisių reguliavimo sferas. Tai lemia skaitmeninių duomenų, kaip reguliavimo objekto, išskirtinės savybės: a) duomenys gali būti naudojami daugelio vartotojų vienu metu ir vieno asmens naudojimas duomenimis nesumažina galimybių naudotis tais pačiais duomenimis kitiems; b) išskirti kitus asmenis nuo prieigos yra techniškai įmanoma, bet nėra natūraliai neišvengiama; c) duomenys, priešingai nei fiziniai išteklių, nesunaudojami juos naudojant. Dėl to duomenų teisinis reguliavimas siekia nustatyti prieigos prie duomenų ir jų naudojimo sąlygas.

2.1.2. ES mokslinių tyrimų duomenų valdymo strateginės nuostatos ir teisinis reguliavimas

Analizė atlikta remiantis ES, Lietuvos ir kitų ES šalių dokumentais, reglamentuojančiais AM klausimus: ES teisės aktai; LR teisės aktai; Lietuvos mokslo ir studijų įstatymas; EOSC dokumentacija; Mokslą finansuojančių institucijų reikalavimai; Atvirosios prieigos prie mokslinių tyrimų ir eksperimentinės plėtos rezultatų tvarkos aprašas (Lietuvos mokslo taryba, 2024a).

ES strateginės nuostatos dėl atvirojo mokslo ir mokslinių tyrimų duomenų valdymo

AM apibrėžiamas kaip mokslinių tyrimų praktika, kur žinios yra atveriamos kuo anksčiau, remiantis skaitmeninėmis ir komunikacijos technologijomis. Pagrindinis tikslas – padaryti Europos mokslą efektyvesnį, skaidresnį, patikimesnį ir labiau orientuotą į visuomenės poreikius (plačiau žr. EK AM politikos aprašą¹⁶). EK dokumentuose akcentuojamos AM praktikos, susijusios su tyrimų rezultatų ankstyva, neribota AP iš karto po jų paskelbimo saugyklose, užtikrinant tyrimų patikimumą ir atkuriamumą, atsakingą mokslinių tyrimų rezultatų (publikacijų, duomenų ir kitų rezultatų) valdymą pagal FAIR principus (daugiau apie juos 2.2.1 poskyryje) ir reikalavimą, kad MTD bei kiti rezultatai būtų randami, prieinami, sąveikūs ir pakartotinai naudojami, išlaikant patikimumą ir pakartojamumą. Taip pat skatinamas visuomenės įtraukimas į mokslinius tyrimus ir inovacijas, piliečių mokslo stiprinimas ir visuomenės pasitikėjimo mokslu didinimas.

EK išklė ambicingus tikslus įgyvendinti skaitmeninę transformaciją, kuri yra platus ir visa apimantis procesas integruojant skaitmenines technologijas į visus ekonomikos ir visuomenės aspektus. ES skaitmeninės transformacijos koncepcija ir iniciatyvos vystėsi palaipsniui, tačiau vienas esminių dokumentų, kuriame skaitmeninė transformacija buvo aiškiai ir išsamiai apibrėžta, yra Europos duomenų strategija, paskelbta 2020 m. vasario 19 d. (European Commission, 2020). Jos tikslas – pakeisti Europos veikimo būdą, pasitelkiant skaitmenines inovacijas siekti ekonomikos augimo, didinti konkurencingumą, spręsti visuomenės iššūkius ir gerinti piliečių gyvenimo kokybę.

Europos duomenų strategijoje¹⁷ akcentuojama, kad duomenys yra esminis ekonomikos augimo, konkurencingumo, inovacijų, darbo vietų kūrimo ir visuomenės pažangos augimo šaltinis, o juos naudojant sukurtos įvairios taikomojo pobūdžio programos suteiks naudos visuomenei ir verslui, padės sveikatos priežiūros sistemai, kurs saugesnes ir švaresnes transporto sistemas, pasiūlys naujus produktus ir paslaugas, mažins viešųjų paslaugų sąnaudas, didins tvarumą ir energijos vartojimo efektyvumą.

Europos duomenų strategija siekiama sukurti bendrąją duomenų rinką, kuri užtikrintų Europos konkurencingumą pasaulyje ir duomenų suverenumą. Tam tikslui kuriama Bendroji Europos duomenų erdvė¹⁸. Siekiama, kad visuomenė ir verslas galėtų pasinaudoti kuo didesniu kiekiu duomenų bei kad duomenis kuriančios įmonės, įstaigos ir fiziniai asmenys išlaikytų jų kontrolę.

Europos duomenų erdvė jungia 14 sektorių, tarp jų ir mokslinius tyrimus, eksperimentinę plėtrą, ir inovacijas (toliau – MTEPI). EOSC yra pagrindinė Bendrosios Europos duomenų erdvės sudedamoji dalis, skirta MTEPI.

¹⁶ https://research-and-innovation.ec.europa.eu/strategy/strategy-research-and-innovation/our-digital-future/open-science_en

¹⁷ <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>

¹⁸ <https://digital-strategy.ec.europa.eu/en/policies/data-spaces>

EOSC siekiama tyrėjams ir suinteresuotosioms šalims suteikti sklandžią prieigą prie kokybiškų duomenų ir skaitmeninių paslaugų, kartu remti tarpvalstybinį ir tarpdalykinį bendradarbiavimą.

Taip pat EK 2025 m. rudenį paskelbė dvi naujas strategijas: 2025 m. spalio 8 d. komunikatą „Europos dirbtinio intelekto naudojimo moksle strategija. Klojami Europos DI mokslų išteklių centro (RAISE) pamatai“ (European Commission, 2025c) ir 2025 m. lapkričio 19 d. komunikatą „Duomenų sąjungos strategija: duomenų atvėrimas dirbtiniam intelektui“ (European Commission, 2025e).

EK parengta Europos dirbtinio intelekto moksle strategija iškelia MTD į pačią Europos mokslinio konkurencingumo šerdį. Vienas svarbiausių siūlymų – sukurti Europos DI išteklių mokslui centrą (angl. *Resource for AI Science in Europe*) (toliau – RAISE), virtualiąją ES lygmens struktūrą, kuri sujungtų talentus, skaičiavimo išteklius, finansavimą ir, svarbiausia, MTD. RAISE užduotis – nustatyti trūkstamus duomenų segmentus, užtikrinti prieigą prie DI parengtų duomenų rinkinių, koordinuoti jų prieinamumą tarp valstybių bei mokslų sričių. Strategijoje pabrėžiama, kad DI gali veiksmingai veikti tik su kokybiškais, suderinamais ir plačiai prieinamais duomenimis. Norėdama tai spręsti, EK tyrimų duomenis įtraukia į platesnę skaitmeninę ekosistemą – ypač į EOSC platformą ir kitas Europos duomenų erdves, kurios turėtų užtikrinti sklandžią prieigą prie bendrų MTD ir įrankių. Vienas iš inovatyviausių pasiūlymų – kurti duomenų laboratorijas DI gamyklose (angl. *AI factories*). Iš esmės tai sukuria aiškią grandinę: nuo pirminių MTD iki aukštos kokybės, DI parengtų duomenų rinkinių.

EK parengtoje Duomenų sąjungos strategijoje (European Commission, 2025e) pabrėžiama, kad Europa vis labiau susiduria su ryškiu duomenų trūkumu, kuris pradeda riboti ne tik DI plėtrą, bet ir platesnę mokslinių tyrimų ekosistemą. Strategijoje pabrėžiama, kad, nesustiprinus didelio masto, tarpusavyje suderinamų ir patikimai valdomų MTD sistemų, Europa rizikuoja atsilikti nuo pasaulio lyderių, kurių MTI jau dabar veikia daug didesniu mastu. Atsakydama į šią problemą, EK numato kryptingai plėsti tyrimams svarbius duomenų išteklius – nuo anonimizuotų klinikinių vaizdų ar genominių rinkinių iki aplinkosauginių stebėjimų, kultūros paveldo archyvų ir daugiakalbių tekstynų. Specializuotų duomenų laboratorijų kūrimas žymi svarbų poslinkį link praktinių, tyrėjams orientuotų duomenų paslaugų: šios laboratorijos kaupys, tvarkys, pseudonimizuos ir sintetins duomenis, padėdamos tyrėjams įveikti kliūtis, kurias sukelia skirtingos taisyklės, institucijų duomenų izoliacija ir nevienodi techniniai pajėgumai. Suteikdamos saugias duomenų apdorojimo aplinkas ir pažangias privatumo užtikrinimo technologijas, duomenų laboratorijos turėtų sumažinti dalijimosi jautriais ar komerciškai saugomais duomenimis barjerus.

MTD aptariami ir platesniame AM kontekste. Naujausia ES MTD politika ir tikslai yra nustatyti Europos Tarybos 2025 kovo 3 d. rekomendacijoje Nr. 6693/25 „Dėl 2025–2027 m. Europos mokslinių tyrimų erdvės politikos darbotvarkės“ (European Commission, 2025f). Joje pažymėta, kad pirmoji Europos mokslinių tyrimų erdvės (toliau – EMTE) struktūrinės politikos kryptis – sudaryti sąlygas AM dalintis MTD ir juos pakartotinai naudoti, be kita ko, pasitelkiant EOSC. Šia politikos kryptimi siekiama esminių pokyčių Europos mokslinių tyrimų bendruomenėje ir MTI sukuriant geresnę ES teisinę sistemą, skirtą atviram dalijimuisi MTD ir kitais skaitmeniniais mokslinių tyrimų objektais, naudojamais ir sukurtais per visą mokslinių tyrimų gyvavimo ciklą bei užtikrinti sklandžią prieigą prie jų ir patikimą pakartotinį naudojimą. Numatyti šie tikslai:

- **Mokymai ir atsilyginimas už atvirojo mokslo praktikas.** Skatinama, kad už AM praktiką ir įgūdžius būtų atlyginama.
- **Standartų, priemonių ir paslaugų kūrimas.** Siekiama sudaryti sąlygas tyrėjams rasti, gauti, pakartotinai naudoti ir derinti mokslo rezultatus pasitelkiant standartus, priemones ir paslaugas.
- **Susietos bendruomeninės infrastruktūros diegimas.** Skatinama diegti ir prižiūrėti susietą bendruomeninę infrastruktūrą, sudarančią sąlygas atvirai dalytis mokslo rezultatais.
- **FAIR MTD tinklo kūrimas.** Siekiama, kad Europa pirmautų kuriant FAIR MTD tinklą, sąveikas su kitomis duomenų erdvėmis.
- **Geresnių teisinių sąlygų sudarymas.** Siekiama sudaryti geresnes teises sąlygas ir suteikti išteklių tyrėjams, kad jie galėtų susipažinti su viešosiomis lėšomis finansuojamų mokslinių tyrimų rezultatais ir juos pakartotinai naudoti, taip pat naudoti leidinius ir duomenis mokslo tikslais.

2025–2027 m. EMTE politikos darbotvarkėje nustatytas veiksmas, skirtas lygybei AM srityje užtikrinti. Pažymima, kad, įgyvendinant AM principus, reikia tvirtos politinės paramos ir didelių investicijų į išteklius, koordinavimą, technologinę parengtį, atvirąją prieigą ir skaitmeninius įgūdžius, taip pat reikia veiksmingų skatinimo, pripažinimo ir atlygio sistemų tyrėjams kūrimo. Visa tai labai skiriasi ne tik Europos šalyse ir organizacijose, bet ir visame pasaulyje, todėl EMTE reikia lygybę užtikrinančių požiūrių. Šio EMTE veiksmo tikslai:

- **Kliūčių ir iššūkių nustatymas.** Siekiama nustatyti kliūtis ir iššūkius, susijusius su lygybe AM srityje, ir pasiūlyti tolesnius su politiniu atsaku susijusius veiksmus.
- **Nepelno AP akademinė leidybos modelių plėtra.** Siekiama plėtoti ne pelno AP akademinis leidybos modelius, plačiai taikomus ir pripažįstamus įvairiose disciplinose, įskaitant mokslinių tyrimų vertinimą.
- **Gebėjimų stiprinimas.** Skatinama stiprinti gebėjimus, susijusius su AM įgūdžiais ir lygiomis galimybėmis EMTE mokslinių tyrimų bendruomenei naudotis AM infrastruktūra.
- **ES politikos metodų plėtotė.** Siekiama plėtoti ES politikos metodus, siekiant nustatyti būdus, kaip užtikrinti didesnę lygybę, kiek tai susiję su AM praktika ir infrastruktūra.

MTD svarba pabrėžiama DI vystymosi kontekste „Išvadose dėl ES DI moksle strategijos“ (2025-05-23 9292/25) (Council of the EU, 2025). DI yra universalus duomenų analizės ir modeliavimo, vedančio į naujus atradimus, įrankis. Valstybės narės kviečiamos aktyviai prisidėti prie esamų bendrųjų Europos duomenų erdvių, tokių kaip EOSC, ir naujų duomenų erdvių įgyvendinimo, kai tai būtina, siekiant palaikyti DI pagrindu vykdomus mokslinius tyrimus ir DI modelių mokymą. MTD svarba akcentuojama ir Europos inovacijų darbotvarkėje¹⁹, kur MTD yra pagrindinė naujų inovacijų ir technologijų, įskaitant DI, kūrimo varomoji jėga.

Analizuojant EK AM ir EOSC AM reglamentuojančius dokumentus, galima išskirti pagrindinius strateginius tikslus:

- Sukurti susietą, daugiadalykę aplinką, kurioje tyrėjai ir mašinos gali rasti, pasiekti ir pakartotinai naudoti MTD bei paslaugas (EOSC).
- Paversti AM numatytąją mokslo veiklos formą (*modus operandi*) – atviri publikacijų, duomenų, programinės įrangos, protokolų šrautai, paremti FAIR principais.
- Sustiprinti EK duomenų ekonomiką ir inovacijas – per atvirų duomenų direktyvą ir duomenų erdves (Open Data Directive (Directive (EU) 2019/1024); Data Governance Act (Regulation (EU) 2022/868); Data Act (Regulation (EU) 2023/2854)).
- Užtikrinti pagrindinių teisių apsaugą, ypač asmens duomenų apsaugą (BDAR) ir intelektualinę nuosavybę, naudojant anonimizavimą, ribotą prieigą ir pan.
- Keisti tyrimų vertinimą – kad būtų vertinami ne tik straipsniai, bet ir duomenų rinkiniai, programinė įranga ir kiti atviri rezultatai (EK rekomendacija 2018/790 (European Commission, 2018), CoARA²⁰ ir pan.).

Per pastarąjį dešimtmetį ES institucijos priėmė ne vieną teisės aktą, skirtą ES MTD strategijai įgyvendinti. Plačiau aptarsime:

- Direktyvą (ES) 2019/1024 (Directive (EU) 2019/1024) dėl atvirų duomenų ir viešojo sektoriaus informacijos pakartotinio naudojimo (Atvirųjų duomenų direktyva, toliau – ADD);
- Reglamentą (ES) 2022/868 (Regulation (EU) 2022/868) dėl duomenų valdymo (Duomenų valdymo aktas, toliau – DVA);
- Reglamentą (ES) 2023/2854 (Regulation (EU) 2023/2854) dėl sąžiningų duomenų prieigos ir naudojimo taisyklių (Duomenų aktas, toliau – DA).

¹⁹ <https://lithuania.representation.ec.europa.eu/news/ek-pristate-nauja-europos-inovaciju-darbotvarke-kuria-siekiama-europos-isitvirtinimo-naujos-2022-07-05.lt>

²⁰ <https://www.coara.org/>

Pirmasis ir svarbiausias teisės aktas MTD valdymo srityje yra ADD. Pagrindinis šios direktyvos tikslas – nustatyti gaires dėl viešojo sektoriaus informacijos pakartotinio naudojimo. Tai apima geografinius, nekilnojamojo turto registro, statistinius ir teisinius duomenis, kuriuos turi viešojo sektoriaus institucijos ar viešojo sektoriaus įmonės, taip pat iš viešųjų lėšų finansuojamų MTD. Direktyvoje pateikiama tiksli MTD apibrėžtis, kuri yra itin svarbi, nes užtikrina vienodą šios sąvokos supratimą visose ES valstybėse narėse.

Remiantis direktyva, mokslinių tyrimų duomenys – tai skaitmeninės formos dokumentai, išskyrus mokslines publikacijas, kurie yra surinkti ar sukurti vykdant mokslinių tyrimų veiklą ir naudojami kaip įrodymai tyrimų procese arba mokslo bendruomenėje paprastai pripažįstami kaip būtini mokslinių tyrimų išvadoms ir rezultatams patvirtinti (ADD 2 str. 9 d.). „Mokslinių tyrimų duomenys apima statistiką, eksperimentų rezultatus, matavimus, lauko tyrimų metu gautus stebėjimus, apklausų rezultatus, interviu įrašus ir vaizdus. Jie taip pat apima metaduomenis, specifikacijas ir kitus skaitmeninius objektus“ (ADD 27 konstatuojamoji dalis). Tačiau administraciniai duomenys ir mokomoji medžiaga į ADD taikymo sritį nepatenka.

Kitas svarbus ADD aspektas – nustatyti, kurioms mokslinių tyrimų institucijoms ji taikoma. ADD nėra specialiai pritaikyta universitetams ar mokslinius tyrimus vykdančioms organizacijoms, todėl kartais gali kilti sunkumų nustatant, kokios normos joms taikytinos. Šis sudėtingumas atsiranda todėl, kad ADD nustato skirtingus reguliavimo režimus viešojo sektoriaus institucijoms, viešojo sektoriaus įmonėms ir MTD, todėl ne visada lengva nustatyti, kurios normos taikomos universitetams ir mokslinių tyrimų institucijoms. Universitetas apibrėžiamas kaip „viešojo sektoriaus institucija, organizuojanti pavidurines aukštojo mokslo studijas, kurias baigus suteikiamas akademinis laipsnis“ (ADD 2 str. 4 d.). Svarbu pažymėti, kad ADD universitetams taikoma tik MTD atžvilgiu, o švietimo įstaigų ir administracinių padalinių turimi dokumentai neįtraukiami (ADD 1 str. 2 d. l p.). Be to, ADD nustato atskirus reguliavimo režimus bibliotekoms, įskaitant universitetų bibliotekas (ADD 1 str. 2 d. j p.).

Direktyvoje taip pat vartojamos sąvokos „mokslinius tyrimus vykdančios organizacijos“ ir „mokslinius tyrimus finansuojančios organizacijos“, tačiau jos aiškiai neapibrėžiamos. Šios organizacijos patenka į direktyvos taikymo sritį tik tuo atveju, jei atitinka „viešojo sektoriaus institucijos“ apibrėžimą.

ADD 10 straipsnio 1 dalis įpareigoja valstybes nares nustatyti AP prie MTD politiką. ADD daugiausia remiasi nacionalinėmis AP politikomis ir suteikia valstybėms narėms plačią diskreciją nustatyti sąlygas, kada MTD turi būti atverti. Tai grindžiama principu „kiek įmanoma atvira, kiek būtina uždara“.

Be to, ADD aiškiai nurodo, kad ji neįpareigoja atverti duomenų, jei yra pagrįstų priežasčių juos laikyti uždaru. AP pagal „kaip praktika, pagal kurią internetinė prieiga prie mokslinių tyrimų rezultatų galutiniam naudotojui suteikiama nemokamai ir netaikant apribojimų naudojimui bei pakartotiniam naudojimui, išskyrus galimybę reikalauti nurodyti autorystę“ (ADD, konstatuojamoji dalis 27p.). Ji gali būti užtikrinama skelbiant pagal *Creative Commons Public Domain Dedication (CC0)*, *Creative Commons Attribution (CC-BY)* ar kitas atviro turinio licencijas.

ADD nepaneigia MTD apsaugos, kurią suteikia intelektinės nuosavybės teisės, komercinės paslaptys ar BDAR. Direktyva reikalauja atverti MTD tik tuomet, kai tai neprieštarauja kitiems ES ar nacionaliniams teisės aktams, laikantis principo „kiek įmanoma atvira, kiek būtina uždara.“ Pagrindinis MTD politikos reikalavimas – kad duomenys, kaip numatyta, būtų atviri ir atitiktų FAIR duomenų principus.

ADD 10 straipsnio 2 dalis įpareigoja valstybes nares sudaryti sąlygas pakartotinai naudoti MTD komerciniais ir nekomerciniais tikslais, jei jie yra „viešai prieinami per institucinę ar teminę saugyklą“. Čia matyti valstybių narių kompetencijos nustatyti AP politiką ir ADD sąveiką. Nors ADD įpareigoja užtikrinti pakartotinį naudojimą tik viešai prieinamiems duomenims, viešo prieinamumo sąlygas nustato pačios valstybės narės per AP politiką.

ADD nustato kelis esminius pakartotinio MTD naudojimo reikalavimus: skaidrumą (ADD 7 str.), nediskriminavimą (ADD 11 str.) ir principą, kad pakartotinis naudojimas turi būti nemokamas (ADD 6 str.). Šios nuostatos užtikrina, kad MTD būtų prieinami ir naudojami, taip skatinant skaidrumą ir inovacijas visoje ES.

Galiausiai pabrėžtina, kad ADD orientuota į viešųjų duomenų atvėrimą, kad jie galėtų būti laisvai naudojami mokslinių tyrimų tikslais. Todėl galima teigti, kad ADD ne tik nustato pareigas tyrėjams atverti MTD, bet ir suteikia prieigą prie didžiulio kiekio viešųjų duomenų, kurie gali būti naudojami moksliniams tyrimams.

Taigi, ADD yra reikšmingas reguliavimas, susijęs su MTD valdymu ir pakartotiniu naudojimu visoje ES. Ji užtikrina nuoseklumą ir skatina bendradarbiavimu grindžiamą mokslinių tyrimų aplinką tarp valstybių narių,

pateikdama aiškią ir bendrą MTD apibrėžtį. Direktyvos lankstumas leidžia kiekvienai valstybei narei nustatyti savo AP politiką, laikantis principo „kiek įmanoma atvira, kiek būtina uždara.“ Toks požiūris gerbia intelektinės nuosavybės, komercinių paslapčių ir privatumo apsaugą, kartu skatindamas skaidrumą ir inovacijas. Be to, įtvirtindama pareigą užtikrinti iš viešųjų lėšų finansuotų MTD pakartotinį naudojimą ir nustatydamas pagrindinius reikalavimus – skaidrumą, nediskriminavimą ir nemokamą pakartotinį naudojimą – ADD reikšmingai padidina MTD prieinamumą ir naudojamumą. Taigi ADD ne tik nustato pareigas tyrėjams atverti MTD, bet ir sudaro sąlygas prieigai prie plačios viešųjų duomenų bazės, taip remdama mokslinius tyrimus visoje ES.

Nors ADD nustato viešojo sektoriaus informacijos pakartotinio naudojimo sistemą, joje pripažįstama, kad tam tikrų rūšių duomenys negali būti atverti dėl įvairių teisinių ir etinių apribojimų. Siekiant papildyti ADD ir spręsti šiuos ribojimus, buvo priimtas DVA, kuris taikomas nuo 2023 m. rugsėjo 24 d. DVA tikslas – reglamentuoti tam tikrų viešojo sektoriaus duomenų, įskaitant MTD, kurių negalima paskelbti kaip atvirų duomenų, pakartotinį naudojimą. Toks dvejetainis požiūris leidžia panaudoti plačią duomenų įvairovę visuomenės labui, kartu laikantis būtinų ribojimų.

DVA 3 straipsnyje išskiriamos keturios konkrečios duomenų kategorijos, patenkančios į jo taikymo sritį:

1. Komercinis konfidencialumas, įskaitant verslo, profesines ir įmonių paslaptis.
2. Statistinis konfidencialumas.
3. Trečiųjų šalių intelektinės nuosavybės teisių apsauga.
4. Asmens duomenų apsauga.

DVA sudaro sąlygas pakartotinai naudoti šias saugomų duomenų kategorijas taip, kad būtų derinamas atvirumas ir konfidencialumas. Ši sistema būtina pasitikėjimui užtikrinti ir garantuoti, kad saugomi duomenys būtų pakartotinai naudojami tik laikantis griežtų sąlygų. Viešosios mokslinius tyrimus vykdančios organizacijos taip pat patenka į DVA reguliavimo sritį, išskyrus universitetų bibliotekas. Tai svarbu todėl, kad viešosios mokslinius tyrimus vykdančios organizacijos dažnai tvarko duomenis, priklausančius DVA numatytoms saugomoms kategorijoms. DVA 4 straipsnyje aiškiai draudžiami išimtiniai susitarimai dėl duomenų, kuriems taikomas šis aktas, taip užkertant kelią monopolijoms ir užtikrinant platesnę prieigą prie vertingų duomenų.

Be šių priemonių, DVA taip pat reglamentuoja duomenų altruizmo sąvoką. 2 straipsnio 16 dalyje nustatyta, kad duomenų altruizmas – tai savanoriškas dalijimasis duomenimis, grindžiamas duomenų subjektų sutikimu tvarkyti su jais susijusius asmens duomenis arba duomenų turėtojų leidimu naudoti jų ne asmens duomenis, nesiekiant ar negaunant atlygio, viršijančio su duomenų pateikimu susijusių sąnaudų kompensavimą, kai duomenys teikiami bendrojo intereso tikslams, numatytiems nacionalinėje teisėje (kai taikoma), pavyzdžiui, sveikatos priežiūrai, kovai su klimato kaita, judumui gerinti, oficialiajai statistikai rengti, gamybai ir sklaidos palengvinti, viešųjų paslaugų teikimui gerinti, viešajai politikai formuoti ar moksliniams tyrimams bendrojo intereso labui. Tai skatina asmenis ir organizacijas dalytis duomenimis bendrojo intereso tikslais, tokiais kaip sveikatos apsauga, kova su klimato kaita, judumo gerinimas, oficialiosios statistikos rengimas, viešųjų paslaugų gerinimas, viešosios politikos formavimas ar moksliniai tyrimai. Reglamentuodamas duomenų altruizmą, DVA sukuria sistemą, leidžiančią panaudoti duomenų turėtojų geranoriškumą visuomenės labui ir skatina dalijimosi bei bendradarbiavimo kultūrą.

Taigi DVA papildo ADD, sukurdamas saugomų duomenų pakartotinio naudojimo reguliavimo sistemą. Remiantis DVA, mokslinius tyrimus vykdančios organizacijos atlieka dvejetainį – jos yra ir duomenų gavėjos, ir duomenimis besidalijančios institucijos. Toks dvejetainis vaidmuo leidžia joms tiek gauti naudos iš platesnės duomenų ekosistemos, tiek prie jos prisidėti, skatinant inovacijas ir visuomenės gerovę atsakingai valdant saugomus duomenis ir pakartotinai naudojant.

Duomenų aktas (DA) nuo ADD ir DVA skiriasi keliais esminiais aspektais, ypač tuo, kokioms institucijoms jis taikomas ir kokio pobūdžio duomenis apima. Nors ADD ir DVA daugiausia orientuoti į viešojo sektoriaus institucijas ir jų turimus duomenis, DA daugiausia dėmesio skiria privačių įmonių turimiems duomenims. Svarbu pažymėti, kad DA nepanaikina jokių nacionalinių ar ES teisės aktų, ribojančių dalijimąsi duomenimis, pavyzdžiui, susijusių su intelektinės nuosavybės teisėmis ar asmens duomenų apsauga.

DA nustato taisykles kelioms pagrindinėms sritims, kaip numatyta DA 1 straipsnyje:

- a) produktų duomenims ir susijusių paslaugų duomenims pateikti susietojo produkto ar susijusios paslaugos naudotojui;
- b) duomenims pateikti iš duomenų turėtojų duomenų gavėjams;
- c) duomenims pateikti iš duomenų turėtojų viešojo sektoriaus institucijoms, kai šių duomenų išimtinai reikia konkrečiai viešojo intereso užduočiai atlikti;
- d) perėjimui tarp duomenų tvarkymo paslaugų palengvinti;
- e) apsaugos priemonėms nuo neteisėtos trečiųjų šalių priegios prie ne asmens duomenų įdiegti;
- f) sąveikumo standartams, skirtiems duomenų prieigai, perduoti, naudoti, kurti.

DA II skyrius reglamentuoja modelių „verslas verslui“ (angl. *business to business*, B2B) ir „verslas vartotojui“ (angl. *business to consumer*, B2C) dalijimąsi duomenimis, ypač daiktų interneto (angl. *Internet of Things*) rinkoje. Šis reguliavimas taikomas visiems neapdorotiems ir iš anksto apdorotiems duomenims, generuojamiems naudojant susietuosius produktus ar susijusias paslaugas, kurie yra lengvai prieinami duomenų turėtojui. Tai apima asmens duomenis, ne asmens duomenis ir metaduomenis. Pažymėtina, kad tokiais sąlygomis *sui generis* duomenų bazių teisė netaikoma.

Mokslinius tyrimus vykdančios organizacijos gali produktyviai pasinaudoti šiomis nuostatomis. Kaip naudotojos, mokslinius tyrimus vykdančios organizacijos gali gauti duomenis iš duomenų turėtojų DA nustatytomis sąlygomis ir naudoti juos tyrimų tikslais arba dalytis su trečiosiomis šalimis. Be to, mokslinius tyrimus vykdančios organizacijos gali veikti kaip trečioji šalis, kai naudotojai paprašo duomenų turėtojų perduoti duomenis joms. Vis dėlto svarbu pažymėti, kad duomenų turėtojai neprivalo dalytis lengvai prieinamais duomenimis su trečiosiomis šalimis naujų, dar rinkoje nepasirodžiusių susietųjų produktų, medžiagų ar procesų bandymams, nebent tai numatyta sutartyje.

Mokslinius tyrimus vykdančios organizacijos taip pat gali veikti kaip duomenų turėtojos, ypač kai jos pačios gauna ar generuoja produktų ar susijusių paslaugų duomenis, teikdamos susijusią paslaugą. Jei mokslinius tyrimus vykdančios organizacijos laikomos duomenų turėtojos, jos privalo laikytis visų DA nustatytų pareigų, užtikrindama skaidrumą ir sąžiningumą dalijantis duomenimis.

DA III skyrius reglamentuoja situacijas, kai verslas (duomenų turėtojas) pagal ES ar nacionalinę teisę privalo pateikti duomenis kitam verslui (duomenų gavėjui). Šis skyrius taikomas visiems įmonių turimiems duomenims, neapsiribojant II skyriuje reglamentuojamais duomenimis. Juo užtikrinama, kad įmonės negalėtų nepagrįstai sukaupti duomenų, kai egzistuoja teisinė pareiga jais dalytis, taip skatinant sąžiningą ir konkurencingą rinką.

Mokslinius tyrimus vykdančios organizacijos, kaip potencialios duomenų gavėjos, gali prašyti duomenų turėtojų joms pateikti duomenis. Tačiau, kaip minėta, duomenų turėtojai neprivalo dalytis lengvai prieinamais duomenimis naujų, dar rinkoje nepasirodžiusių produktų ar procesų bandymams, nebent tai leidžia sutartis. Tai leidžia verslui apsaugoti inovacinius procesus, kartu laikantis teisinių dalijimosi duomenimis reikalavimų.

DA leidžia duomenų turėtojams imti mokesť už dalijimąsi duomenimis. DA 9 straipsnio 4 dalis riboja kainą, taikomą ne pelno mokslinius tyrimus vykdančioms organizacijoms, nustatant, kad ji gali padengti tik duomenų pateikimo sąnaudas, pavyzdžiui, formatavimo, elektroninio perdavimo ir saugojimo išlaidas. Taip užtikrinama, kad dalijimasis duomenimis būtų sąžiningas ir nesukeltų neproporcingos finansinės naštos ne pelno tyrimų organizacijoms.

DA IV skyrius užtikrina sąžiningas sutartines sąlygas duomenims gauti, ypač kai viena įmonė turi stipresnę derybinę poziciją ir nustato nediskutuotinas priegios prie duomenų sąlygas. Vis dėlto neaišku, ar mokslinius tyrimus vykdančios organizacijos visais atvejais bus laikomos silpnesniąja šalimi. Bet kurios nesąžiningos sutartinės sąlygos, vienašališkai primestos silpnesnei šaliai, neturėtų būti privalomos. Tai leidžia mokslinius tyrimus vykdančioms organizacijoms derėtis dėl sąžiningų sąlygų ir gauti reikalingus duomenis nepatiriant išnaudojimo.

Be to, IV skyrius leidžia viešojo sektoriaus institucijoms tam tikromis sąlygomis gauti duomenis, paprastai

laikomus privačių subjektų, kai egzistuoja išimtinis poreikis. Ši nuostata taikoma visiems duomenims, tačiau prieiga prie asmens duomenų yra griežčiau ribojama. DA taip pat suteikia teisę viešosioms institucijoms dalytis gautais duomenimis su mokslinius tyrimus vykdančiomis organizacijomis. DA 21 straipsnio 1 dalies a punktas leidžia viešojo sektoriaus institucijoms teikti duomenis asmenims ar organizacijoms mokslinių tyrimų ar analitikos tikslais, suderinamais su pradiniu duomenų prašymo tikslu. Mokslinius tyrimus vykdančios organizacijos turi laikytis tam tikrų reikalavimų, įskaitant pareigą saugoti duomenis ne ilgiau kaip šešis mėnesius. Tačiau gali būti taikomos išimties, pavyzdžiui, kai archyvavimas būtinas pagal ES ar nacionalinę teisę dėl viešosios prieigos ir skaidrumo pareigų (DA 19 str. 1 d. c p.).

DA VIII skyrius skirtas duomenų sąveikumui ir nustato esminius reikalavimus, kad skirtingų šaltinių duomenys galėtų būti naudojami bendrose Europos duomenų erdvėse ir tarp jų. Šie reikalavimai ypač svarbūs tyrėjams ir organizacijoms, veikiančioms EOSC. Sąveikumo standartai yra būtini, kad duomenys iš įvairių šaltinių galėtų būti integruojami ir veiksmingai naudojami. Tai ypač svarbu tyrimų organizacijoms, kurios remiasi įvairiais duomenų rinkiniais. Laikydamosi šių standartų, mokslinius tyrimus vykdančios organizacijos gali stiprinti savo tyrimų pajėgumus ir prisidėti prie labiau susietos bei bendradarbiavimo grindžiamos mokslinių tyrimų aplinkos kūrimo.

Taigi DA nustato išsamią duomenų prieigos ir dalijimosi sistemą, kuri daro didelį poveikį mokslinius tyrimus vykdančioms organizacijoms. Sutelkdamas dėmesį į privačių įmonių turimus duomenis ir skatindamas sąžiningą bei skaidrą dalijimąsi duomenimis, DA papildo ADD ir DVA. Jis užtikrina duomenų prieinamumą, kartu saugodamas jautrią informaciją ir gerbdamas galiojančias teisinės apsaugos priemones. DA akcentuojamas sąveikumas ir sąžiningos sutartinės sąlygos dar labiau stiprina inovatyvių tyrimų ir bendradarbiavimo potencialą visoje ES.

2.1.3. LR teisės aktai ir Lietuvos mokslą finansuojančių institucijų reikalavimai

Lietuvoje MTD dalijimąsi reglamentuoja keli teisės aktai. Lietuvos Respublikos teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatymas (Lietuvos Respublikos Seimas, 2000) yra pagrindinis teisės aktas, perkeltantis ADD nuostatas į nacionalinę teisę. Jame perkeltas direktyvoje pateiktas MTD apibrėžimas (3 straipsnio 12 dalis). Šio įstatymo 10 straipsnio 8 dalyje nustatyta: „Valstybės lėšomis finansuojamų mokslinių tyrimų duomenys ir didelės vertės duomenų rinkiniai pakartotinai naudoti teikiami neatlygintinai.“ Šio įstatymo 2 straipsnio 2 dalies 8 punktas numato, kad įstatymas netaikomas duomenims, kuriuos valdo valstybinės aukštosios mokyklos ir valstybiniai mokslinių tyrimų institutai, išskyrus MTD, jeigu šie tyrimai finansuojami valstybės lėšomis, o tyrėjai, mokslinius tyrimus atliekančios organizacijos arba mokslinius tyrimus finansuojančios organizacijos jau yra juos viešai paskelbę. Ši nuostata taip pat detalizuojama Lietuvos mokslo ir studijų įstatyme (Lietuvos Respublikos Seimas, 2009), kurio 51 straipsnyje nustatyta: „Siekiant užtikrinti valstybės biudžeto lėšomis atliekamų mokslinių tyrimų kokybę, valstybės biudžeto lėšų panaudojimo skaidrumą, padidinti mokslinių tyrimų rezultatų panaudojimo galimybes, paskatinti mokslo pažangą, visi mokslo ir studijų institucijose valstybės biudžeto lėšomis atliekamų mokslinių tyrimų ir eksperimentinės plėtros rezultatai turi būti skelbiami viešai atvirosios prieigos būdais ir priemonėmis, kiek tai neprieštarauja asmens duomenų, intelektinės nuosavybės, profesinių, komercinių ar valstybės ir tarnybos paslapčių apsaugai, nacionalinio saugumo ir gynybos, taip pat teisėsaugos ir viešojo saugumo veiklą reglamentuojantiems teisės aktams <...>.“

Ši nuostata atsirado anksčiau nei 2012 m. ir 2018 m. EK rekomendacijos dėl AP, tad galima teigti, kad Lietuva anksti įteisino atvirumo principą, kas formaliai suteikia privalomą pagrindą atverti tiek publikacijas, tiek MTD, jei jie sukurti finansuojant mokslinius tyrimus valstybės lėšomis. Strategiškai tai reiškia, kad nacionalinis tikslas sutampa su EK reikalavimais užtikrinti, kad viešaisiais pinigais finansuotų tyrimų rezultatai būtų prieinami visuomenei, didintų mokslo kokybę, skaidrumą ir efektyvumą. LR mokslo ir studijų įstatymas įpareigoja LMT patvirtinti AP prie mokslinių tyrimų ir eksperimentinės plėtros rezultatų įgyvendinimo tvarką.

Lietuvos mokslą finansuojančių institucijų reikalavimai

Lietuvos MTD politikos raida pastaraisiais metais rodo aiškų perėjimą nuo rekomendacinio prie normatyvinio reguliavimo, tačiau kartu atskleidžia struktūrinę įtampą tarp teisinių nuostatų, administracinių reikalavimų ir realių įgyvendinimo pajėgumų. Ši įtampa tampa ypač reikšminga vertinant Lietuvos pasirengimą integruotis į EOSC ne formaliai, o sistemiškai.

Teisinį pagrindą sudarantis jau minėtas Mokslo ir studijų įstatymo 51 straipsnis (Lietuvos Respublikos Seimas, 2009) ir jame įvardintos nuostatos yra būtinos, tačiau savaime nepakankamos MTD valdymo sistemai sukurti, nes apima rezultato (viešumo) siekį, bet neapibrėžia nei MTD gyvavimo ciklo, nei infrastruktūros ar kompetencijų mechanizmų, reikalingų tam siekiui pasiekti. Kitaip tariant, įstatymas sukuria teisėtą pareigą, bet ne institucinę architektūrą.

Šią spragą ilgą laiką bandė užpildyti LMT atvirosios prieigos gairės (Lietuvos mokslo taryba, 2016), kurios atliko svarbų vaidmenį formuojant bendrą supratimą apie AM, DVP, FAIR principus ir institucijų atsakomybes. Tačiau pats gairių pobūdis lėmė, kad jų įgyvendinimas buvo nevienodas ir fragmentuotas: dalis institucijų jas traktavo kaip strateginę kryptį, dalis – kaip formalią prievolę, o realus poveikis duomenų valdymo praktikoms stipriai priklausė nuo pavienių ekspertų, projektų ar bibliotekų iniciatyvų. Tai reiškia, kad iki 2024 m. Lietuvoje dominavo vadinamosios minkštosios politikos (angl. *soft policy*) modelis, kuris skatino, bet neįpareigojo.

Įgyvendindama Mokslo ir studijų įstatymo 51 str. nuostatas LMT 2024 m. rugpjūčio 30 d. priėmė įsakymą V-385, kuriuo patvirtino naująjį Atvirosios prieigos prie mokslinių tyrimų ir eksperimentinės plėtos rezultatų tvarkos aprašą (Lietuvos mokslo taryba, 2024a). Šis teisės aktas žymi kokybinį pokytį valstybės mokslo politikoje: nuo 2028 m. gruodžio 31 d. neteks galios 2016 m. patvirtintos gairės (VIII-2), o jų vietą užims privalomojo pobūdžio nuostatos, pradedamos taikyti naujiems kvietimams jau nuo 2025 m. sausio 1 d. Šis dokumentas žymi perėjimą nuo gairių prie privalomo reguliavimo, jame pirmą kartą sistemingai sujungiamos publikacijos ir MTD į vieną AP objektą, nustatomos aiškesnės pareigos institucijoms ir projektams, numatomi pereinamieji laikotarpiai ir būsimos stebėsenos mechanizmų kūrimas.

AP tvarkos apraše suprantama kaip nemokama ir nevaržoma galimybė skaityti, kopijuoti, platinti bei atlikti automatinę turinio analizę nepažeidžiant autorių teisių. Pagrindinis šio proceso siekis – užtikrinti valstybės lėšomis finansuotų tyrimų rezultatų pakartotinį naudojimą, didinti mokslo patikimumą ir puoselėti akademinę etiką. Naujasis reglamentavimas remiasi tarptautiniais standartais, įskaitant UNESCO ir EK rekomendacijas bei „Europos horizonto“ programos principus.

MTD apibrėžiami kaip tyrimo metu gauta medžiaga, kuri turi turėti nuolatinius identifikatorius ir būti aprašyta metaduomenimis. Centrinė reglamentavimo ašis – FAIR principų įgyvendinimas.

Kiekvienas projektas privalo turėti DVP, kuriame dokumentuojama, kaip MTD bus kaupiami, saugomi ir atveriami. DVP turėtų apimti duomenų tipų ir formatų aprašymus, duomenų saugumo priemones, ypač dirbant su neskelbtina informacija, teisinių reikalavimų (intelektinės nuosavybės, asmens duomenų) vykdymą, ilgalaikio saugojimo vietą ir sąlygas.

Nors laikomasi principo „kiek įmanoma, atviri, kiek būtina, uždari“, numatytos teisėtos išimtys, kai duomenys gali būti neviešinami dėl nacionalinio saugumo, valstybės paslapčių ar trečiųjų šalių teisių. Svarbu tai, kad net jei patys duomenys yra uždari ar sunaikinti, jų metaduomenys privalo išlikti visiškai atviri ir pasiekiami paieškai.

Tyrėjai skatinami rinktis patikimas saugyklas, kurios yra sertifikuotos ir integruotos į EOSC. Taip pat pabrėžiama piliečių mokslo svarba, įtraukiant visuomenę į duomenų rinkimo ar vertinimo procesus.

Svarbu pažymėti, kad Atvirosios prieigos prie mokslinių tyrimų ir eksperimentinės plėtos rezultatų tvarkos aprašas įsigalioja ne iš karto visa apimtimi, o palaipsniui. Dokumente numatytas etapinis įgyvendinimo modelis, pagal kurį dalis nuostatų pradedama taikyti nuo 2025 m., kitos – nuo 2026 m., o tam tikri sudėtingesni ar sistemiskai reikšmingi reikalavimai atidedami net iki 2030 m. Toks sprendimas rodo, kad teisės akto rengėjai pripažįsta praktinį poreikį suteikti mokslo ir studijų institucijoms pereinamąjį laikotarpį, per kurį būtų galima adaptuoti vidaus procedūras, sustiprinti duomenų valdymo kompetencijas ir pasirengti techniniams bei organizaciniams pokyčiams.

Lietuvos mokslo ir studijų institucijų padėtis rodo, kad už LMT tiesiogiai nustatytų reikalavimų ribų MTD valdymas vis dar plėtojamas netolygiai. Apklausos duomenys leidžia matyti, kad daugelyje institucijų MTD valdymas dar nėra įtvirtintas kaip nuosekli ir savarankiška vidaus politika. Dažniau jis reguliuojamas atskirais elementais – projektų reikalavimais, doktorantūros tvarkos aprašais, etikos nuostatomis ar finansuojančių institucijų nustatytomis sąlygomis. Tai rodo, kad institucijose MTD valdymas neretai vystomas kaip atsakas į išorinius reikalavimus, o ne kaip vientisa sistema, apimanti visą duomenų gyvavimo ciklą. Dėl to praktika Lietuvoje išlieka

nevienoda: vienose institucijose jau ryškėja aiškesnės vidaus taisyklės ir duomenų valdymo planavimo praktika, o kitose MTD klausimai vis dar sprendžiami fragmentiškai, atskiromis procedūromis ar pavieniais projektais (žr. apklausos rezultatus 4.1 poskyryje). Kitų mokslinius tyrimus finansuojančių institucijų, ypač veikiančių inovacijų politikos ir verslo MTEP srityje, viešuose dokumentuose daugiausia dėmesio skiriama MTEP veikloms dokumentuoti, apskaitai taryti ir rezultatams pagrįsti. MTD valdymo reikalavimai šiose srityse dažniausiai nėra taip aiškiai ir nuosekliai išplėtoti. Nors ES fondų priemonėse atviri skaitmeniniai išteklių, duomenų bazės ar atvirojo kodo programinė įranga buvo pripažinti tinkamais MTEP rezultatais, tai labiau rodo atvirumo elementų įtraukimą į rezultatų vertinimą, o ne nuoseklios nacionalinės MTD valdymo politikos susiformavimą už LMT ribų. Todėl apskritai galima teigti, kad Lietuvoje finansuotojų požiūris į MTD valdymą išlieka nevienodas: LMT juda aiškesnio ir privalomo reguliavimo kryptimi, o kitose sistemose MTD klausimai vis dažniau siejami su projektine atskaitomybe, inovacijų diegimu ir rezultatų naudojimu.

2.1.4. Apibendrinimas

ES MTD politika per pastarąjį dešimtmetį iš fragmentiškų AP iniciatyvų išaugo į nuoseklia, horizontalią skaitmeninės transformacijos strategiją, kurioje duomenys suvokiami kaip esminis ekonomikos augimo, inovacijų, dirbtinio intelekto plėtros ir geopolitinio konkurencingumo išteklius. ADD, DVA ir DA sudaro tarpusavyje papildančią reguliavimo sistemą, kuri apima tiek viešojo, tiek privataus sektoriaus duomenis ir nustato sąlygas jų pakartotiniam naudojimui. Nacionaliniu lygmeniu šie tikslai perkeliama per Lietuvos mokslo ir studijų įstatymą bei LMT 2024 m. patvirtintą Atvirojo prieigos prie MTEP rezultatų tvarkos aprašą, kuris transformuoja ankstesnę rekomendacinę praktiką į privalomą ir aiškiai apibrėžtą procedūrą. Tokiu būdu atvirumas tampa ne pasirinkimu, o numatyta taisykle, grindžiama principu „kiek įmanoma atvira, kiek būtina uždara“.

Šioje sistemoje centrinę infrastruktūrinę reikšmę įgyja EOSC, veikiantis kaip federuota, daugiadalykė ir sąveiki platforma, jungianti nacionalines ir temines saugyklas į bendrą Europos MTD erdvę. EOSC neapsiriboja techniniu duomenų saugojimu – ji kuria aplinką, kurioje duomenys, paslaugos ir skaitmeniniai įrankiai tampa tarpusavyje suderinami, atitinkantys FAIR principus ir tinkami pakartotinai naudoti tiek žmonių, tiek mašinų, įskaitant dirbtinio intelekto sistemas.

Strateginiu požiūriu tai reiškia, kad Lietuvos mokslinių tyrimų teisinis reguliavimas turi būti ne tik formaliai suderintas su ES teise, bet ir techniškai bei semantiškai integruotas į EOSC federaciją. Priešingu atveju kyla rizika, kad Lietuva liks duomenų teikėja be galimybės visapusiškai dalyvauti aukštos pridėtinės vertės duomenų analizės, DI modelių mokymo ir tarptautinių tyrimų grandinėse.

Nepaisant sukurto reguliacinio pagrindo, išlieka kelios esminės grėsmės. Pirmą, reguliacinė fragmentacija ir lėtas nacionalinių įgyvendinamųjų teisės aktų atnaujinimas gali sukelti teisinį neapibrėžtumą, kuris ypač pavojingas aukštųjų technologijų ir DI srityse. Antra, technologinė pažanga, ypač dirbtinio intelekto ir didžiųjų duomenų analizės srityje, vyksta sparčiau nei teisėkūros procesai, todėl yra rizika, kad pernelyg formalistinis teisės normų aiškinimas taps inovacijų stabdžiu. Trečia, praktikoje išlieka duomenų sąveikumo problema – nevienodi metaduomenų standartai, skirtingos saugyklų architektūros ir nevienodas FAIR principų taikymas apsunkina tarpdisciplininį ir tarpvalstybinį bendradarbiavimą.

Atsižvelgiant į tai, rekomenduotina stiprinti nuolatinę ES duomenų teisės stebėseną ir operatyvų jos perkėlimą į nacionalinę praktiką, plėtoti sąveikumo standartų darninimą bei užtikrinti sistemingą institucinių infrastruktūrų integraciją su EOSC. Taip pat būtina institucionalizuoti duomenų valdymo kompetencijas, stiprinti tyrėjų mokymą FAIR ir duomenų teisės klausimais, integruoti AM rodiklius į mokslininkų vertinimo sistemas. Tik tokiu būdu AM taps ne deklaratyvia, o veiksminga inovacijų, konkurencingumo ir tarptautinio bendradarbiavimo platforma.

Apibendrinant teigtina, kad MTD politika ES pereina iš normatyvinės fazės į infrastruktūros ir strateginę fazę, kurioje duomenys tampa ne tik skaidrumo, bet ir ekonominės galios šaltiniu. EOSC šiame kontekste yra praktinė priemonė, leidžianti realizuoti FAIR principus, stiprinti tarpvalstybinį bendradarbiavimą ir sudaryti sąlygas DI pagrindu vykdomiems tyrimams. Lietuvos uždavinys – užtikrinti, kad nacionalinė teisinė ir infrastruktūros sistema būtų lanksti, sąveiki ir strategiškai orientuota į aktyvų dalyvavimą Europos MTD ekosistemoje.

2.2. Mokslinių tyrimų duomenų FAIR, CARE, TRUST principai

2.2.1. FAIR principai

Kaip jau buvo užsiminta Įvade ir 2.1.2 poskyryje, ES MTD politika ir EOSC neatsiejami nuo FAIR duomenų principų įgyvendinimo. FAIR duomenų ir sąveikių paslaugų tinklo kūrimas nurodomas kaip vienas EOSC europinio mazgo tikslų (EOSC Association, 2026), o organizacijos, kandidatuojančios tapti EOSC mazgais, privalo pasirašyti susitarimo memorandumą, į kurį įtrauktas įsipareigojimas palaikyti FAIR principus ir siekti jų įgyvendinimo (*Memorandum of Understanding on Preparing Operational Integration within the Envisaged EOSC Federation*, n.d.; EOSC Association, 2026).

FAIR principai – tai gairės, padedančios didinti MTD tinkamumą pakartotinai naudoti, ypatingą svarbą teikiant automatizuotam jų suradimui ir naudojimui (Mons et al., 2020). FAIR principai buvo suformuluoti 2014 m. Leidene, Nyderlanduose, vykusiam seminare²¹ ir po dvejų metų paskelbti kaip straipsnis žurnale *Scientific Data* (Wilkinson et al., 2016).

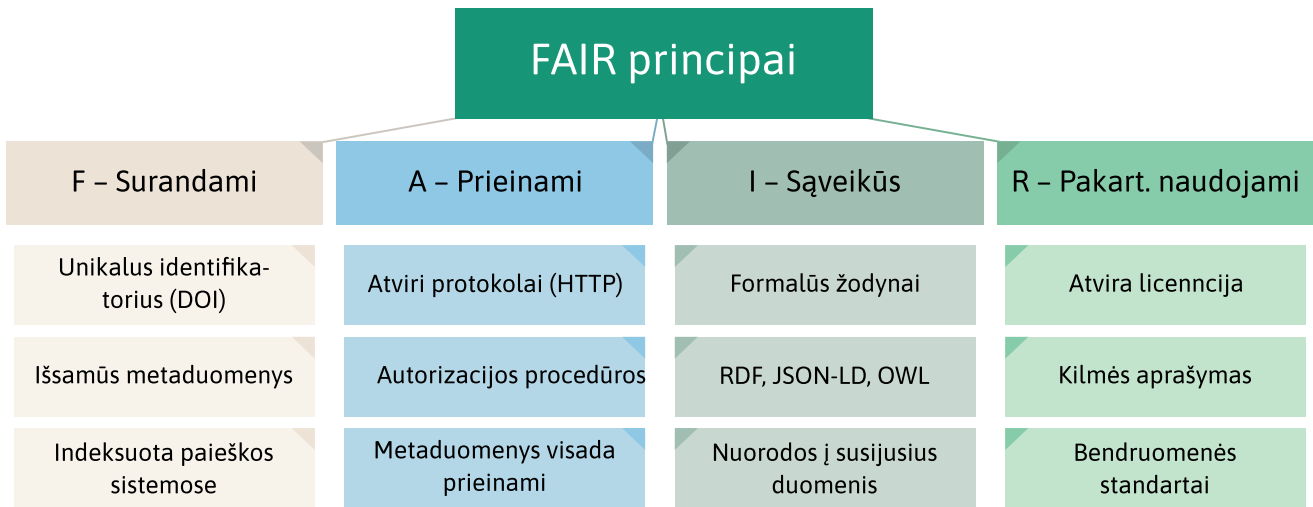
Jau netrukus po jų paskelbimo, FAIR principai susilaukė plataus politinio palaikymo. 2016 m. jie paminėti Europos Komisijos Mokslinių tyrimų ir inovacijų generalinio direktorato informaciniame leidinyje, jų įgyvendinimą nurodant kaip vieną iš tuo metu dar tik planuojamo kurti EOSC tikslų (European Commission, 2016b). Tais pačiais metais jie buvo įtraukti į programos „Horizontas 2020“ (2014–2020) atvirųjų duomenų bandomosios programos *Open Research Data Pilot* rekomendacijas (European Commission, 2016a), palaikymą jiems išreiškė ir G20 valstybių lyderiai, susitikę Kinijos Hangzhou mieste²². Šiuo metu FAIR principų laikytis reikalauja nemažai mokslinius tyrimus finansuojančių organizacijų, įskaitant didžiausią šiuo metu veikiančią Europos mokslinių tyrimų finansavimo programą „Europos horizontas“ (European Commission, 2025a). Verta pažymėti, kad FAIR principai atgarsio sulaukė ne tik tarp mokslo politikos formuotojų ir finansuojančių organizacijų, bet ir mokslo bendruomenėje. Principus pristatęs straipsnis nuo paskelbimo 2016 m. iki 2026 m. pavasario buvo pacituotas beveik 15 tūkstančių kartų (žurnalo *Scientific Data* duomenimis). Be to, per šį laikotarpį paskelbta nemažai kitų straipsnių, aptariančių FAIR principų įgyvendinimą skirtingose mokslo srityse (Mons et al., 2020; van Reisen et al., 2020) arba jų pritaikymą kitiems nei duomenys skaitmeniniams objektams: darbų sekoms (Goble et al., 2020; Wolf et al., 2021), programiniam kodui (Chue Hong et al., 2022; Lamprecht et al., 2019), kompiuterinėms simuliacijoms (Stueber & Frey, 2022) ir objektų skaitmeniniams dvyniams (Peters & Schindler, 2024). Nuolat gausėjančiose mokslinėse publikacijose ir konferencijų pranešimuose aptariami ne tik teoriniai FAIR principų pritaikymo aspektai, bet ir siūlomos praktinių sprendimų koncepcijos ar aprašomi jau įvykę bandomieji projektai (pvz. Celebi et al., 2020; da Silva Santos et al., 2023a; da Silva Santos et al., 2023b; Frey & Hellmann, 2021; Paulus & Biskup, 2023; Schultes et al., 2020; Singh et al., 2025).

FAIR principų turinys

FAIR – tai akronimas, lengvai įsimenama forma pateikiantis duomenų paruošimo mašiniam apdorojimui ir tinkamam bei patogiam tvarkymui gaires. Kaip minėta GS Įvade, akronimas sudarytas iš rekomenduojamas duomenų savybes nusakančių anglišku būdvardžių pirmųjų raidžių: *Findable* (Surandami), *Accessible* (Prieinami), *Interoperable* (Sąveikūs), *Reusable* (Pakartotinai naudojami). Toliau pateikiamas kiekvienos iš šių sąvokų paaiškinimas ir jai įgyvendinti keliami reikalavimai.

²¹ <https://www.lorenzcenter.nl/jointly-designing-a-data-fairport.html>

²² http://www.g20chn.com/xwxzEnglish/sum_ann/201609/t20160906_3397.html

1 pav. FAIR principų apibūdinimas**F – Findable (Surandami duomenys)**

Surandamumo principas reikalauja, kad:

- (meta)duomenims būtų priskirtas unikalus ir nuolatinis identifikatorius (pvz., DOI, URN, „Handle“). Naudojant tokius identifikatorius, išvengiama dviprasmybių automatiškai ar rankiniu būdu ieškant duomenų rinkinių (McMurry et al., 2017). Dažniausiai unikalūs identifikatoriai priskiriami pačios duomenų saugyklos. Šie identifikatoriai turi būti ilgalaikiai – dažnai naudojamos registrų paslaugos, užtikrinančios, kad internetinės nuorodos su identifikatoriais būtų aktyvios. Puikus globalaus ir unikalaus identifikatoriaus pavyzdys – DOI, priskiriamas duomenų rinkiniams tokiose saugyklose kaip „Zenodo“ ar „Figshare“. Kiekvienam įkeltam duomenų rinkiniui automatiškai suteikiamas atskiras DOI (pvz., 10.5281/zenodo.1234567), kuris išlieka aktyvus nepriklausomai nuo to, ar originalus URL pasikeičia;
- duomenys būtų aprašyti išsamiais metaduomenimis, o metaduomenyse aiškiai būtų nurodytas duomenų identifikatorius. Pagrindinis šios rekomendacijos tikslas – kad vartotojas (arba algoritmas) galėtų tiksliai ir greitai rasti tam tikrą duomenų rinkinį, naudojant šiam rinkiniui priskirtus metaduomenis, net nežinant identifikatoriaus.
- duomenys ir (ar) metaduomenys būtų užregistruoti ir indeksuoti paieškos sistemoje ar duomenų bazėse. Siekiant, kad duomenų rinkiniai iš tikrųjų galėtų būti lengvai surandami, jie turi būti indeksuoti ir registruoti įvairiose paieškos sistemoje (tokiose kaip „Google“, „OpenAire Explore“, „PubMed“). Net labai tvarkingas duomenų rinkinys gali būti neprieinamas nei automatiškai, nei rankiniu būdu, jei jis nėra randamas paieškos sistemomis.

A – Accessible (Prieinami duomenys)

Prieinamumo principas reikalauja, kad:

- (Meta)duomenys ir duomenys būtų pasiekiami pagal identifikatorių, naudojant atvirus interneto protokolus. Jei prieiga apribota, turi būti taikomos aiškiai reglamentuotos autorizacijos procedūros.
- Metaduomenys ir/arba duomenys būtų prieinami nemokamais ir atvirais protokolais, pasiekiamais kiekvienam interneto naudotojui (pvz., HTTP, FTP, SMTP). Metaduomenų nereikėtų laikyti mokamose arba nuosavybinėse (angl. *proprietary*) sistemoje (pvz., „Microsoft“, „Apple“ duomenų bazėms tvarkyti skirti produktai).
- Metaduomenys liktų pasiekiami, net jei pats duomenų rinkinys nebėra pasiekiamas.

I – Interoperable (Sąveikūs duomenys)

Sąveikumo principas reikalauja, kad:

- (Meta)duomenyse būtų naudojami formaliai apibrėžti, plačiai pripažinti, atnaujinami žodynai, formalūs terminai.
- Vartotojai galėtų lengvai nuskaityti ir interpretuoti vieni kitų metaduomenis. Turi būti naudojamos aktyvios, palaikomos ir plačiai naudojamos programavimo kalbos bei įrankiai – RDF, JSON LD, OWL.
- Metaduomenyse būtų pateikiamos nuorodos į susijusius (meta)duomenų rinkinius ar semantinius modelius.

R – Reusable (Pakartotinai naudojami duomenys)

Pakartotinio panaudojimo principas reikalauja, kad:

- Duomenų atributai būtų tinkamai aprašyti, būtų nurodyta aiški, atvira licencija, pateiktas nuodugnus duomenų kilmės aprašymas.
- Metaduomenys atitiktų disciplinos arba bendruomenės standartus.

Atitikties FAIR principams vertinimas

FAIR principų kūrėjai sąmoningai pasirinko ganėtinai laisvą rekomendacijų formuluotę, taip atverdami kelią plačiai techninių realizavimo sprendimų įvairovei (Mons et al., 2017). Vis dėlto galima įžvelgti ir šio sprendimo trūkumų. Pagrindinis jų – dėl principų formuluotės neapibrėžtumo kylantys iššūkiai vertinant, ar ir kaip kokybiškai vienas ar kitas skaitmeninis objektas atitinka FAIR principus. Sprendžiant šią problemą, buvo sukurta keletas atitikties FAIR principams vertinimo įrankių. Dalis siūlomų įrankių, pvz., „ARDC FAIR Data Self-Assessment Tool“²³ ar „How FAIR are your Data?“ (Jones & Grootveld, 2017), yra skirti savarankiškam duomenis kuriančių mokslininkų bei tyrėjų įsivertimui. Tokio pobūdžio įrankiai dažniausiai veikia paprasčiausio kontrolinio klausimyno principu. Kitos sudėtingesnės vertinimo metodikos orientuotos į duomenų tvarkytojus ir juos saugančias infrastruktūras. Dalis šių metodikų (pvz., Wilkinson et al., 2018, 2019) buvo pasiūlytos pačių FAIR principų kūrėjų, kitas pasiūlė atskiros FAIR principus taikančios bendruomenės kitos suinteresuotos šalys (pvz., Ammar et al., 2020; Clarke et al., 2019; F-UJI²⁴; Gaignard et al., 2023; Huber, 2025; RDA FAIR Data Maturity Model Working Group, 2020). FAIR principų atitiktį racionaliau matuoti ne binariškai („taip“ / „ne“), bet laipsniškai, todėl sudėtingesni vertinimo metodai rezultatai išreiškia konkrečiais skaitiniais rodikliais. Taip pat dažnu atveju, be teorinio metodikos aprašymo, siūlomi ir programiniai įrankiai, leidžiantys tą metodiką automatizuotai pritaikyti praktikoje.

Su FAIR principais susiję tarptautiniai projektai ir informaciniai ištekliai

FAIR principams populiarinti ir įgyvendinimo sprendimams kurti buvo vykdyta įvairių tarptautinių ir regioninių projektų. Didžiausi iš jų – GO FAIR ir FAIRsFAIR. GO FAIR projektas buvo inicijuotas Vokietijos, Prancūzijos ir Nyderlandų švietimo ir mokslo ministerijų, o vėliau išsiplėtė į kelias kitas šalis. Viena projekto dalis – 2017 m. įsteigta GO FAIR iniciatyva²⁵ – iki 2022 m. koordinavo metodinės medžiagos bei įrankių, padedančių įgyvendinti FAIR principus, kūrimą, veikiant trimis pagrindinėmis kryptimis: technologijų vystymo („GO Build“), kultūros formavimo („GO Change“) ir kompetencijų ugdymo („GO Train“). Greta iniciatyvos įkurtas GO FAIR fondas²⁶ veiklą vykdo iki šiol. Jis rūpinasi FAIR principų populiarinimu ir principų įgyvendinimo specialistų rengimu bei sertifikavimu. Iš programos „Horizontas 2020“ lėšų 2019–2022 m. finansuoto projekto FAIRsFAIR²⁷ tikslas buvo skatinti MTD valdymą pagal FAIR principus Europoje ir pasiūlyti praktinių šių principų įgyvendinimo sprendimų, orien-

²³ <https://ardc.edu.au/resource/fair-data-self-assessment-tool/>

²⁴ <https://www.f-uji.net>

²⁵ <https://www.go-fair.org>

²⁶ <https://www.gofair.foundation>

²⁷ <https://cordis.europa.eu/project/id/831558>

tuotų į visą MTD gyvavimo ciklą. Daugiausia dėmesio projekte buvo skiriama FAIR duomenų kultūros ir gerosios praktikos puoselėjimui. Projekto interneto svetainė vis dar pasiekiamą, joje galima rasti metodinės medžiagos²⁸ bei naudingų įrankių²⁹ katalogus. Be šių dviejų projektų, paminėtina ir FAIRsharing³⁰ – bendruomeninė iniciatyva, koordinuojama Oksfordo universiteto mokslininkų. Ji parengė ir prižiūri duomenų ir metaduomenų standartų, duomenų bazių bei kitų su FAIR principu susijusių išteklių katalogą³¹.

CARE ir TRUST principai

Nors ES mokslo politika ir EOSC vadovaujasi visų pirma FAIR principais, egzistuoja ir keletas kitų principų ar kokybės sertifikavimo schemų, į kuriuos būtų naudinga atsižvelgti kuriant nacionalinę MTD ekosistemą. Nemažai jų buvo sumanyti kaip FAIR principų papildymas: CARE principai duomenims (Carroll et al., 2020), TRUST principai duomenų saugykloms (Lin et al., 2020), „CoreTrustSeal“ sertifikavimas duomenų saugykloms (CoreTrustSeal Board, 2016), FIDELIS projekto Skaidrių ir patikimų saugyklų charakteristikų matrica (TTRAMatrix) (L’Hours et al., 2025) ir pan.

FAIR principų (Wilkinson et al., 2016) paskelbimas sujungia MTD valdymo pasaulį ir, viena vertus, paskatino ieškoti praktinių sprendimų, kaip juos įgyvendinti (Bailo et al., 2020; Cabrera et al., 2020; Krisnawijaya et al., 2025; RDA FAIR Data Maturity Model Working Group, 2020), kita vertus, inicijavo svarbius susijusius procesus, kurių metu buvo sukurti (arba išplėtoti) papildomi su MTD valdymu susiję principai, tvarkos ir taisyklės.

Toliau aptariami CARE principai duomenims ir TRUST principai duomenų saugykloms.

2.2.2. CARE principai

CARE principai duomenims buvo pradėti kurti 2018 m. Tarptautinės duomenų savaitės praktiniame seminare „Vietinių bendruomenių duomenų suvereniteto principai jų duomenų valdymui“, kuris organizuotas kartu su Tyrimų duomenų aljanso Tarptautine vietinių bendruomenių duomenų suvereniteto interesų grupe. Principai buvo išdėstyti straipsnyje „The CARE Principles for Indigenous Data Governance“ (Carroll et al., 2020) ir koncentravosi į vietinių bendruomenių (čiabuvių) teises pačioms kontroliuoti duomenis apie save siekiant gerovės ir lygybės. Šiuo metu linkstama į tai, kad šie principai gali būti aktualūs ne tik konkrečių vietinių bendruomenių (čiabuvių) duomenims, bet ir kitų (ypač nedidelių) kultūrinių bendruomenių (pvz., Lietuvos ar jos regionų) duomenims (O’Brien et al., 2024).

Vietinių bendruomenių (čiabuvių) duomenys CARE principuose apima (Carroll et al., 2020, p. 3): 1) informaciją ir žinias apie aplinką, teritorijas, dangų, išteklius ir ne žmones (gyvūnus ir pan.), su kuriais vietinių bendruomenių atstovai palaiko kokius nors santykius; 2) informaciją apie atskirus vietinių bendruomenių asmenis, įskaitant administracinius, surašymo, sveikatos, socialinius, prekybos ir verslo duomenis; 3) informaciją ir žinias apie vietines bendruomenes kaip kolektyvus, įskaitant tradicinę ir kultūrinę informaciją, žodinę istoriją, informaciją apie protėvius ir klanus, kultūrai svarbias vietas, pasakojimus ir kultūrinius artefaktus.

CARE principų esmė ir aktualumas

CARE principai yra orientuoti ne į technologijas, o į žmones, jų interesus ir duomenų naudojimo tikslus. Todėl jie nėra FAIR principų alternatyva, o juos papildo. Jie apibrėžia bendro pobūdžio principus, taikytinus mokslinių tyrimų, valdžios institucijų ir organizacijų duomenų valdymo aplinkose (Carroll et al., 2020, p. 5):

- **Kolektyvinė nauda:** duomenys turi kurti naudą vietinėms bendruomenėms ir prisidėti prie jų savivaldos, gerovės, paslaugų gerinimo, inovacijų.
- **Teisė kontroliuoti:** vietinės bendruomenės turi teisę nustatyti duomenų prieigos, naudojimo ir saugojimo sąlygas.

²⁸ <https://www.fairsfair.eu/competence-centre/training-library>

²⁹ <https://www.fairsfair.eu/tools-software>

³⁰ <https://fairsharing.org/communities#governance>

³¹ <https://fairsharing.org/search>

- **Atsakomybė:** duomenų naudotojai turi įsipareigoti gerbti vietos bendruomenių teises, kultūrą, stiprinti jų gebėjimus naudoti duomenis keliant vietinių bendrovių gerovę.
- **Etika:** duomenų naudojimas turi būti etiškas, nedaryti žalos, būti suderintas su vietinių bendruomenių vertybėmis ir gerovės plėtra.

Taigi CARE principai papildo FAIR principus pridėdami socialinius, kultūrinius ir etinius aspektus prie duomenų valdymo sampratos ir užtikrina, kad AM praktikos būtų etiškos ir teisingos. CARE principai aktualūs ne tik vietinių bendruomenių (čiabuvių) kontekste, bet ir kitoms bendruomenėms, siekiančioms kolektyvinių, su jomis susijusių duomenų teisių apsaugos: etninėms mažumoms, religinėms bendruomenėms, istorinėms bendruomenėms (pvz., regioninėms tapatybėms, vietinių tradicijų paveldui) ar skaitmeninėms bendruomenėms, kurių duomenys gali būti jautrūs. CARE principų laikymasis skatina atsakingą duomenų rinkimą, lygiavertę partnerystę su bendruomenėmis, duomenų grįžtamąją vertę, jautrių duomenų apsaugą, o tai atliepia Lietuvos bioetikos, kultūros paveldo ir duomenų apsaugos reglamentavimą. CARE principų laikymasis taip pat užtikrina, kad atviri duomenys netaptų nauja duomenų kolonizacijos forma, kai dominuoja anglų kalba, bet prisidėtų prie teisingumo ir partnerystės skatinimo.

Lietuvos mokslininkai jau dabar naudoja duomenis, kurie patenka į CARE reglamentavimo sritį: jūrų ir vandenynų duomenis, genominius duomenis, antropologinius rinkinius, muziejų artefaktus. Šiose srityse CARE principų taikymas jau yra privalomas arba rekomenduojamas tarptautiniuose projektuose, o tai prisideda prie institucinio pasirengimo laikytis tarptautinių etinių standartų, kurie bus svarbūs jungiantis į EOSC.

CARE principų svarba EOSC kontekste

EOSC vis labiau orientuojasi į atsakingą duomenų valdymą ir socialinį poveikį, taigi CARE principai tiesiogiai papildo šiuos tikslus. FAIR principų (kaip pagrindinio EOSC ekosistemos kūrimo komponento) kontekste CARE principų integracija vertinama kaip būtina norint kurti etiškai tvarius duomenų ekosistemų standartus. Daugelis EOSC rekomendacijų skatina įtraukti teisių turėtojų valdymo mechanizmus, kuriuos CARE principai dar aiškiau ir tiksliau apibrėžia (pvz., prieigos kontrolės).

CARE principai svarbiausi šiose integruojamos į EOSC srityse:

- Metaduomenų praplėtimas vietinio konteksto (vietinėms bendruomenėms svarbiomis) žymomis ir nuorodomis (duomenų kilmė).
- Prieigos valdymo instrumentai (embargai, ribotos prieigos modeliai).
- Duomenų kilmės fiksavimas ir skaidri komunikacija.
- CARE principų įtraukimas į EOSC sąveikumo sistemą.

CARE principų integracija į EOSC dar nėra baigta, tačiau pasaulinės duomenų organizacijos (RDA, GIDA) ją aktyviai formuoja.

2.2.3. TRUST principai

TRUST principai duomenų saugykloms buvo pradėti kurti 2019 m. Tyrimų duomenų aljanso 13-ajame plenariniame susitikime, o jų kūrimą koordinavo Tyrimų duomenų aljanso (ir Pasaulio duomenų sistemos) Skaitmeninių saugyklų sertifikavimo interesų grupė. Principai buvo išdėstyti straipsnyje „The TRUST Principles for digital repositories“ (Lin et al., 2020) ir apibendrino charakteristikas, kuriomis turi pasižymėti patikimos ir gerai valdomos skaitmeninės duomenų saugyklos.

TRUST principai suteikia aiškų, tarptautinį ir mokslinės bendruomenės priimamą pagrindą duomenų saugykloms. Tai nėra tik gairės – tai strateginė būtinybė siekiant, kad Lietuvos akademinė bendruomenė taptų visa-verte Europos atvirosios mokslo erdvės dalyve. TRUST principų įgyvendinimas yra bendrasis vardiklis patikimoms saugykloms, padedantis įgyvendinti gerąsias praktikas, užtikrinti duomenų valdymo procesų skaidrumą, patiki-

mumą ir atskaitomybę. Saugyklos, kurios įgyvendina savo veikloje TRUST principus, geriau užtikrina duomenų integralumą, prieinamumą ir ilgalaikį išsaugojimą.

Ypač svarbu tai, kad FAIR duomenys gali egzistuoti tik saugyklose, kurios laikosi TRUST principų, nes jie indikuoja tvarią infrastruktūrą, metaduomenų standartų įgyvendinimą, patikimą prieigą prie duomenų ir ilgalaikį jų kuravimą. Be to, TRUST principai užtikrina tinkamą jautrių duomenų tvarkymą (ypač teisinius ir etinius aspektus), ilgalaikį finansinį ir organizacinį tvarumą bei atsižvelgimą į kintančius vartotojų poreikius. Galų gale TRUST principų laikymasis ir įgyvendinimas yra kelias į saugyklų sertifikavimą („CoreTrustSeal“, „nestor Seal“, ISO16363), nes šie principai yra tiesiogiai susieti su tarptautiniais patikimų skaitmeninių saugyklų (angl. *Trustworthy Digital Repository*) standartais (Kleemola & Wong, 2025).

TRUST principų esmė ir aktualumas

Nors šiuo metu akademinė bendruomenė yra susitelkusi į FAIR principų įgyvendinimą duomenų kūrimo ir valdymo procesuose, tačiau vis labiau akivaizdu, kad vien FAIR principų neužtenka. FAIR principai kalba apie tai, kokie turi būti duomenys, tačiau šie principai negali būti įgyvendinti be patikimų saugyklų. Pavyzdžiui, CESSDA ERIC infrastruktūra pažymi, kad FAIR principai gali veikti tik tada, kai saugyklos turi tvarią, skaidrią ir technologiškai patikimą infrastruktūrą, o tai yra tiesioginiai TRUST principų reikalavimai (Kleemola & Wong, 2025).

ES siekia, kad FAIR duomenys būtų ne tik tinkamai kuriami, bet ir patikimai saugomi, prieinami ir pakartotinai naudojami ilgą laiką. Todėl TRUST principai yra esminė EOSC federacinės infrastruktūros dalis, o EOSC mazgams ir paslaugų teikėjams būtina įrodyti patikimumą, skaidrumą ir ilgalaikį finansinį bei organizacinį tvarumą. Šiuo metu Europoje ir pasaulyje egzistuoja didžiulė duomenų fragmentacija, o TRUST principai suteikia galimybę šią fragmentaciją sumažinti, kurti bendrą infrastruktūrą, bendrus vertinimo kriterijus, konfigūruojamus, bet bendrus reikalavimus saugykloms. TRUST principai padeda standartizuoti skirtingų ES šalių ir disciplinų duomenų tvarkymą, gerinti sąveikumą, lengvina duomenų dalijimąsi tarp tyrėjų ir institucijų. Taigi TRUST principai sudaro bendrą standartų sistemą, kurią gali priimti skirtingos šalys, infrastruktūros ir disciplinos.

TRUST principai (Lin et al., 2020):

- *Skaidrumas*: duomenų saugyklos viešai ir aiškiai pateikia savo misiją, politiką, duomenų naudojimo sąlygas, duomenų saugojimo trukmę, paslaugas ir jautrių duomenų valdymo galimybes.
- *Atsakomybė*: duomenų saugyklos prisiima atsakomybę už duomenų autentiškumą, integralumą, ilgalaikį saugojimą, teises ir jautrių duomenų apsaugą. Tai apima metaduomenų standartų laikymąsi, (meta)duomenų kokybės kontrolę ir nuoseklią bei ilgalaikę duomenų priežiūrą.
- *Orientacija į naudotojus*: duomenų saugyklos turi būti įsitvirtinusios savo bendruomenių praktikose ir reaguoti į jų poreikius: užtikrinti duomenų radimą, suprantamumą, grįžtamąjį ryšį, suteikti reikalingus įrankius rasti ir naudoti duomenis.
- *Tvarumas*: turi būti užtikrinamas ilgalaikis ir tinkamas duomenų saugyklų finansavimas ir organizacinis tvarumas, sukurti tinkami dokumentai, rizikų valdymo, tęstinumo ir perdavimo planai.
- *Technologinis pajėgumas*: duomenų saugyklos privalo turėti patikimą technologinę infrastruktūrą, standartus, saugumo priemones, technologinę brandą, palaikyti reikiamus semantinius ir techninius standartus.

TRUST principų praktinį įgyvendinimą užtikrina įvairios sertifikavimo (pvz., CoreTrustSeal³²) ir patikimų skaitmeninių saugyklų (TDR) iniciatyvos (pvz., FIDELIS³³). Svarbu paminėti, kad TRUST principų įgyvendinimas yra finansiškai ir organizaciškai brangus (santykinai) procesas. Todėl, siekiant Lietuvoje įgyvendinti TRUST principus, reikia gerai suplanuoti, kiek ir kokių saugyklų reikia, nes tinkamam jų išlaikymui reikia nemažai išteklių. Kita vertus, valstybinis finansavimas ir TRUST principų palaikymas svarbus diegiant ir užtikrinant technologinę saugyklų infrastruktūrą bei jų organizacinį tvarumą (ypač kalbant apie kibernetinį saugumą ir perdavimo politiką).

³² www.coretrustseal.org

³³ <https://eden-fidelis.eu>

TRUST principų svarba EOSC kontekste

EOSC siekia sukurti patikimą ir sąveikią europinę bendros duomenų prieigos ekosistemą. EOSC federacijos pagrindas – patikimi mazgai, todėl saugyklos, įgyvendinančios TRUST principus, yra natūralūs ir būtini EOSC ekosistemos elementai. TRUST principai garantuoja, kad duomenų paslaugos yra skaidrios, gerai dokumentuotos, technologiškai suderinamos, pritaikytos vartotojų poreikiams ir atitinkančios ilgalaikio saugojimo reikalavimus. O tai yra pagrindiniai reikalavimai, integruojant saugyklas į EOSC federaciją.

TRUST principų svarba Lietuvos akademinėi bendruomenei

Visų pirma TRUST principų įgyvendinimas yra svarbi sąlyga kuriant tvarią nacionalinę duomenų valdymo ekosistemą. TRUST principai padeda atliepti poreikį sukurti stabilų, ilgalaikį patikimų nacionalinių duomenų saugyklų finansavimą ir valdymą Lietuvoje, tai ypač aktualu akademinėi bendruomenei, kuri susiduria su fragmentuota duomenų laikymo praktika. Lietuvoje dar nėra nė vienos sertifikuotos (taigi TRUST principus visiškai atitinkančios) duomenų saugyklos, ir vos kelios institucijos planuoja sertifikavimą ateityje (remiantis antrojo institucijų tyrimo rezultatais, žr. 4.1 poskyrį). Todėl Lietuvos nacionalinė politika turėtų ypač akcentuoti TRUST principų įgyvendinimą saugyklose, to naudą ir svarbą.

Ne mažiau svarbu ir tai, kad TRUST principų įgyvendinimas (kurio pagrindinis instrumentas – sertifikavimas) užtikrina tinkamą Lietuvos saugyklų integraciją į tarptautines infrastruktūras, nes tik patikimos ir tvarios saugyklos gali būti priimamos į tarptautinius tinklus (pavyzdžiui, tokius kaip EOSC). Be to, TRUST principų laikymasis didina Lietuvos mokslinių duomenų matomumą ir naudojimą, nes skaidrūs ir vartotojams suprantami duomenys tampa lengviau randami, interpretuojami ir pakartotinai naudojami tarptautinėse tyrimų ekosistemose.

Galų gale TRUST principų įgyvendinimas turėtų stiprinti ir duomenų valdymo kultūrą universitetuose, nes jis padėtų sukurti vientisą sistemą, kuri prisidėtų prie kokybiškesnio duomenų valdymo planavimo, patikimesnio etinių ir teisinių standartų laikymosi bei didesnės atskaitomybės finansuotojams.

2.2.4. Apibendrinimas

Atsižvelgiant į FAIR principų svarbą EOSC ir į jų privalomumą EOSC federacijos nariams, rekomenduojama, vystant MTD ekosistemą Lietuvoje ir rengiant ją integracijai į EOSC, stengtis užtikrinti bent jau minimalų šių principų įgyvendinimo lygį.

Išteklius, reikalingus FAIR principams įgyvendinti, galima skirti į dvi grupes: technologinius ir žmogiškuosius. Technologijos ypač svarbų vaidmenį atlieka užtikrinant duomenų prieinamumą (A) ir sąveikumą (I), yra reikalingos ir įgyvendinant kitus du principus (F ir R). Vis dėlto pastaruosius du principus įgyvendinti neįmanoma be žmogaus įsikišimo. Kad būtų užtikrintas duomenų surandamumas (F), duomenų autoriai arba jiems padedantys duomenų valdytojai turi pateikti kuo išsamesnius metaduomenis. Kad būtų užtikrintas duomenų pakartotinis naudojimas (R), duomenų autoriai turi aprašyti duomenų kilmę, suteikti jiems licenciją ir pasirūpinti, kad metaduomenys būtų pateikti pagal jų dalyko bendruomenės standartus. Kad FAIR principai būtų įgyvendinami nacionaliniu mastu, Lietuva turės, viena vertus, užtikrinti, kad visuose duomenų ekosistemos infrastruktūriniuose komponentuose būtų naudojamos FAIR principams įgyvendinti tinkamos ir tarpusavyje sąveikios technologinės priemonės, kita vertus, plėsti duomenis kuriančių tyrėjų ir institucijų duomenų valdymo kompetencijas, siekti, kad duomenų aprašymas metaduomenimis, identifikatorių bei standartinių licencijų naudojimas ir disciplininė bendruomenėse priimtų standartų laikymasis taptų neatskiriama mokslinių tyrimų kultūros dalimi. Taip pat prasminga būtų investuoti į duomenų valdytojų, padėsiančių tyrėjams sutvarkyti duomenis pagal FAIR principus, tinklo sukūrimą.

Kad būtų užtikrintas tvarus FAIR principų laikymasis, svarbu ne vien supažindinti duomenų ekosistemos dalyvius su FAIR principų reikalavimais, bet ir, naudojantis jau egzistuojančiais sprendimais, numatyti konkrečią atitikties šiems principams vertinimo metodiką bei sistemingai ją taikyti.

FAIR principus papildo CARE principai, kurie orientuoti ne į technologijas, o į žmones, jų interesus ir duomenų naudojimo tikslus. CARE principų laikymosi užtikrinimas galimas tik tinkamai tobulinant duomenis kuriančių

tyrėjų ir institucijų duomenų valdymo kompetencijas bei puoselėjant etišką duomenų kūrimo ir naudojimo kultūrą. Tai vėlgi susiję ir su investicijomis į duomenų valdytojų kompetencijų ugdymą.

FAIR (ir CARE) principus įgyvendinti duomenų kūrimo ir valdymo procesuose neužtenka. FAIR principai kalba apie tai, kokie turi būti duomenys, tačiau šie principai negali būti įgyvendinti be patikimų saugyklų. FAIR principus atitinkantys duomenys gali egzistuoti tik saugyklose, kurios laikosi TRUST principų, nes jie indikuoja tvarią infrastruktūrą, metaduomenų standartų įgyvendinimą, patikimą prieigą prie duomenų ir ilgalaikį jų kuravimą. Todėl TRUST principai yra esminė EOSC federacinės infrastruktūros dalis, o EOSC mazgams ir paslaugų teikėjams būtina įrodyti patikimumą, skaidrumą ir ilgalaikį finansinį bei organizacinį tvarumą. TRUST principai garantuoja, kad duomenų paslaugos yra skaidrios, gerai dokumentuotos, technologiškai suderinamos, pritaikytos vartotojų poreikiams ir atitinkančios ilgalaikio saugojimo reikalavimus. TRUST principai suteikia aiškų, tarptautinį ir mokslinės bendruomenės priimamą pagrindą duomenų saugykloms. Tai nėra tik gairės, tai strateginė būtinybė norint, kad Lietuvos akademinė bendruomenė būtų visavertė Europos atvirosios mokslo erdvės dalyvė. TRUST principų įgyvendinimas turėtų stiprinti ir duomenų valdymo kultūrą universitetuose, nes jis padėtų sukurti vientisą sistemą, kuri prisidėtų prie kokybiškesnio duomenų valdymo planavimo, patikimesnio etinių ir teisinių standartų laikymosi bei didesnės atskaitomybės finansuotojams.

2.3. Mokslinių tyrimų duomenų tipai ir standartai, sąveikumas, valdysena ir tvarkysena, duomenų saugyklos

2.3.1. Mokslinių tyrimų duomenų tipai ir standartai

Norint užtikrinti, kad MTD būtų randami, suprantami, sąveikūs ir pakartotinai naudojami, nepakanka vien jų sukaupti ar paskelbti. EOSC kontekste esminę reikšmę turi du susiję dalykai: aiškus duomenų tipų atpažinimas ir tinkamų standartų taikymas. Duomenų tipai leidžia suprasti, kokio pobūdžio objektai kuriami ir saugomi, o standartai nustato, kaip šie objektai aprašomi, kokiais formatais pateikiami, kaip susiejami su metaduomenimis ir kaip gali būti naudojami kitų sistemų, tyrėjų ar paslaugų.

MTD kontekste standartai yra svarbūs ne vien techniniam duomenų saugojimui, bet ir jų prasmingam interpretavimui, integracijai ir pakartotiniam naudojimui skirtingose institucinėse, disciplininėse ir tarptautinėse aplinkose. Siekiant užtikrinti sąveikumą EOSC ekosistemoje, aktualūs yra keli tarpusavyje susiję standartizacijos lygmenys. Pirma, metaduomenų standartai apibrėžia, kaip nuosekliai aprašomi duomenų rinkiniai, jų kilmė, autoriai, sukūrimo aplinkybės, prieigos ir naudojimo sąlygos. Antra, specifinės srities standartai leidžia atsižvelgti į konkrečių mokslo sričių specifiką, pavyzdžiui, socialinių mokslų, geoduomenų, kalbos išteklių ar bioinformatikos srityse. Trečia, failų ir duomenų formatų standartai užtikrina, kad duomenys būtų saugomi atvirais, dokumentuotais ir ilgalaikiam naudojimui tinkamais formatais. Ketvirta, kontroliuojami žodynai, taksonomijos ir ontologijos sudaro prielaidas vienodai interpretuoti duomenų reikšmes ir sąvokas skirtingose sistemose. Penkta, nuolatiniai identifikatoriai, tokie kaip DOI, ORCID (angl. *Open Researcher and Contributor ID*) ar organizacijų identifikatoriai, leidžia patikimai susieti duomenų rinkinius su autoriais, institucijomis ir kitais mokslinės komunikacijos objektais. Būtent šių lygmenų visuma sudaro prielaidas ne tik techniniam, bet ir semantiniam bei organizaciniam duomenų sąveikumui.

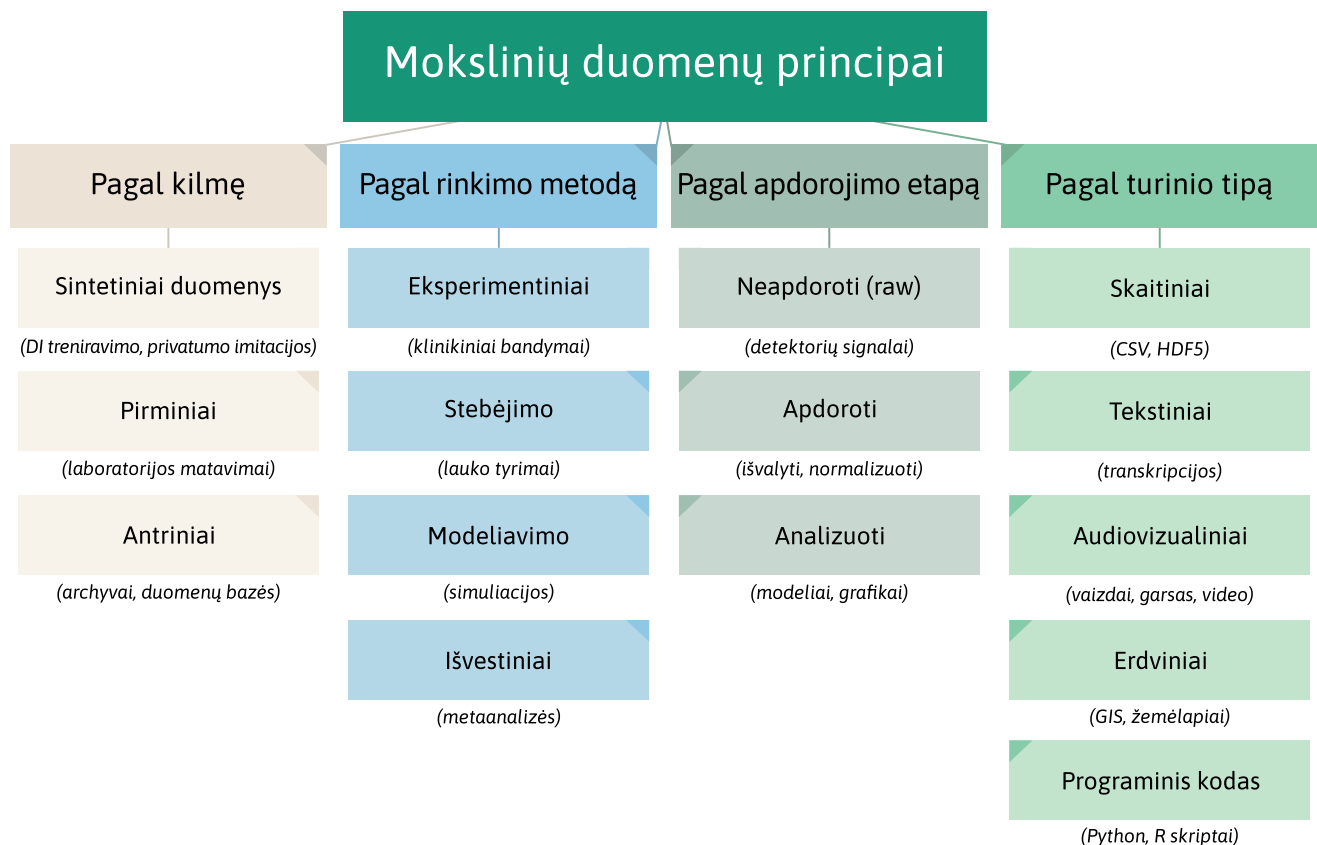
Informacinių technologijų (toliau – IT) srityje egzistuoja daugybė duomenų ir failų formatų. Pavyzdžiui, dar 1996 m. buvo dokumentuota apie 100 paplitusių formatų vien grafiniams duomenims kompiuteriuose užrašyti (Murray & vanRyper, 1996). Kai kurie iš minimų formatų (pvz., TIFF, PNG) naudojami mokslo duomenims užrašyti iki šiol. Be to, mokslo tyrimuose naudojamų duomenų formatų skaičius gerokai padidėjo. Tą lėmė IT vystymasis, naujų tyrimo metodų ir kompiuterių programų atsiradimas bei augantis supratimas apie duomenų mainų ir standartizacijos svarbą.

Atskiri duomenų formatai yra pritaikyti tam tikrų matavimų, skaičiavimų rezultatų arba tam tikros mokslo srities duomenims užrašyti. Gerai specifikuoti duomenų formatai įgalina verifikuoti mokslinių rezultatų duomenis

ir palengvina keitimąsi duomenimis bei duomenų publikavimą pagal FAIR principus. Dėl šios priežasties išsami naudojamų duomenų formatų Lietuvoje apžvalga yra naudinga MTD kokybei įvertinti ir Lietuvos institucijų parengtumo lygiui FAIR duomenų mainams nustatyti.

Mokslinių duomenų tipai įprastai skirstomi į keturias grupes pagal jų kilmę, rinkimo metodą, apdorojimo etapą ir pagal turinio tipą (žr. 2 pav.).

2 pav. Mokslinių duomenų tipai



2026-ieji žymi lūžio tašką, kai AM principai susilieja su DI poreikiais. Duomenys nebevertinami tik kaip publikacijų priedai. Kaip minėta 2.1.2 poskyryje, ši transformacija grindžiama ES DI akto³⁴ taikymo termino pradžia (2026 m. rugpjūčio 2 d.) bei ES DI mokslui strategijos (European Commission, 2025c) nuostatomis, kurios numato RAISE (European Commission, 2025c) infrastruktūros plėtrą. Pagal EOSC SRIA 2.0 gaires, duomenys evoliucionuoja iš pasyvių objektų į DI agentams pritaikytus (angl. *Agent-Ready*) išteklius, kurie būtini siekiant realizuoti autonominių laboratorijų ir didelio masto skaičiavimų viziją EMTE (European Commission, 2025b; European Commission, 2025c; Regulation (EU) 2023/2854).

2026 m. mokslinių duomenų erdvės nebeapsiriboja vien tekstiniais failais ar lentelėmis. Duomenų tipų įvairovė ir jų standartizacijos lygis tiesiogiai lemia šalies mokslinį konkurencingumą EOSC ekosistemoje.

³⁴ <https://artificialintelligenceact.eu/>

Pasaulyje įsitvirtina supratimas, kad tradicinis FAIR principų taikymas yra būtina, bet nepakankama sąlyga DI proveržiui³⁵. Įvyksta standartų poslinkis – nuo žmogui skaitomų metaduomenų pereita prie mašininio būdu nuskaitymų FAIR. Dominuoja tokie formatai kaip „Zarr“ ir „Parquet“, optimizuoti debesijos skaičiavimams ir mašininio mokymosi modelių treniravimui. Dėl griežtėjančių privatumo reikalavimų (pvz., JAV ir Azijos rinkose) vis dažniau standartizuojami sintetiniai duomenų rinkiniai, leidžiantys treniruoti DI modelius nepažeidžiant duomenų subjektų teisių. Pasaulinėje praktikoje plinta vadinamųjų duomenų pasų (angl. *Datasheets for Datasets*) standartas, reikalaujantis fiksuoti ne tik tai, kas yra duomenyse, bet ir jų kilmę (angl. *provenance*), etinį šališkumą ir paruošimo algoritmams metodiką.

EOSC federacija ir teisinis sąveikumas

EOSC kontekste teisinė darna ar suderinamumas pirmiausia reiškia teisinį sąveikumą, t. y. gebėjimą skirtingoms organizacijoms, veikiančioms pagal nevienodas teisines sistemas, politiką ir procedūras, dalyvauti bendroje federacijoje pagal suderintas bendras taisykles. EOSC plėtroje pabrėžiama, kad vien techninio ir semantinio sąveikumo nepakanka. Tam, kad duomenys ir paslaugos galėtų būti saugiai bei teisėtai naudojami tarpvalstybinio ir tarpdisciplininiu mastu, būtina aiški taisyklių visuma dėl prieigos, priimtino naudojimo, privatumo ir duomenų apsaugos, kibernetinio saugumo, licencijavimo ir atsakomybės pasidalijimo. Būtent todėl EOSC federacijos kūrimo procese greta EOSC-IF formuojamos ir bendros federacijos politikos, apimančios prieigos, priimtino naudojimo, privatumo ir duomenų apsaugos bei kibernetinio saugumo reikalavimus. Europa įgyvendina EOSC federacijos modelį, kuriame atskiros duomenų erdvės jungiamos į tarpusavyje sąveikas ekosistemas. Šią kryptį stiprina ir teisinė aplinka: įsigaliojęs ES DI aktas ir Duomenų aktas didina reikalavimus duomenų kokybei, standartizacijai ir saugiam dalijimuisi, ypač aukštos rizikos srityse, tokiose kaip medicina.

Praktiniu lygmeniu šį procesą atspindi EOSC mazgų plėtra. 2025 m. pabaigoje pradėję veikti pirmieji 14 EOSC mazgų prisidėjo prie bendros Europos MTD erdvės formavimo. Šioje sistemoje svarbų vaidmenį įgyja ir „Gaia-X“ pasitikėjimo sistema (Gaia-X European Association for Data and Cloud, 2024), orientuota į duomenų suverenitetą ir saugų dalijimąsi tarp mokslo bei pramonės.

Kartu vis labiau akcentuojamas semantinis suderinamumas Europos mastu: plėtojami ir standartizuojami tarpdisciplininiai žodynai, leidžiantys DI algoritmams jungti duomenis iš skirtingų šaltinių (pvz., klimato kaitos duomenis su sveikatos statistika).

Mokslinių duomenų tipų klasifikacija (Globali tendencija)

Pasaulinėje praktikoje duomenys skirstomi ne tik pagal mokslo sritis, bet ir pagal jų generavimo pobūdį bei tinkamumą mašininiam mokymuisi:

- Stebėjimų duomenys (angl. *observational data*): realiojo laiko jutiklių rodmenys, teleskopų nuotraukos, nuotolinio zondavimo (palydovų) duomenys.
- DI reikalingi duomenų rinkiniai: reikalauja didelio pralaidumo srautinių standartų.
- Eksperimentiniai duomenys (angl. *experimental data*): laboratorinių tyrimų rezultatai (pvz., genų sekos, chromatogramos). Paprastai saugomi specifiniais gamintojų (angl. *proprietary*) formatais, kuriuos siekiama konvertuoti į atvirus standartus.
- Simuliacijų ir modeliavimo duomenys (angl. *simulation data*): aukšto našumo skaičiavimų infrastruktūra (angl. *high-performance computing*) (toliau – HPC) generuoti skaitmeniniai modeliai (pvz., klimato prognozės).
- *Status quo*: tai milžiniškos apimties duomenys, kur standartizacija vyksta per masyvų formatus (pvz., *NetCDF, HDF5*).
- Išvestiniai / apdoroti duomenys (angl. *derived/compiled data*): duomenų gavybos būdu (angl. *data mining*), tekstų analizės ar mašininio mokymo modelių sugeneruoti rezultatai.

³⁵ Svarbu pastebėti, kad atviri ir FAIR duomenys svarbūs ir klasikiniam algoritmomams, ne tik DI/mašininio mokymosi metodams.

• Sintetiniai duomenys (angl. *synthetic data*): algoritmiškai sukurti duomenys, naudojami mašininio mokymosi modeliams treniruoti ten, kur tikri duomenys yra riboti arba jautrūs. Šie duomenys aktualūs mažų imčių problemoms spręsti (kai tam tikrų reiškinų stebėjimų yra per mažai DI mokyti).

Europa per EOSC-IF sąveikumo sistemą remiasi ne vien duomenų formatų suvienodinimu, bet platesniu techninio, semantinio, organizacinio ir teisinio sąveikumo modeliu. Praktiniu lygmeniu tai reiškia, kad EOSC turi būti suderinama su skirtingose mokslo srityse jau taikomais bendruomeniniais standartais ir metaduomenų schemomis, pavyzdžiui, gyvybės moksluose naudojamasi FASTQ, BAM, VCF (genomika) formatais, socialiniuose moksluose taikomas DDI (apklausos), SDMX (statistika) ar humanitariniuose moksluose paplitusiomis

TEI (tekstų kodavimas) gairėmis. Kartu vis didesnė reikšmė teikiama mašininio būdu interpretuojamiems ir sąveikiams metaduomenims, kurie leistų duomenis ne tik aprašyti, bet ir automatiškai atrasti, susieti bei pakartotinai naudoti skaitmeninėse tyrimų ir mašininio mokymosi bei DI aplinkose. Europa akcentuoja ir mažųjų duomenų (angl. *small data*) kokybę – net ir nedideli duomenų rinkiniai turi būti aprašyti taip, kad DI agentai galėtų juos automatiškai integruoti į didesnes imtis. Itin svarbu metaduomenų standartai: pereinama nuo aprašomojo lygio (*kas tai?*) prie funkcinio (*ką su tuo gali daryti algoritmas?*), naudojant Schema.org ir DCAT-AP schemas.

Techniniai ir semantiniai reikalavimai duomenims DI ir EOSC kontekste

DI plėtra keičia ne pačius MTD valdymo principus, bet jų taikymo mastą ir griežtumą. Jei anksčiau pakakdavo, kad duomenys būtų išsaugoti, dokumentuoti ir tam tikru mastu prieinami žmogui, tai šiuolaikinėje duomenų aplinkoje vis svarbiau, kad jie būtų tinkami ir automatizuotam apdorojimui, jungimui, analizei bei pakartotiniam naudojimui mašininio būdu. Dėl to duomenų kokybė vis dažniau vertinama ne vien pagal jų išsamumą ar prieinamumą, bet ir pagal tai, kiek jie yra mašininio būdu interpretuojami, struktūruoti, semantiškai aprašyti ir atsekami.

Šiuolaikinėje MTD aplinkoje duomenų tinkamumas DI taikyti neturėtų būti suprantamas kaip vieno DI formato turėjimas. Kur kas svarbiau tai, ar duomenys ir jų metaduomenys yra pakankamai struktūruoti, standartizuoti ir mašininio būdu interpretuojami, kad juos būtų galima automatiškai aptikti, susieti, apdoroti ir pakartotinai naudoti (Scardaci et al., 2023; Wilkinson et al., 2016). Būtent tokį požiūrį įtvirtina FAIR principai, pabrėžiantys ne tik žmogaus, bet ir mašinos galimybę automatiškai rasti ir naudoti duomenis. Panašia kryptimi plėtojamas ir EOSC sąveikumo karkasas, kurio stiprinimas siejamas su papildomais metaduomenimis, leidžiančiais mašinoms dinamiškai suprasti, kaip pasiekti paslaugas ir išteklius, taip sudarant prielaidas automatizuotam išteklių komponavimui federuotoje aplinkoje (Scardaci et al., 2023).

Techniniu požiūriu tai reiškia, kad DI taikymams tinkama duomenų aplinka turi remtis disciplinoje pripažintais formatais ir metaduomenų modeliais, o ne vien bendrojo pobūdžio failų saugojimu. Gyvybės moksluose tokį vaidmenį atlieka genomikos formatai, tokie kaip SAM/BAM ir VCF, naudojami sekoskaitos skaitymų ir genetinės variacijos duomenims aprašyti bei saugoti. Socialiniuose, elgsenos ir ekonomikos moksluose analogišką funkciją atlieka DDI standartų šeima, skirta kiekybinių ir kokybinių tyrimų duomenims aprašyti per visą jų gyvavimo ciklą, o statistinių duomenų mainams plačiai taikomas SDMX standartas. Humanitariniuose moksluose panašią reikšmę turi TEI gairės, skirtos struktūruotam tekstų kodavimui ir mašininio būdu apdorojamiems tekstiniams ištekliams kurti. Todėl duomenų tinkamumas DI pirmiausia kyla ne iš universalaus AI formato, bet iš to, ar duomenys pateikiami pagal konkrečiai sričiai prasmingas ir mašininio būdu apdorojamas bendruomenines specifikacijas³⁶

Ne mažiau svarbus yra anotavimo, proveniencijos ir versijavimo sluoksnis. DI sistemoms nepakanka turėti vien žaliuosius duomenis. Privalo būti aišku, kas, kaip, kada ir kokiomis transformacijomis tuos duomenis sukūrė, pažymėjo ar pakeitė. W3C PROV modelis³⁷ provenienciją apibrėžia kaip informaciją apie subjektus, veiklas ir asmenis, dalyvavusius sukuriant konkretų duomenų objektą, ir tiesiogiai sieja ją su kokybės, patikimumo ir pasitikėjimo vertinimu. W3C Web Annotation Data Model³⁸ pateikia struktūruotą modelį anotacijoms kurti ir keistis jomis tarp skirtingų sistemų. Praktikoje tai reiškia, kad DI mokymui tinkama infrastruktūra turi palaikyti ne tik duomenų laikymą, bet ir nuoseklų anotacijų, proveniencijos bei eksperimentų istorijos fiksavimą. Tokie įrankiai kaip DVC

³⁶ Plačiau žr. <https://ddialliance.org/ddi-codebook/>

³⁷ <https://www.w3.org/TR/prov-dm/>

³⁸ <https://www.w3.org/TR/annotation-model/>

(angl. *Data Version Control*) šiuo požiūriu svarbūs kaip praktinė įgyvendinimo priemonė, nes leidžia bendro darbo metu susieti duomenų, kodo ir modelių versijas bei atkurti eksperimentų eigą, tačiau jie laikytini ne universaliu standartu, o viena iš priemonių atsekamumui ir atkartojamumui užtikrinti.³⁹

Europos duomenų erdvės kontekste papildomai svarbus katalogavimo ir atrandamumo lygmuo. DCAT-AP (angl. *Data Catalog Vocabulary Application Profile*) Europoje naudojamas duomenų rinkinių aprašams ir paieškai tarp portalų, o jo semantinis suderinamumas su kitais žodynais ir schemomis leidžia skirtingų šaltinių aprašus padaryti palyginamus ir pakartotinai naudojamus. Todėl duomenų tinkamumas DI taikymams turėtų būti siejamas ne su pavienėmis technologijomis, pavyzdžiui, vektorinių bazių naudojimu, bet su platesniu techninių savybių rinkiniu: nuolatiniais identifikatoriais, turtingais ir mašininio būdu interpretuojamais metaduomenimis, disciplininiais formatais, struktūruotu anotavimu, proveniencija ir versijavimo mechanizmais⁴⁰ (žr. taip pat Scardaci et al., 2023).

Lietuvos situacijoje ši perspektyva atskleidžia ir potencialą, ir ribojamuosius veiksnius. Apklausų duomenys rodo, kad skirtingose mokslo srityse kuriami nevienodi duomenų tipai, todėl ir jų standartizacijos poreikiai skiriasi: gamtos, medicinos ir technologijų moksluose dažniau dominuoja eksperimentiniai, modeliavimo, vaizdiniai ir programinės įrangos duomenys, o humanitariniuose ir socialiniuose moksluose – apklausų, tekstiniai, dokumentiniai, audiovizualiniai bei skaitmenizuoti archyviniai išteklių (žr. 1 lentelę). Tai patvirtina, kad Lietuvos integracija į EOSC negali būti grindžiama vienu techniniu modeliu ar vienodu reikalavimų rinkiniu visiems duomenims. Kita vertus, apklausos taip pat rodo, kad metaduomenų išsaugojimo, FAIR principų supratimo ir nuoseklus standartų taikymo lygis tebėra netolygus, todėl būtent metaduomenų valdysena ir institucinių gebėjimų stiprinimas išlieka viena reikšmingiausių prielaidų pasirengimui dalyvauti federuotoje Europos mokslo duomenų erdvėje (apklausų rezultatai žr. 4 skyrių).

1 lentelė. Apklausos duomenimis išryškėję pagrindiniai duomenų tipai

Srities grupė	Vyraujantys duomenų tipai	Pastabos standartizacijai
GMT	Eksperimentiniai (84,0 %), modeliavimo ir skaičiavimų (54,0 %), vaizdiniai (52,7 %), programinė įranga ir kodas (28,0 %).	Aktualūs didelės apimties, instrumentiniai ir skaičiavimų duomenys; svarbūs atviri formatai, versijavimas ir aiškus techninis aprašymas.
HSM	Apklausų (79,3 %), tekstiniai ir dokumentiniai (74,4 %), garso ir vaizdo (35,4 %), skaitmenizuoti archyviniai šaltiniai.	Ypač svarbūs semantiniai aprašai, kontekstas, licencijos, autoriaus teisės ir domeniniai metaduomenų standartai.
Bendri tipai	Stebėjimų, išvestiniai ir erdviniai duomenys paplitę abiejose sričių grupėse panašiu mastu.	Tinkami horizontaliems sprendimams, bendriems metaduomenų profiliams ir tarpdisciplininiam atradimui EOSC aplinkoje.

Atsižvelgiant į šiuos aspektus, Lietuvos pasirengimas dalyvauti EOSC turėtų būti vertinamas ne vien pagal turimų saugyklų ar sukauptų rinkinių skaičių, bet ir pagal tai, kaip nuosekliai institucijos identifikuoja savo kuriamus MTD tipus, kaip sistemiškai taiko atitinkamus standartus ir kiek jų MTD gali būti integruojami į federuotą, mašininio būdu veikiančią Europos MTD ekosistemą. Šiuo požiūriu prioritetu turėtų tapti bendro minimalaus metaduomenų ir sąveikumo sluoksnio stiprinimas, domeninių standartų taikymo plėtra, anotavimo, proveniencijos ir versijavimo praktikų brandinimas bei aiškesnis MTD valdymo reikalavimų įtvirtinimas instituciniu lygmeniu.

³⁹ <https://doc.dvc.org/example-scenarios/versioning-data-and-models> ir <https://www.w3.org/TR/annotation-model/>

⁴⁰ <https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/solution/dcat-application-profile-data-portals-europe/release/300>

2.3.2. Mokslinių tyrimų duomenų ir sistemų semantinis, organizacinis, teisinis ir techninis sąveikumas

Kaip jau minėta 2.1.2 poskyryje, tarptautinėje erdvėje MTD valdysena evoliucionuoja nuo statiško duomenų saugojimo prie dinamiško jų įveiklinimo, kurio ašis – AM paradigma. Šis požiūris iš esmės keičia žinių kūrimo paradigmą, traktuojant MTD kaip viešąjį gėrį, kuris turi būti ne tik prieinamas, bet ir visiškai sąveikus. Šio kokybinio lūžio politinį pamatą įtvirtina naujoji Europos duomenų sąjungos strategija (European Commission, 2025e), kurios tikslas – eliminuoti duomenų fragmentaciją ir sukurti vientisą „duomenų greitkelį“. Duomenų sąjungos vizijoje MTD pozicionuojami kaip strateginis kuras Europos DI modeliams, pabrėžiant, kad atvirumas ir aukšta duomenų kokybė yra būtinos mokslo proveržio sąlygos.

Strateginė ES kryptis, įtvirtinta DGA veikia kaip pasitikėjimo mechanizmas, įgalinantis duomenų altruizmą ir saugų dalijimąsi jautriais ištekliais. Kartu Sąveikios Europos aktas (angl. *Interoperable Europe Act*) (toliau – IEA) nustato sąveikumo pagal dizainą standartus, kurie užtikrina, kad mokslo infrastruktūros būtų suderinamos su bendrosiomis Europos duomenų erdvėmis. Ypatinę postūmį šiai evoliucijai suteikia Europos strategija dėl DI moksle (European Commission, 2025c), nubrėžianti gaires RAISE ekosistemai kurti. Ši iniciatyva reikalauja, kad atvirojo mokslo valdysena būtų orientuota į DI parengtumą (angl. *AI-readiness*), integruojant DI etikos ir mašininio mokymosi reikalavimus, numatytus DI taikymo strategijoje.

MTD ir sistemų sąveikumas operacionalizuojamas per Europos sąveikumo pagrindus (angl. *European Interoperability Framework*) (toliau – EIF) bei specializuotą mokslo srities plėtinį – EOSC-IF. Europos sąveikumo sistema sąveikumą apibrėžia kaip organizacijų gebėjimą sąveikauti siekiant abipusiškai naudingų tikslų, įskaitant informacijos ir žinių mainus tarp šių organizacijų per jų palaikomus procesus, keičiantis duomenimis tarp jų informacinių ir ryšių technologijų sistemų⁴¹. EOSC sąveikumo sistema skirta naudotis visiems subjektams, prisidedantiems prie EOSC kūrimo ir diegimo, suteikia bendrą supratimą apie reikalavimus, iššūkius ir rekomendacijas, į kurias jie turėtų atsižvelgti, taip pat apie bendrą principų rinkinį, kaip šias rekomendacijas galima įgyvendinti.

Semantinis sąveikumas

Semantinį sąveikumą (angl. *semantic interoperability*) galima apibrėžti kaip „kompiuterinių sistemų gebėjimą perduoti duomenis su nedviprasmiška bendra reikšme. Semantinis sąveikumas yra reikalavimas, leidžiantis naudoti mašininio būdu skaičiuojamą logiką, išvadas, žinių atradimą ir duomenų federaciją tarp informacinių sistemų“ (Lehväsliho et al., 2019).

Tai reiškia, kad semantinis sąveikumas pasiekiamas, kai perduodama informacija, jos perduota forma turi visą reikšmę, reikalingą priimančiajai sistemai, kad galėtų ją teisingai interpretuoti, net jeigu siuntėjas nežino, kokiais algoritmais remsis gavėjas. Sintaksinis sąveikumas (kuris dažnai siejamas su techniniu sąveikumu) kartais įvardijamas kaip semantinio sąveikumo prielaida. Jis užtikrina, kad tikslus keičiamų duomenų ir informacijos formatus ir reikšmė būtų išsaugomi ir suprantami per visų šalių mainus, kitaip tariant, „kas siunčiama, tas ir suprantama“.

Semantikos samprata išplečiama ir peržengia „vien mašininio“ lygmens ribas: semantinis sluoksnis reikalauja, kad ir žmonės (bendruomenės, paslaugų teikėjai, naudotojai) būtų pakankamai susiderinę, kad komunikacija būtų sklandi. Pavyzdžiui, paslaugų lygio susitarimai (angl. *Service Level Agreement*, toliau – SLA) turi remtis bendrais semantiniais artefaktais, reikia turėti minimalų bendrų metaduomenų formatų rinkinį ir bendrus semantinius artefaktus (pvz., ontologijas, tezaurus) tiek tyrėjų bendruomenių viduje, tiek tarp jų, kartu užtikrinant aiškią šių artefaktų valdyseną.

EOSC-IF minimalūs reikalavimai ir rekomendacijos semantiniam sąveikumo sluoksniui

EOSC sąveikumo sistema (IF) (European Commission, 2021a) semantinio sąveikumo sluoksnį nagrinėja per FAIR skaitmeninių objektų (angl. *FAIR Digital Objects*, toliau – FDO) modelį. EOSC ekosistemoje siekiama, kad paslaugų teikėjai, duomenų teikėjai ir paslaugų vartotojai turėtų bendrą sutartą būdą, kaip skaitmeniniai objektai yra

⁴¹ <https://interoperable-europe.ec.europa.eu/collection/iopeu-monitoring/1-introduction-european-interoperability-framework>

skaitomi, interpretuojami ir naudojami. Tam reikalingas susitarimas dėl nuorodų į bendrus išteklius, kurie aprašo objekto reikšmę ir kitus sąveikumui kritinius aspektus.

Esminė mintis – FDO metaduomenys nėra vienas monolitinis blokas. Jie suprantami kaip sluoksniuota metaduomenų struktūra, kur atskiri metaduomenų elementai nukreipia į skirtingus sąveikumo poreikius (techninius, semantinius, organizacinius, teisinius). Semantinio sluoksnio požiūriu tai reiškia, kad prasmės užtikrinimas įgyvendinamas nuorodomis (dažnai su nekintančiu identifikatoriumi (angl. *Persistent Identifier*, toliau – PID)) į bendrus semantinius artefaktus, kurie suteikia aiškias, mašinai skaitomas sąvokų ir savybių reikšmes.

Kartu pabrėžiama detalumo (smulkumo) idėja: FDO gali egzistuoti keliais lygiais (pvz., visas duomenų rinkinys ir jo dalys ar net kintamasis), o sąveikumui reikalingos nuorodos į FAIR išteklius turi būti užtikrinamos nepriklausomai nuo to, kuriame lygyje objektas vartojamas. Tokia dekompozicija leidžia semantines nuorodas (į sąvokas, apibrėžimus, schemas) perkelti į smulkesnius vienetus ir taip padidinti paieškos, jungimo ir pakartotinio naudojimo tikslumą.

Semantinis sąveikumas praktiškai remiasi bendriniais ir tyrėjų bendruomenių specifiniais semantiniiais artefaktais, į kuriuos FDO ir jo metaduomenys gali dėti nuorodas. EOSC-IF rekomenduoja visą semantikos spektrą: nuo formalių (ontologijos, tezaurai) iki mažiau formalių artefaktų (UML (angl. *Unified Modeling Language*, toliau – UML) modeliai, duomenų bazių modeliai, XML (angl. *Extensible Markup Language*) schemas). Akcentuojama, kad net mažiau formalūs artefaktai gali stiprinti sąveikumą, jei jie pateikia mašinai interpretuojamas nuorodas į formalesnius apibrėžimus (pvz., UML klasę „Asmuo“ susiejant su ontologijos sąvoka).

Semantinis sluoksnis sukonkretinamas parodant, kaip jį praktiškai įgyvendinti pasitelkiant konceptualias metaduomenų sistemas (angl. *generic metadata frameworks / conceptual metadata standards*) ir duomenų tipų registrų modelius. Jos apibrėžia, kaip metaduomenys turi būti konstruojami ir kokios metaduomenų sąvokos vartojamos, paliekant srities specifiką srities standartams.

Organizacinis sąveikumas

Organizacinis sąveikumas (angl. *organisational interoperability*) reiškia būdą, kuriuo viešojo administravimo institucijos derina savo procesus, pareigas ir lūkesčius, kad pasiektų bendrai sutartus ir abipusiai naudingus tikslus. Praktiškai organizacinis sąveikumas reiškia procesų ir atitinkamos keičiamos informacijos dokumentavimą ir integravimą arba suderinimą. Organizacinio sąveikumo tikslas – taip pat patenkinti naudotojų poreikį turėti paslaugų pasiekiamumą, lengvą atpažįstamumą, prieinamumą ir orientaciją į naudotojus (European Commission, 2017).

Atsižvelgiant į bendrą AM tikslą, kuriuo grindžiama visa EOSC veikla, šis sąveikumo lygis turėtų būti sukelktas į skirtingų organizacijų, teikiančių paslaugas EOSC, procesų dokumentavimą, integravimą ar suderinimą, siekiant užtikrinti, kad tyrėjai galėtų pasiekti AM tikslus. Taip pat turėtų būti aišku, kas yra atsakingas už bendrų sąveikumo paslaugų, tokių kaip paslaugų katalogai, registrai ir bendros PID paslaugos, teikimą (taip pat kūrimą, priežiūrą ir kuravimą) (European Commission, 2021a, p. 12).

EOSC sąveikumo sistema pateikia šias rekomendacijas dėl organizacinio sąveikumo (European Commission, 2021a, p. 27–28):

- Į sąveikumą orientuotos dalyvavimo rekomendacijos.
- Standartizuotų duomenų formatų ir (arba) žodynų naudojimo rekomendacijos ir atitinkami metaduomenys.
- Aiškus nuolatinių organizacijos pavadinimų ir funkcijų valdymas.

Teisinis sąveikumas

Teisinis sąveikumas (angl. *legal interoperability*) apima įstatymų, politikos, procedūrų ir bendradarbiavimo susitarimų aplinką, reikalingą sklandiems informacijos mainams ir duomenų pakartotiniam naudojimui tarp skirtingų asmenų, organizacijų ir skirtingose jurisdikcijose. Teisinis sąveikumas yra tarp kelių duomenų rinkinių iš skirtingų šaltinių, kai (GEO Data Sharing Working Group, 2014):

- Kiekvieno duomenų rinkinio teisinio naudojimo sąlygos yra aiškiai ir lengvai nustatomos, paprastai automatizuotomis priemonėmis.
- Kiekvienam duomenų rinkiniui nustatytos teisinio naudojimo sąlygos leidžia kurti ir naudoti kombinuotus arba išvestinius produktus.
- Vartotojai gali teisėtai pasiekti ir naudoti kiekvieną duomenų rinkinį nepašydamu duomenų turėtojų leidimo kiekvienu konkrečiu atveju, darant prielaidą, kad tenkinamos sukauptos kiekvieno ir visų duomenų rinkinių naudojimo sąlygos.
- Kai prieiga prie duomenų yra ribojama, metaduomenys yra FAIR, t. y. naudojami priimti standartai duomenims apibūdinti ir taip sudaromos sąlygos juos atrasti.

Atsižvelgiant į EOSC ir FAIR principus, teisinis sąveikumas visų pirma reikalauja, kad duomenis būtų galima pakartotinai naudoti. Tai susiję su galimybe derinti duomenų rinkinius iš kelių šaltinių, nesukeliant konfliktų su kiekvieno duomenų rinkinio licencijos nustatytais apribojimais (Doldirina et al, 2018, p. 8).

EOSC sąveikumo sistema nustato šias teisinio sąveikumo rekomendacijas (European Commission, 2021a, pp. 27–28):

- Standartizuotos žmogaus ir kompiuterio skaitomos licencijos su centralizuotu žinių ir palaikymo šaltiniu autorių teisių ir licencijų klausimais.
- Leidžiančios licencijos metaduomenims (ir, pageidautina, duomenims, kai tik įmanoma). Pirmenybė teikiama CC0 licencijai, o ne CC BY 4.0.
- Skirtingų duomenų rinkinio dalių su skirtingomis licencijomis identifikavimas.
- Aiškiai pažymėti pasibaigusio arba neegzistuojančio autorių teisių atvejai, taip pat ir duomenys, skirti kūriniams „našlaičiams“.
- Aiškus EOSC rekomenduojamų licencijų sąrašas ir jų suderinamumas su valstybių narių rekomenduojamomis licencijomis.
- Duomenų rinkinių licencijų raidos stebėjimas laikui bėgant.
- Suderinta politika ir gairės, kaip elgtis tais atvejais, kai patentų paraiškos pateikimas ar komercinės paslaptys gali būti pažeistos dėl atskleidimo.
- BDAR atitiktis asmens duomenų tvarkymui.
- Papildomi prieigos prie duomenų ir jų naudojimo apribojimai taikomi tik esant atitinkamiems teisės aktams arba esant teisėtiems pagrindams.
- Suderintos naudojimo sąlygos įvairiose saugyklose.
- Valstybių narių nacionalinių teisės aktų ir EOSC suderinimas.

Techninis sąveikumas

Techninis sąveikumo sluoksnis (angl. *technical interoperability*) apibrėžiamas kaip skirtingų informacinių sistemų ir programinių aplikacijų gebėjimas tarpusavyje komunikuoti ir apsikeisti informacija taip, kad sąveika būtų įmanoma nepaisant skirtingų technologinių realizacijų. Šis sluoksnis apima technines priemones ir taisykles, kurios leidžia sukurti patikimą ryšį tarp sistemų bei užtikrina, kad duomenys (ar kiti skaitmeniniai objektai) galėtų būti perduodami iš vienos sistemos į kitą.

Svarbus ne vien duomenų persiuntimo faktas, bet ir operacinis sąveikumo rezultatas. Sistemos turi ne tik išsiųsti ir priimti informaciją, bet ir priimti viena kitos pateikiamus duomenis tokia forma, kad galėtų atlikti reikiamą veiksmą be papildomos žmogaus (operatoriaus) intervencijos. Todėl techninis sluoksnis siejamas su automatizuotu, nuosekliu paslaugų veikimu (angl. *end-to-end*), kai rankinis suderinimas tampa išimtimi, o ne taisykle.

EOSC kontekste techninis sąveikumas peržengia tradicines duomenų mainų ribas: kalbama ir apie kitų tyrimų artefaktų (pvz., programinės įrangos, paslaugų, darbo srautų, protokolų, net aparatinės įrangos dizainų) perdavimą ir suderinamą naudojimą. Tai reiškia, kad techninis sluoksnis turi užtikrinti sąlygas ne tik duomenų sklaidai, bet ir suderintam skaitmeninių tyrimų komponentų naudojimui skirtingose infrastruktūrose ir paslaugų ekosistemose.

Remiantis EIF logika, kuria remiasi EOSC-IF, techninis sluoksnis apima programas ir infrastruktūras, kurios įgalina sistemų sujungimą. Sąsajų specifikacijas, integravimo ir tarpininkavimo paslaugas, duomenų pateikimą ir apsikeitimą, taip pat saugius komunikacijos protokolus. Kartu pažymima, kad kai kuriais atvejais šis sluoksnis gali būti skaidomas į infrastruktūros ir aplikacijų dalis. Tokiu būdu techninis sąveikumas traktuojamas kaip bazinė (bet nepakankama) prielaida platesniam sąveikumui, nes jis sukuria kanalus ir vykdymo terpę, per kurią vėliau gali būti realizuojami semantiniai, organizaciniai ir teisiniai derinimai.

EOSC-IF minimalūs reikalavimai ir rekomendacijos techniniam sluoksniui

EOSC-IF reikalavimai techniniam sluoksniui yra laikomi minimaliais reikalavimais EOSC sąveikumui, išvedamais iš tipinių bendruomenių patiriamų problemų ir poreikių.

Pirma reikalavimų grupė siejama su tapatybės ir prieigos valdymu: praktikoje autentifikavimas ir autorizavimas dažnai vyksta atskirai kiekvienai bendruomenei / paslaugai, todėl naudotojui tenka kartoti prisijungimo ir teisių suteikimo procesus. EOSC-IF pabrėžia poreikį užtikrinti kuo mažiau trikdantį, nuo vienos konkrečios bendruomenės nepriklausomą prisijungimą ir teisių gavimą (angl. *Authentication and Authorization Infrastructure*, AAI), taip pat platesnį saugumo ir privatumo pagrindą patikimiems mainams. Su tuo susijusi ir rekomendacija turėti aiškiai suprantamus paslaugų lygių susitarimus (SLA), kad paslaugų patikimumas ir įsipareigojimai būtų suderinti skirtingoms disciplinoms.

Antra reikalavimų grupė kyla iš duomenų ir tyrimų objektų heterogeniškumo: skirtingos disciplinos naudoja tiek bendrinius formatus (pvz., CSV, JSON, XML), tiek specifinius modelius, kuriuos sunku tarpusavyje suderinti, ypač kai trūksta antraščių, dokumentų ar susitarimų dėl bendrų atributų. Todėl įvardijamas poreikis EOSC kontekste turėti minimalų metaduomenų taikymo profilį (angl. *minimal metadata application profile*), kuris leistų vienodai atrasti ir susitvarkyti su duomenimis, nepaisant jų formatų įvairovės. Rekomenduojama užtikrinti lengvą prieigą prie skirtingų formatų duomenų šaltinių ir įrankių, kurie padeda įveikti heterogeniškumą ir integruoti duomenis tarp bendruomenių.

Trečia reikalavimų grupė – techniniai reikalavimai, siejami su paieškos ir suradimo (angl. *findability*) infrastruktūra. Turi būti galimybė tyrimų duomenis (ir kitus tyrimų objektus) atrasti ne tik stambiu masteliu (angl. *coarse-grained*), bet ir smulkiu masteliu (angl. *fine-grained*), t. y. kartais ieškoma kolekcijų ar teminių rinkinių, o kartais – konkrečių elementų jų viduje. Rekomenduojama turėti paieškos įrankius abiem detalumo lygiams kaip bazinį federuotos paieškos mechanizmą.

Techniniame sluoksnyje svarbus PID ir atvirų specifikacijų vaidmuo. Yra daug PID sistemų, skiriasi jų taikymo politikos, identifikatoriai kartais ne visada tinkamai atpažįstami (angl. *resolved*), todėl reikalavimuose įvardijamas poreikis turėti bendrą, aiškiai suprantamą PID politiką. Kartu rekomenduojama EOSC paslaugas kurti remiantis atviromis specifikacijomis, nes jos yra pagrindinis mechanizmas techniniam suderinamumui užtikrinti ir ilgalaikiam pakartotiniam naudojimui bei vystymui palaikyti.

EOSC sąveikumo sistema – ateities perspektyvos

EOSC sąveikumo sistema buvo paskelbta 2021 m., todėl labai naudinga atsižvelgti į naujas EOSC suderinimo rekomendacijas, paskelbtas 2025 m. rugsėjo 30 d. tyrime „Skaitmeninių teisių valdymo padėtis“ (Hugo et al., 2025).

Tyrime „Skaitmeninių teisių valdymo padėtis“ rekomendacijų rinkiniu siekiama suderinti EOSC praktiką su išsamia teisių valdymo strategija. Šios rekomendacijos apima technines priemones, politikos lygmens veiksmus ir metaduomenų standartizavimo pastangas, atspindint daugialypį iššūkio pobūdį. Pateikiamos tokios pagrindinės rekomendacijos:

- Parengti ir priimti visame EOSC taikomą teisių žymėjimo kalbos (angl. *Rights Markup Language*, toliau – RML) žodyną. Pagrindinis žingsnis – EOSC turėtų užbaigti specialų RML žodyną, kuris plėstų arba profilizuotų esamus standartus (pvz., atvirosios skaitmeninių teisių kalbos (angl. *Open Digital Rights Language*⁴², toliau – ODRL), kad atitiktų EOSC specifinius reikalavimus. Šiame RML turėtų būti numatyti terminai visiems pagrindiniams identifikuotiems teisių elementams. Standartinėms licencijos leidimams / draudimams, embargų laikotarpiams, prieigos sąlygoms (atvira / ribota / uždaryta), įsipareigojimams (šaltinio nurodymas, finansavimo citavimas, pakartotinio naudojimo pranešimas), asmens duomenų indikatoriui, etinio naudojimo apribojimams ir pan. Labai svarbus suderinamumas su nusistovėjusiomis ontologijomis, pavyzdžiui, naudoti ODRL klases ir savybes leidimams bei susijusiems įsipareigojimams, integruoti „DataCite“ terminus pagrindinei licencijų informacijai, įtraukti duomenų naudojimo ontologijos (angl. *Data Use Ontology*)⁴³ ar panašius kodus sutikimo reikalavimams. Žodynas turėtų būti publikuojamas suteikiant jam nuolatinį unifikuotą išteklių identifikatorių (angl. *unified resource identifier*) ir pateikiamas mašininio būdu skaitomu formatu (RDF/OWL), kad būtų lengva integruoti. Daugiafunkcės (daugiakalbės) žymos ir apibrėžtys padės užtikrinti pritaikymą visoje Europoje. Techniniu požiūriu šis žodynas bus teisių metaduomenų schema EOSC aplinkoje.

- Suderinti valdomos prieigos (angl. *access control*) sąvoką ir naujus terminus, susijusius su jautriais duomenimis. Nors tam tikrais atvejais galima diegti teisių išraiškos kalbos (angl. *Rights Expression Language*) ir RML profilius bei plėtinius, yra jautrių duomenų rinkinys, kuriam reikalingas suderintas ir nuoseklus terminų rinkinys, apibūdinantis valdomos prieigos elementus, kurie, tikėtina, bus išreikšti draudimais ir susijusiais įsipareigojimais. Pagrindinis šių terminų tikslas – subjektų teisių apsauga, apimanti tokias nuostatas kaip galutinio naudotojo duomenų įrašymas, gautos duomenų išvestinės patikrinimas prieš ir/arba po sukūrimo, siekiant užtikrinti subjektų teisių apsaugą, bei draudimas naudoti duomenis už kontroliuojamų aplinkų ribų. Kai kuriais atvejais pradiniai duomenys gali būti nematomi galutiniam naudotojui, o prieinami tik išvestiniai duomenys.

- Integruoti RML į saugyklų platformas (techninė integracija). Pagrindinės EOSC paslaugos ir duomenų saugyklų platformos turėtų būti atnaujintos taip, kad palaikytų naują RML. Tai reiškia, kad tokios programinės įrangos kaip „Invenio“ („Zenodo“), „Dataverse“, „DSpace“ ir kt. turėtų leisti naudotojams pasirinkti ir priskirti teisių pareiškimus iš kontroliuojamo žodyno. Pavyzdžiui, duomenų įkėlimo sąsaja galėtų pateikti ne tik licencijos pasirinkimą, bet ir papildomas sąlygas (pvz., „Ar yra asmens duomenų? T/N“, „Ar pakartotiniam naudojimui reikalingas sutikimas? T/N“, „Tik nekomercinis naudojimas? T/N“). Šie pasirinkimai būtų automatiškai įrašomi į metaduomenis kaip RML teiginių rinkinys. Rekomenduojama sukurti ir API bei įrankius teisių informacijai apdoroti. Rinkimo protokolai (OAI-PMH, ResourceSync) turėtų perduoti RML duomenis, o paketavimo standartai, tokie kaip RO-Crate⁴⁴, turėtų turėti teisių informacijos, koduotos RML formatu, skyrių. Labai svarbu, kad teisių metaduomenys būtų įtvirtinti EOSC techniniu lygmeniu, t. y. taptų duomenų mainų struktūros dalimi. Tai galėtų būti palengvinta įtraukiant RML palaikymą į EOSC pagrindines paslaugas, pavyzdžiui, EOSC metaduomenų katalogas ir PID paslaugas galėtų turėti laukus ar susietus duomenis, nurodančius teisių žodyną.

- Numatyti privalomą teisių metaduomenų pateikimą EOSC politikoje. Politikos lygmeniu EOSC turėtų įpareigoti teikti mašininio būdu skaitomus teisių metaduomenis kaip sąlygą prijungti ir finansuoti saugyklas. Kaip „OpenAIRE“ gairės įtvirtino prieigos teisių lauką kaip privalomą, taip EOSC gali pakelti kartelę. Bet kuris duomenų rinkinys, programinė įranga ar išteklius EOSC aplinkoje turi turėti RML pagrįstą teisių pareiškimą. Tai nereiškia, kad viskas turi būti atvira. Tai reiškia, kad turi būti užtikrintas skaidrumas. Net ir ribojamos prieigos duomenys turėtų turėti RML pareiškimą, pvz., „prieiga ribojama + priežastis (asmens duomenys)“ (angl. „*restrictedAccess + reason (personal data)*“), o ne būti tiesiog tyliai apriboti.

- Suderinti metaduomenų schemas / profilius, įtraukiant naujus RML elementus. Rekomenduojama atnaujinti esamas metaduomenų schemas / profilius, kad jos apimtų naujus RML elementus. Pavyzdžiui, naujojo-

⁴² <https://www.w3.org/TR/odrl-vocab/>

⁴³ <https://ega-archive.org/access/data-access-committee/data-use-ontology/#>

⁴⁴ <https://www.researchobject.org/ro-crate/>

je „OpenAIRE“ gairių versijoje turėtų būti įtrauktas skyrius apie tai, kaip įterpti EOSC teisių žodyną (galbūt kaip papildomą rekomenduojamą lauką arba kaip išplėsto teisių elemento dalį). „DataCite“ galėtų apsvarstyti kai kuriuos iš šių elementų būsimose schemas versijose arba pateikti rekomendacijas dėl susijusių identifikatorių ar papildomų teisių metaduomenų naudojimo. Reikėtų paskelbti suderinamumo lentelę, rodančią, kaip įprasti saugyklų metaduomenys („Dublin Core“, „DataCite“, „schema.org“) gali perteikti RML informaciją, kad įgyvendinimas būtų lengvesnis. Iš esmės tikslas – užtikrinti, kad nepriklausomai nuo to, kokį metaduomenų standartą saugykla naudoja viduje, būtų aiškus kelias standartizuotą teisių informaciją atskleisti kitiems.

- Integruoti į EOSC tarpusavio sąveikumo sistemą. EOSC tarpusavio sąveikumo sistema turėtų įtraukti teisinį tarpusavio sąveikumą naudodama RML kaip pagrindinį komponentą. Tai reikšia, kad apibrėžiant, kaip EOSC ištekliai tarpusavyje sąveikauja, teisių metaduomenys būtų traktuojami taip pat svarbiai kaip techniniai metaduomenys. EOSC-IF galėtų pateikti gaires, kaip naudoti RML visuose EOSC mainų sluoksniuose.

EOSC-IF savo sąveikumo logiką grindžia EIF, kuris jau seniai įtvirtintas kaip bendras atskaitos taškas viešojo sektoriaus sąveikumui Europoje. Todėl EOSC-IF keturių sąveikumo sluoksnių (teisinio, organizacinio, semantinio ir techninio) modelis nėra atskiras išradimas – jis perima EIF konceptualų karkasą ir jį pritaiko MTD bei paslaugų federacijos kontekstui.

Pastaruoju metu paskelbtas naujos EIF versijos projektas rodo, kad EIF evoliucionuoja. Nuo bendro principų rinkinio vis labiau pereinama prie įgyvendinimo mechanizmų, aiškesnių privalomų nuostatų ir sąsajų su naujesniu ES reguliaciniu bei technologiniu kontekstu (pvz., duomenų erdvėmis, tapatybės infrastruktūra, atitikties vertinimu). Dėl to šio projekto naujovių supratimas tampa svarbus ne tik viešųjų paslaugų sąveikumui, bet ir EOSC-IF interpretavimui – nes EIF kryptis formuoja platesnę aplinką, kurioje EOSC-IF bus taikomas ir vystomas.

Šioje evoliucijoje reikšmingą vaidmenį atlieka „Next Generation EIF“ neformali EK ekspertų grupė, kuri teikia strategines gaires atnaujinti EIF, siekiant jį pritaikyti skaitmeninei transformacijai ir padaryti veiksmingesnį praktikoje. Grupės siūlymai orientuoti į tai, kad EIF taptų labiau įgyvendinamas, aiškiau suderintas su naujomis technologijomis (pvz., DI, duomenų erdvėmis, skaitmeniniais dvyniais) ir integruotas su IEA instrumentais (pvz., sąveikumo vertinimais, Sąveikios Europos portalu). Kartu akcentuojama, kad EIF atnaujinimas turi sustiprinti esminius principus – orientaciją į vartotoją ir prieinamumą, atvirumą ir pakartotiną panaudojimą, technologinį neutralumą ir perkeliamumą, saugumą ir privatumą, taip pat daugiakalbystę ir skaidrumą.

Atsižvelgiant į tai, EIF projekto analizė leidžia geriau orientuotis, kur link juda sąveikumo politika ir praktika ES mastu, kokių papildomų reikalavimų ar lūkesčių ateityje gali atsirasti EOSC ekosistemai. Kitaip tariant, ku aiškiau suprantamos EIF projekto siūlomos kryptys, tuo tiksliau galima vertinti, kaip EOSC-IF nuostatos derės su naujuoju horizontaliu sąveikumo pagrindu ir kokius pokyčius verta numatyti EOSC įgyvendinimo planuose.

Pirmiausia skiriasi paskirtis ir taikymo laukas. EOSC-IF (European Commission, 2021a) yra sektoriškai orientuotas į mokslinių duomenų ir paslaugų federaciją EOSC ekosistemoje, o jo „minimalūs reikalavimai“ pateikiami kaip tyrėjų bendruomenėse identifikuotų problemų / poreikių ir rekomendacijų rinkinys, stipriai remiantis FDO idėja. Be to, pats dokumentas aiškiai pasako, kad neduoda konkrečios rekomendacijos, kaip technologiškai įgyvendinti skaitmeninius objektus ar juos sudarančius elementus – tai paliekama už jo apimties ribų.

„Next Generation EIF Blueprint“ projektas, priešingai, kuriamas kaip visam ES viešajam sektoriui skirtas veikimo karkasas, kurio naujumas (jei jis taptų patvirtintu reguliavimo) būtų privalomumo ir valdysenos (angl. *governance*) sugriežtinimas. Projektas save aiškiai įtvirtina IEA logikoje. Akcentuoja, kad EIF turi būti laikomas atskaitos tašku ir turi sąveikauti su akte numatytais praktikomis (įskaitant sąveikumo vertinimą (angl. *interoperability assessments*) prieš diegiant ar keičiant skaitmenines sistemas).

Projektas akivaizdžiai dėmesio centrą perkelia nuo aukšto lygio principų prie vykdytinų rekomendacijų. Daug rekomendacijų projekte turi „Nature: Mandatory“, jos pririšamos prie konkrečių suinteresuotųjų šalių vaidmenų, sąveikumo sluoksnių, principų, teisėkūros atramų ir net Sąveikios Europos sprendinių. Tai esminis poslinkis nuo EOSC-IF, kuriame labiau fiksuojami poreikiai ir kryptys, bet mažiau standartizuojamas atskaitomybės ir atitikties mechanizmas.

Projektas atneša naują politinių / technologinių realiųjų paketą, kurio EOSC-IF (European Commission, 2021a) natūraliai neintegruo kaip „privalomo karkaso“. Aiškiai įtraukiami DVA, DI aktas „EU Digital Identity Wallet / eIDAS“ evoliucija, „EU Data Spaces“, duomenų suverenitetas ir panašūs horizontalūs režimai. Taip EIF iš principų rinkinio tampa mazgu, jungiančiu viešųjų paslaugų sąveikumą su duomenų erdvėmis, tapatybės sluoksniu ir DI skaidrumo / sąveikumo reikalavimais.

Jei projektas taptų patvirtintu reguliavimo standartu, praktiškai tai reikštų vienodesnę, labiau išmatuojamą ir audituojamą bazę visoms valstybėms narėms: Nacionalinės sąveikumo sistemos (angl. *National Interoperability Framework*, NIF) suderinimą su EIF, privalomus stebėsenos mechanizmus, aiškesnę atsakomybę (CIO/koord.), taip pat sąsajas su finansavimu ir pirkimais. EOSC-IF tokiu atveju išliktų kaip sektorinis (mokslinių duomenų) karkasas, bet veiktų griežtesnio horizontalaus pagrindu principu, pvz., semantinių artefaktų ir minimalių metaduomenų idėjos galėtų lengviau susikalbėti su EIF skatinamomis „Core Vocabularies/DCAT-AP“, atitiktis testavimo ir pernaudojimo katalogų praktikomis.

IŠŠŪKIAI IR REKOMENDACIJOS

Teisinis sąveikumas

Esminis teisinio sąveikumo tikslas – sudaryti sąlygas derinti duomenų rinkinius iš kelių šaltinių, pagrindiniai iššūkiai teisiniam sąveikumui užtikrinti.

Iššūkiai

- Nustatytos skirtingos licencijos ar naudojimo sąlygos neleidžia sujungti ar kitaip derinti atskirų duomenų rinkinių naudojimo.
- Nebuvimas standartizuotų, tiek žmogui, tiek mašinai skaitomų licencijų ir kitų duomenų rinkinio naudojimą nustatančių sąlygų.
- Metaduomenyse dažnai nėra nurodoma standartinių mašina skaitomų licencijų ir kitų naudojimo sąlygų.

Rekomendacijos

- Diegti tiek žmogui, tiek mašinai skaitomas licencijas ir kitas duomenų rinkinių naudojimo sąlygas.
- Sudaryti aiškų licencijų ir kitų naudojimo sąlygų rekomenduojamą sąrašą ir užtikrinti jų suderinamumą su EOSC rekomendacijomis.
- Privalomai reikalauti užpildyti metaduomenis apie licencijas ir kitas duomenų naudojimo sąlygas.
- Ribojant prieigą prie duomenų, metaduomenyse aiškiai nurodyti ribojimo priežastis.

Organizacinis sąveikumas

Iššūkiai

- EOSC yra sparčiai besivystanti ir nuolat besikeičianti organizacija, todėl siekiantys dalyvauti joje neturi aiškių ir stabilių rekomendacijų dėl procesų suderinamumo, atsakomybių paskirstymo ir paslaugų prieinamumo.

Rekomendacijos

- Nustatyti aiškią institucinę struktūrą, numatančią aiškius ryšius tarp institucijų, bendravimo tvarką, funkcijas ir atsakomybes.
- Aiškiai apibrėžti, kas yra atsakingas už bendrų sąveikumo paslaugų (pvz., paslaugų katalogų, registru, bendrų PID paslaugų) teikimą, kūrimą, priežiūrą ir kuravimą.
- Vadovautis rekomendacijomis dėl standartizuotų duomenų formatų, žodynų ir atitinkamų metaduomenų naudojimo, kad užtikrintų duomenų ir paslaugų suderinamumą.

Semantinis sąveikumas

Semantinis sąveikumas yra kritinė grandis, užtikrinanti, kad perduodami duomenys būtų ne tik techniškai pasiekiami, bet ir vienareikšmiškai interpretuojami tiek žmonių, tiek mašinų.

Iššūkiai

- Semantinių artefaktų fragmentacija: skirtingos mokslinių tyrimų bendruomenės naudoja specifines ontologijas, tezaurus ir kontroliuojamus žodynus, kurie dažnai nėra tarpusavyje suderinti arba neturi aiškių mašininio būdu nuskaitymų sąsajų.
- Metaduomenų detalumo trūkumas: dažnai metaduomenys traktuojami kaip monolitinis blokas, neįgyvendinant sluoksnuotos struktūros, kuri leistų referuoti į smulkesnius skaitmeninio objekto vienetus (pvz., kintamuosius), taip apsunkinant tikslią duomenų paiešką ir federaciją.
- Mašininio skaitomumo problematika: nors egzistuoja įvairūs semantiniai modeliai, dalis jų išlieka mažiau formalūs (pvz., tik tekstiniai aprašymai), o tai trukdo automatizuotai loginei analizei ir žinių atradimui.

Rekomendacijos

- FDO modelio diegimas: rekomenduojama skaitmeninius objektus konstruoti remiantis FDO principais, užtikrinant nuorodas per PID į bendrus semantinius artefaktus.
- Ontologijų ir žodynų darninimas: skatinti tyrėjų bendruomenes naudoti pripažintus konceptualius metaduomenų karkasus ir publikuoti semantinius artefaktus RDF/OWL formatais su nuolatiniais URI.
- Sluoksnuotos metaduomenų struktūros taikymas: įgyvendinti dekompozicijos principą, leidžiantį semantines nuorodas perkelti į smulkesnius duomenų vienetus, taip didinant pakartotinio naudojimo tikslumą.

Techninis sąveikumas

Techninis sluoksnis sudaro bazinę prielaidą visai sąveikumo ekosistemai, tačiau jo operacionalizavimas susiduria su iššūkiais.

Iššūkiai

- Infrastruktūrinis heterogeniškumas: skirtingos mokslo disciplinos naudoja nesuderinamus duomenų formatus, perdavimo protokolus ir API, o tai reikalauja didelių rankinių pastangų integruojant sistemas.
- Tapatybės ir prieigos valdymo (AAI) fragmentacija: autentifikavimas ir autorizavimas dažnai vyksta izoliuotai kiekvienos paslaugos lygmeniu, sukuriant perteklinius trikdžius vartotojams ir apsunkinant saugų dalijimąsi jautriais ištekliais.
- Nepakankamas DI parengtumo (angl. *AI-readiness*) lygis: infrastruktūros dažnai ne visiškai užtikrina automatizuotą paslaugų veikimą nuo pradžios iki galo (angl. *end-to-end*) be žmogaus įsikišimo, o tai riboja galimybes mokslo duomenis naudoti kaip strateginį šaltinį DI modeliams mokytis.

Rekomendacijos

- Bendros AAI infrastruktūros integravimas: diegti nuo konkrečios bendruomenės nepriklausomus prisijungimo ir teisių suteikimo mechanizmus, užtikrinant sklandų vartotojų judėjimą tarp EOSC ekosistemos paslaugų.
- Atvirų specifikacijų ir minimalių metaduomenų profilių taikymas: paslaugas kurti remiantis atviromis specifikacijomis ir taikyti minimalų metaduomenų taikymo profilį, kuris užtikrintų duomenų atradimą nepriklausomai nuo jų formato.
- Federuotos paieškos mechanizmų plėtra: užtikrinti infrastruktūrą, kuri leistų tyrimų objektų paiešką tiek stambiu (kolekcijų), tiek smulkiu (konkrečių elementų) masteliu per bendrą PID politiką.

2.3.3. Mokslinių tyrimų duomenų saugyklos – patikimumas ir sąveikumas

Remiantis EK parengtu anotuoto dotacijos susitarimo dokumentu (angl. *Annotated Grant Agreement*, toliau – AGA) (European Commission, 2025a), patikima saugykla (angl. *trusted repository*) apibrėžiama kaip skaitmeninis archyvas, užtikrinantis ilgalaikį tyrimų rezultatų išsaugojimą ir prieigą prie jų. Pagrindinis patikimumo kriterijus – saugyklos gebėjimas garantuoti deponuoto turinio autentiškumą ir vientisumą.

Pagal AGA saugyklos klasifikuojamos į tris pagrindines kategorijas, kurios laikomos patikimomis:

- Sertifikuotos saugyklos. Jos yra oficialiai pripažintos tarptautinių organizacijų ar vyriausybinių įstaigų (pvz., „CoreTrustSeal“, „nestor Seal“ (DIN31644) arba ISO 16363).
- Disciplininės (srities) saugyklos. Jos yra plačiai pripažintos ir palaikomos konkrečių mokslo bendruomenių bei turi tarptautinį pripažinimą.
- Institucinės ar bendrosios saugyklos. Jos privalo atitikti specifines kokybės charakteristikas, net jei neturi formalaus sertifikato.

Esminiai techniniai ir organizaciniai reikalavimai patikimai saugyklai apima:

- PID suteikimą. Turinys privalo būti žymimas unikaliais identifikatoriais (pvz., DOI, „Handle“), užtikrinančiais vienareikšmišką citavimą.
- Standartizuotus metaduomenis. Metaduomenys turi būti išsamūs, nuskaitomi mašininio būdu (angl. *machine-actionable*) ir atitikti mokslo bendruomenės standartus.
- Ekspertinį kuravimą. Saugykla turi vykdyti deponuojamų duomenų rinkinių ir metaduomenų tikslumo kontrolę.
- Ilgalaikį tvarumą. Būtinai viešai prieinamos politikos, apibrėžiančios duomenų saugumą, privatumą ir veiklą užtikrinantį tvarumą.

Pabrėžtina, kad asmeninės svetainės, bendrosios debesijos saugyklos (pvz., „Google Drive“, „Dropbox“) ar akademiniai socialiniai tinklai (pvz., „ResearchGate“) nėra laikomi patikimomis saugyklomis.

Sąveikios saugyklos samprata pagal EOSC-IF modelį

EOSC-IF sąveikumą apibrėžia kaip organizacijų ir sistemų gebėjimą bendrauti siekiant bendrų tikslų, perduodant duomenis ir žinias tarp informacinių ir ryšių technologijų sistemų. Kaip jau minėta (žr. 2.3.2 poskyrį) EOSC-IF modelyje sąveikumas dekonstruojamas į keturis kritinius lygmenis:

- Techninis sąveikumas apima atvirų specifikacijų naudojimą, bendrą saugumo ir privatumo sistemą (įskaitant autentifikavimo ir autorizavimo infrastruktūrą – AAI) bei sąsajas (pvz., REST API), leidžiančias skirtinoms platformoms keistis duomenimis.
- Semantinis sąveikumas užtikrina, kad keičiamų duomenų prasmė būtų vienareikšmiškai suprantama abiejose pusėse. Tam naudojami semantiniai artefaktai (ontologijos, kontroliuojami žodynai, terminologijos), kurie patys privalo atitikti FAIR principus ir būti nuskaitomi mašininio būdu.
- Organizacinis sąveikumas yra susijęs su bendru valdymu, paslaugų lygio susitarimais (SLA) ir ilgalaikio išsaugojimo politikų derinimu tarp skirtingų institucijų.
- Teisinis sąveikumas apima licencijavimo schemų (pvz., „Creative Commons“ (CC BY, CC0)) suderinimą ir asmens duomenų apsaugos reglamento (BDAR) laikymąsi, kad būtų panaikinti teisiniai barjerai dalintis duomenimis.

Sąveiki saugykla pagal EOSC-IF privalo remtis FAIR skaitmeninių objektų koncepcija, kurioje duomenys yra susiejami su jų metaduomenimis ir identifikatoriais per standartizuotus protokolus.

Patikimumo ir sąveikumo sintezė skaitmeninėje ekosistemoje

Apibendrinant AGA ir EOSC-IF nuostatas, darytina išvada, kad patikimumas ir sąveikumas yra neatsiejami elementai, formuojantys saugyklos brandą. Patikimumas (AGA) sukuria organizacinį ir procedūrinį pagrindą, garantuojantį, kad duomenys išliks pasiekiami ir nepakitę dešimtmečius. Sąveikumas (EOSC-IF) suteikia techninius įrankius šiems duomenims efektyviai veikti globalioje tinklinėje struktūroje, nepriklausomai nuo mokslo srities ar geografinės vietos.

Abiejų koncepcijų centre esantys FAIR principai yra tarsi metodologinis tiltas: patikima saugykla užtikrina, kad FAIR principai būtų įgyvendinti praktiškai (naudojant PID ir kokybiškus metaduomenis), o sąveikumo sistema įgalina šių principų veikimą automatizuotose sistemose. Metodologiniu požiūriu Lietuvos AP saugyklų vertinimas turi apimti tiek statinį patikimumą (sertifikavimo statusas, politikos, saugumas), tiek dinaminį sąveikumą (API prieinamumas, semantinis metaduomenų turtingumas, licencijavimo aiškumas). Tik tokia holistinė atitiktis leidžia institucinę saugyklą transformuoti į visavertį nacionalinės ir europinės mokslinių tyrimų infrastruktūros komponentą.

Atlikta Lietuvos MSI apklausa, vertinanti Lietuvos AP saugyklų patikimumą ir sąveikumą, leido nustatyti, ar Lietuvos AP saugyklos užtikrina saugų duomenų deponavimą bei įgalina efektyvų jų radimą bei pakartotinį naudojimą globaliu mastu. Apklausos tikslas nebuvo tik inventorizacija. Apklausa buvo struktūrinis įrankis, leidęs susieti teorinius ES reikalavimus (AGA ir EOSC-IF) su praktine Lietuvos mokslo situacija (apie apklausos rezultatai žr. 4 skyrių).

Be Lietuvos MSI apklausos rezultatų, Lietuvos mokslininkų indėlių į pasaulinę duomenų erdvę ir Lietuvoje generuojamas duomenų apimtis galima vertinti ir GS metodikos 1.3 poskyryje aprašytais būdais – pagal Lietuvoje kuriamų ir palaikomų saugyklų skaičių, pagal Lietuvos mokslininkų naudojamas pasaulines saugyklas ir pagal Lietuvos mokslininkų paskelbtų duomenų įrašų skaičių.

Remiantis „FAIRSharing“ duomenimis galima matyti, kad iš maždaug 4000 pasaulinių saugyklų yra 70 saugyklų, kuriose minimas Lietuvos indėlis: arba palaikant DB, arba deponuojant joje duomenis. „Re3Data“ nurodo penkias duomenų bazines (COD, LiDA, GEOMIND, CLARIN-LT ir MIDAS), kurios šiuo metu plėtojamos Lietuvoje. Taigi Lietuva gali pasiūlyti pasaulyje žinomas, aukštos kokybės duomenų bazines. „FAIRSharing“ ir „Re3Data“ sąrašai, deja, yra neišsamūs. Juose nėra, pavyzdžiui, Lietuvos akademinės elektroninės bibliotekos informacinės sistemos (toliau – eLABa) bibliografinės duomenų bazės. Taigi realiai Lietuvoje turimų duomenų bazių skaičius yra didesnis.

Lietuvos mokslininkai labiau matomi (procentais) tarp sukurtų ir palaikomų duomenų bazių negu tarp atskirų (cituojamų) duomenų rinkinių. Iš maždaug 12 mln. duomenų įrašų, randamų „DataCite“, „CrossRef“ ir DCC, tik apie 500–600 gali būti patikimai identifikuoti kaip sukurti Lietuvos mokslininkų, o tai yra apie 500 kartų blogesnis santykis, palyginti su duomenų bazėmis. Galima daryti išvadą, kad lietuviški duomenų rinkiniai nėra gerai sutvarkyti ir paviešinti pagal FAIR principus – arba jiems trūksta metaduomenų, leidžiančių juos identifikuoti kaip „pagamintus Lietuvoje“, arba jie iš viso nėra įtraukti į tinkamas duomenų bazines.

Lietuva turi gana gerą mokslinių duomenų bazių kūrimo įdirbį, bet Lietuvoje generuojami duomenys vis dar dažnai nėra tinkamai aprašomi, paskelbiami ar atveriami, o tai stabdo Lietuvoje sukurtų duomenų naudojimą tiek pačios Lietuvos mokslininkų, tiek pasauliniu mastu.

Šią situaciją būtų galima keisti, remiantis patyrimu, kaip kurti mokslinių duomenų bazines, kurį jau turi kai kurios Lietuvos tyrėjų grupės. Taip pat sudarant sąlygas išplėsti šias gerąsias praktikas į kitas mokslo sritis, kuriose duomenų tvarkymas dar atsilieka. Būtina sukurti Lietuvos MTD metaduomenų katalogą ir pasirūpinti, kad Lietuvos mokslininkų sukurti duomenų rinkiniai būtų gerai žinomi pasaulyje. Sutvarkyti Lietuvos tyrėjų duomenys ir bendra duomenų erdvė palengvintų Lietuvos duomenų naudojimą, pvz., mašininio mokymo ir DI srityse, bei atvertų galimybes bendradarbiauti ir jungtis į didesnes Europines struktūras, tokias kaip EOSC.

Lietuvos mokslinių tyrimų infrastruktūrų duomenų saugyklos

Tam tikras įžvalgas apie Lietuvos MTD saugyklų situaciją galima gauti iš Lietuvos mokslinių tyrimų infrastruktūrų kelrodžio (Lietuvos mokslo taryba, 2024b) (toliau – Kelrodis), kas ketverius metus rengiamo strateginio dokumento, apžvelgiančio ir numatančio strateginės MTI raidos tendencijas ir kryptis, vystymosi prioritetus bei

galimybes įsitraukti į tarptautines MTI iniciatyvas, tinklus ir organizacijas. Kelrodžio analizė MTD aspektu rodo, kad šešios MTI aiškiai turi savo duomenų saugojimo infrastruktūrą, keturios MTI – aiškiai naudojami kitų paslaugomis, o 14 MTI situacija Kelrodyje neapibrėžta gana aiškiai. Pačiame Kelrodyje gana aiškiai sakoma, kad Lietuvos MTI ekosistemai daugelyje sričių trūksta:

- atvirosios prieigos prie duomenų;
- duomenų atvėrimo ir suderinamumo;
- skaitmeninės infrastruktūros, duomenų kaupimo ir apdorojimo pajėgumų stiprinimo.

Pagal šiuo kriterijus reitinguojant MTI pagal jų tinkamumą integruotis į EOSC drauge su Lietuvos nacionaliniu mazgu būtų galima skirti tokias tris grupes:

- **pirma grupė** – savarankiškas (institucines ar tarpinstitucines) MTD saugyklos turinčios MTI;
- **antra grupė** – duomenis generuojančios ir tarptautiškai integruotos MTI, kurioms dar nepakanka aiškaus MTD saugojimo bei publikavimo sluoksnio aprašymo;
- **trečia grupė** – skaičiavimo ir tinklo infrastruktūros, kurios svarbios ekosistemai, bet nėra duomenų saugyklos siaurąja prasme.

Kelrodžio duomenimis, labiausiai EOSC tinklai tinkamos yra pirmosios grupės kandidatės. Jei žiūrėtume tik į tas MTI, kurioms Kelrodyje aiškiai nurodyta MTD valdymo funkcija, prieiga prie jų bei saugojimas. Stipriausia jų grupė būtų ši:

- LiDA;
- CLARIN-LT;
- LTA TI / LFA IR;
- „BioDatCom Centre“;
- FOREST 4.0 ;
- „CossyBio“;

Bet kuriuo atveju Kelrodžio saugyklos yra svarbus MTD saugojimo ir valdymo išteklius ypač dėl to, kad jomis dalijasi MSI konsorciumai.

2.3.4. Apibendrinimas

Šiuolaikinė MTD valdysena išgyvena kokybinį lūžį: pereinama nuo pasyvaus duomenų saugojimo prie aktyvaus jų panaudojimo, vis labiau suprantant duomenis kaip strateginį išteklių dirbtinio intelekto sistemoms ir duomenimis grindžiamam mokslui. Ši transformacija, grindžiama Europos Sąjungos teisiniu ir strateginiu reguliavimu, reikalauja, kad duomenys būtų ne tik prieinami, bet ir sąveikūs, struktūruoti bei tinkami automatizuotam apdorojimui. Siekiant užtikrinti jų atrandamumą ir pakartotinį panaudojimą EOSC kontekste, būtina kompleksiška taikyti metaduomenų, specialiųjų sričių, failų formatų standartus, kontroliuojamus žodynus ir nuolatinius identifikatorius. Šių elementų visuma sudaro prielaidas techniniam, semantiniam ir organizaciniam duomenų sąveikumui, leidžiančiam sistemoms perduoti informaciją išlaikant vienareikšmę jos prasmę.

Pasaulinėje praktikoje MTD tipai klasifikuojami pagal jų kilmę, rinkimo metodą, apdorojimo etapą ir turinio pobūdį. DI plėtra skatina standartizuoti ne tik tradicinius stebėjimų ar eksperimentinius duomenis, bet ir sintetinius duomenų rinkinius, leidžiančius mokyti algoritmus nepažeidžiant duomenų subjektų privatumo. Šiuolaikinė infrastruktūra turi užtikrinti ne tik duomenų saugojimą, bet ir nuoseklią anotacijų, kilmės bei versijavimo fiksavimą, remiantis pripažintais modeliais, tokiais kaip W3C PROV. Tuo pat metu teisinis sąveikumas tampa būtina federuotos ekosistemos sąlyga, nes tik suderintos prieigos, saugumo ir licencijavimo taisyklės leidžia skirtingoms organizacijoms veikti bendroje duomenų erdvėje.

Lietuvos MSI pasirengimas integruotis į tokią aplinką išlieka netolygus. Gamtos, medicinos ir technologijų moksluose dažniau dominuoja didelės apimties eksperimentiniai ir modeliavimo duomenys, kuriems ypač svarbūs atviri formatai, techninis aprašymas ir versijavimas. Tuo tarpu humanitariniuose ir socialiniuose moksluose vyrauja tekstiniai, audiovizualiniai ir archyviniai išteklių, kuriems kritiškai svarbūs semantiniai aprašai, anotavimas ir licencinis aiškumas. Tai rodo, kad Lietuvos integracija į EOSC negali būti grindžiama vienu universaliu techniniu modeliu. Prioritetu turėtų tapti bendro minimalaus metaduomenų ir sąveikumo sluoksnio stiprinimas, specialiųjų standartų taikymo plėtra ir institucinių gebėjimų stiprinimas. Ilgainiui Lietuvos mokslinis konkurencingumas vis labiau priklausys nuo gebėjimo užtikrinti duomenų sąveikumą ir jų parengtumą mašiniam mokymui.

2.4. Įvairių Europos šalių mokslinių tyrimų duomenų infrastruktūros ir valdysenos modeliai

Kitų Europos valstybių MTD ekosistemų analizė atskleidžia didelę infrastruktūrą, MTD valdymo politikų ir atsakomybių už sritį paskirstymo įvairovę. Šie nacionaliniai MTD ekosistemos ypatumai dažnai ir turi esminę įtaką šalies bendradarbiavimo su EOSC ir įsijungimo į jo plėtrą formato pasirinkimui. Kitų šalių patirtis gali būti naudinga modeliuojant Lietuvos MTD ekosistemos vystymą ir jos integraciją į EOSC, todėl naudinga apžvelgti kelių pasirinktų šalių MTD ekosistemos modelius bei jų patirtį ir strateginius tikslus integracijos į EOSC atžvilgiu. Toliau pateikiama dešimties ES valstybių MTD valdysenos ir infrastruktūros modelių bei jų planų EOSC atžvilgiu apžvalga. Į apžvalgą įtrauktos valstybės, kurios MTD valdymo srityje laikomos Europos lyderėmis, taip pat su Lietuva besiribojančios ar panašumų su ja turinčios šalys.

2.4.1. Airija

Airijos MTD valdymo ekosistemai būdingas nacionalinis strateginis koordinavimas ir mišrus institucinės bei teminės infrastruktūros modelis, kuriame politikos, techninės infrastruktūros ir duomenų kuravimo funkcijos paskirstytos keliems veikėjams. Nacionalinį strateginį koordinavimą vykdo Nacionalinis atvirojo mokslo forumas (angl. *National Open Research Forum*, NORF) – tarpsektorė platforma, jungianti mokslo finansuotojus, aukštojo mokslo institucijas, akademinės bibliotekas ir mokslinių tyrimų infrastruktūrų teikėjus. NORF yra atsakinga už Nacionalinio atvirojo mokslo plano 2022–2030 m. įgyvendinimo stebėseną ir politikos derinimą su EOSC principais (National Open Research Forum, 2022).

Svarbi Airijos atvirojo mokslo infrastruktūros dalis yra Airijos skaitmeninė saugykla (angl. *Digital Repository of Ireland*⁴⁵, toliau – DRI), veikianti kaip patikima ilgalaikio mokslinių tyrimų duomenų saugojimo ir metaduomenų agregavimo platforma. DRI palaiko centralizuotą metaduomenų katalogą, grindžiamą „Dublin Core“ ir „Metadata Object Description Schema“ (MODS) metaduomenų schemomis, ir yra visiškai integruota su OpenAIRE, ORCID ir „DataCite“ infrastruktūromis. DRI teikia ilgalaikio išsaugojimo, metaduomenų ir tyrimų rezultatų sklaidos paslaugas, ypač humanitarinių ir socialinių mokslų bei kultūros paveldo srityse.

Techninę prieigą, tinklų infrastruktūrą ir autentifikavimo sprendimus teikia HEAnet⁴⁶, veikiantis kaip Nacionalinis mokslo ir studijų tinklas (angl. *National Research and Education Network*, NREN). HEAnet užtikrina patikimą prieigą prie DRI ir kitų skaitmeninių mokslinių tyrimų paslaugų, sudarydamas sąlygas federuotai prieigai ir tarptautiniam sistemų sąveikumui.

Airijoje duomenų valdytojų funkcija vystoma decentralizuotu, instituciniu principu, glaudžiai susieta su universitetų bibliotekomis ir mokslinių tyrimų palaikymo padaliniais. Nacionaliniu lygmeniu NORF koordinuoja bendrą strateginę kryptį. Taip pat šalyje veikia nacionalinis duomenų valdytojų tinklas (*Sonraí, Irish Data Stewardship Network*)⁴⁷.

⁴⁵ <https://dri.ie>

⁴⁶ <https://www.heanet.ie>

⁴⁷ <https://datastewards.ie/>

Praktiniu lygmeniu duomenų valdytojų paslaugas teikia instituciniai mokslinių tyrimų duomenų valdymo specialistai ir bibliotekininkai, bendradarbiaujantys su DRI. DRI aktyviai dalyvauja mokymuose, prisideda prie gairių rengimo ir konsultavimo, tačiau duomenų valdytojų vaidmuo Airijoje išlieka integruotas į platesnes tyrimų palaikymo funkcijas, o ne apibrėžtas kaip atskira profesinė kategorija.

Airijos modelis laikytinas funkciškai centralizuotu, tačiau techniškai ir instituciškai lanksčiu. Metaduomenys ir politikos principai telkiami nacionaliniu mastu, o pats duomenų saugojimas gali būti tiek centralizuotas, tiek institucinis.

2.4.2. Austrija

Austrijoje vyrauja koordinuotas decentralizuotas MTD valdymo modelis, kuriame didelis dėmesys skiriamas integracijai į EOSC. Nacionalinį koordinavimą ir metodinę paramą palaiko *FAIR Office Austria*⁴⁸, veikiantis kaip centrinis kompetencijų taškas FAIR duomenų, MTD valdymo politikų ir kompetencijų stiprinimo mokymų srityje.

Techninį pagrindą sudaro ACONET (*Austrian Academic Computer Network*) – nacionalinis mokslinių tyrimų ir švietimo tinklas, teikiantis ryšio, tapatybių ir prieigos valdymo infrastruktūros (angl. *Authentication and Authorization Infrastructure*, AAI) ir kitas elektroninės infrastruktūros paslaugas. Finansavimo lygmeniu Austrijos mokslo fondas (vok. *Fonds zur Förderung der wissenschaftlichen Forschung*, FWF) reikalauja, kad visi finansuojami projektai turėtų DVP, taip sistemingai stiprindamas MTD valdymo praktikas institucijose⁴⁹.

Metaduomenys nacionaliniu mastu jungiami per „EOSC Austria Hub“ – nacionalinį koordinavimo ir prieigos prie EOSC paslaugų tašką – ir „DataHub Austria“, veikiančią kaip nacionalinę MTD ir paslaugų paieškos platformą. Metaduomenų sąveikai užtikrinti taikomos „OpenAIRE“ gairės, naudojami „DataCite DOI“ ir „Research Organization Registry“ (ROR) organizacijų identifikatoriai. Universitetai (pvz., Vienos technikos universitetas – TU Wien) dalyvauja federuotoje AAI sistemoje, leidžiančioje bendrą tyrėjų prisijungimą prie nacionalinių ir tarptautinių skaitmeninių paslaugų⁵⁰.

Duomenų kuravimo kompetencijas Austrijoje stiprina „RDM Exchange Network“ – bendruomeninė iniciatyva, orientuota į duomenų valdytojų ir kitų MTD specialistų tinklaveiką, patirties mainus ir praktinės paramos stiprinimą.

Austrijos modelis laikytinas struktūriškai parengtu nacionaliniam EOSC mazgui įgyvendinti, net jei praktinis įgyvendinimas grindžiamas koordinuota federacija, o ne viena centrine duomenų saugykla.

2.4.3. Čekija

Čekijoje taikomas iš dalies centralizuotas MTD valdymo modelis. AM politiką šalyje koordinuoja nacionalinė technikos biblioteka (*Národní technická knihovna*⁵¹, NTK), tačiau atskiros mokslinius tyrimus finansuojančios arba vykdančios organizacijos gali taikyti savo vidines MTD valdymo politikas. Už MTD skirtą infrastruktūrą atsakingas nacionalinis e. infrastruktūros operatorius CESNET⁵², kuris koordinuoja skaičiavimo, duomenų saugojimo ir kitų duomenų valdymui svarbių paslaugų teikimą. Šis paslaugų tiekėjas, įsteigtas 1996 m., veikia kaip Čekijos mokslų akademija ir visus šalies valstybinius universitetus vienijanti asociacija.

Čekijoje kuriamas nacionalinis EOSC mazgas, grindžiamas federuotu principu ir orientuotas į nacionalinės duomenų infrastruktūros plėtrą. Šiuo metu šaliai EOSC asociacijoje atstovauja CESNET, tačiau nacionaliniam mazgui kurti šalyje 2021 m. pradėtas įgyvendinti projektas EOSC CZ⁵³. Šis projektas finansuojamas Čekijos Respublikos mokslo, jaunimo ir sporto ministerijos bei ES lėšomis, jį koordinuoja Masaryko universitetas. Šiuo metu projekte veikia įvairios atskiroms temoms (pvz., metaduomenys arba kompetencijų kelimas) ar mokslo sritims

⁴⁸ <https://www.go-fair.org/national-offices/go-fair-austria/>

⁴⁹ <https://www.forschungsdaten.at/>

⁵⁰ <https://eosc.eu/tripartite-collaboration/austria>

⁵¹ <https://www.techlib.cz/en/>

⁵² <https://www.cesnet.cz>

⁵³ <https://www.eosc.cz/en>

skirtos darbo grupės, jų interneto svetainėje kaupiama informacija apie duomenų valdymo įrankius ar šalyje ir už jos ribų organizuojamus mokymus. EOSC CZ projektas jau dabar aktyviai bendradarbiauja su šalies MSI AM centrais ir Čekijos universitetų bibliotekų asociacija, kad būtų užtikrintas kuo aktyvesnis šalies akademinės bendruomenės įsitraukimas.

Viena iš EOSC CZ projekto atliekamų funkcijų – duomenų valdytojų telkimas ir jų tinklaveikos užtikrinimas. Projektas iniciavo šalyje dirbančių duomenų valdytojų žemėlapią⁵⁴ parengimą, taip pat administruoja duomenų valdytojams skirtą grupę socialiniame tinkle „Discord“, kurioje, remiantis „OpenAire“ pateikiamais duomenimis⁵⁵, yra per 100 narių. EOSC CZ yra parengę atmintinę pradedantiesiems duomenų valdytojams, o Karolio universitetas Prahoje siūlo šių specialistų rengimo kursus, suteikiančius mikrosertifikatą.

Planuojama, kad centrinis kuriamo nacionalinio EOSC mazgo architektūros komponentas bus nacionalinė duomenų saugyklų platforma „Nrodní repositář dat“ (NRD). Tai bus centrinis metaduomenų katalogas, naudojamas API surenkantis metaduomenis iš įvairių šalyje veikiančių specializuotų ir tarpdisciplininių saugyklų. Ši platforma bus paremta šalyje jau egzistuojančia MTD infrastruktūra e-INFRA⁵⁶. Sujungiant iš įvairių saugyklų surinktus metaduomenis, bus įgyvendintas nacionalinis duomenų paieškos portalas, taip pat planuojama sukurti DVP rengti padedantį įrankį ir teikti kitas paslaugas, tokias kaip centralizuota naudotojų autentifikacija. Pirmas Čekijos EOSC mazgo infrastruktūros kūrimo etapas bus užbaigtas 2026 m. (Antol et al., 2024).

Čekijoje ilgą laiką vyravo decentralizuota MTD valdymo ekosistema, sudaryta iš kelių toje srityje pažengusių MSI, daugybės mažiau pažengusių institucijų, AM politiką koordinuojančios nacionalinės bibliotekos ir nacionalinio e. infrastruktūros operatoriaus. Situacija ėmė keistis šaliai nusprendus kurti nacionalinį EOSC mazgą, kai tikslu įsteigtas projektas ėmėsi lyderystės sujungiant šalies institucijas ir jose dirbančius duomenų valdymo specialistus į bendrą tinklą. Reikia tikėtis, kad projekto įgyvendinami į kompetencijas ir tinklaveiką nukreipti veiksmai padės gerus pagrindus technologiniam EOSC nacionalinio mazgo įgyvendinimui.

2.4.4. Danija

Danijoje MTD valdomi koordinuoto institucijų tinklo principu. Nuo 2021 m. šalis turi bendrą MTD strategiją. Strategijos įgyvendinimo stebėseną vykdo tarpinstitucinė ekspertų grupė, o strategijos įgyvendinimą koordinuoja Danijos skaitmeninės infrastruktūros konsorciumas (angl. *Danish e-infrastructure Consortium*⁵⁷, DeIC). DeIC įkurtas Danijos Karalystės švietimo ir mokslo ministerijos iniciatyva, tačiau yra valdomas tarybos, kurią sudaro įvairių šalies mokslo ir studijų institucijų atstovai. Be nacionalinės duomenų valdymo strategijos įgyvendinimo koordinavimo, šis konsorciumas yra atsakingas už prieigos prie nacionaliniu mastu veikiančios mokslinių tyrimų techninės infrastruktūros, tokios kaip HPC ištekliai ar kvantinių skaičiavimų ištekliai, administravimą. DeIC taip pat koordinuoja tarpinstitucinį keitimąsi patirtimi.

Nemažai Danijos aukštojo mokslo institucijų, pavyzdžiui, Pietų Danijos universitetas (angl. *University of Southern Denmark*⁵⁸), turi gerai išvystytas MTD valdymo paslaugas. Šalyje egzistuoja institucinės (pvz., *DTU Data*⁵⁹) ir tarpinstitucinės (pvz., *Library Open Access Repository*⁶⁰) tarpdisciplininės duomenų saugyklos, taip pat nacionalinė duomenų saugykla *DeiC Dataverse*⁶¹.

Danija buvo subūrusi darbo grupę, kuri parengė studiją, skirtą duomenų valdytojų rengimo nacionaliniu mastu temai (Wildgaard et al., 2020). Taip pat Danija dalyvauja Šiaurės šalių duomenų kuravimo tinklo (angl. *Nordic Data Stewardship Network*⁶²) veiklose. Šio tinklo tikslai – skatinti duomenų valdymo profesionalų tinklo plėtrą ir sukurti erdvę, kurioje šie specialistai galėtų dalintis mokymo ištekliais ir patirtimi.

⁵⁴ <https://experience.arcgis.com/experience/52cf6ad12eb247bf9f43d0cd0a6a238b?org=muni>

⁵⁵ <https://www.openaire.eu/os-czech-republic>

⁵⁶ <https://www.e-infra.cz/en>

⁵⁷ <https://www.deic.dk/en/front>

⁵⁸ <https://www.deic.dk/en/front>

⁵⁹ <https://data.dtu.dk>

⁶⁰ <https://loar.kb.dk/home>

⁶¹ <https://dataverse.deic.dk>

⁶² <https://datice.is/en/nordic-data-stewardship-network>

EOSC asociacijoje Daniją atstovauja DeIC. Taip pat buvo suburtas EOSC Danijos koordinavimo forumas (angl. *EOSC Denmark Coordination Forum*⁶³), kurio tikslas – viešinti Danijos akademinę bendruomenę EOSC siūlomas paslaugas bei koordinuoti šalies institucijų ar infrastruktūrų integraciją į EOSC.

Nors Danija neturi centrinės už MTD valdymo ir valdysenos koordinavimą atsakingos institucijos, jos trūkumą kompensuoja puikiai koordinuotas institucijų, tarp jų ir turinčių nemažą įdirbį duomenų valdymo srityje, tinklas ir bendra nacionalinė strategija tiek paties duomenų valdymo, tiek tam reikalingų specialistų rengimo atžvilgiu.

2.4.5. Estija

Nors pagrindinė mokslinius tyrimus Estijoje finansuojanti organizacija – Estijos mokslo taryba (ETAg⁶⁴) – dar 2016 m. parengė rekomendacijas nacionalinei AM politikai, šalis iki šiol neturi nacionalinio išskirtinai AM politikai ar tuo labiau MTD valdymui skirto dokumento. AM minimas nacionalinėje Mokslinių tyrimų ir eksperimentinės plėtros, inovacijų ir verslumo strategijoje 2021–2035 metams⁶⁵, patvirtintoje 2019 m. (visas tekstas prieinamas tik estų kalba). Joje glaustai deklaruojama, kad Estija turėtų palaikyti AM ir parengti tam reikalingus politinius dokumentus ir infrastruktūrą.

Už technologinių paslaugų šalies aukštojo mokslo ir švietimo sektoriui teikimą atsakinga Estijos mokslinės kompiuterijos infrastruktūra (ETAIS⁶⁶). Ji teikia duomenų saugojimo, debesijos ir HPC paslaugas. ETAIS veikia kaip keturių institucijų konsorciumas, kurio nariai yra Tartu universitetas, Talino technologijos universitetas, Nacionalinis cheminės fizikos ir biofizikos institutas bei Estijos Respublikos švietimo ir mokslinių tyrimų ministerija. ETAIS yra šiaurės šalių jungtinio infrastruktūros tinklo *Nordic e-Infrastructure Collaboration* (NeIC⁶⁷) narys. Konsorciumas atsakingas už EOSC viešinimą Estijoje ir už su EOSC susijusių nacionalinio lygio veiklų koordinavimą.

Estija turi nacionalinę mokslinės informacijos sistemą (ETIS⁶⁸), kurioje kaupiama informacija apie šalyje dirbančius mokslininkus ir tyrėjus. Joje saugoma informacija apie tyrėjų kvalifikaciją, paskelbtas publikacijas, ekspertines veiklas, gautus apdovanojimus, tačiau joje neregistruojami tyrėjų saugyklose publikuoti MTD rinkiniai.

Didžioji dalis šalyje egzistuojančių AM iniciatyvų kyla iš universitetų bibliotekų. Būtent universitetų bibliotekų iniciatyva Estija dalyvavo nemažai tarptautinių projektų: „OpenAire“, „EOSC Nordic“, „Skills4EOSC“⁶⁹, FAIR-IMPACT⁷⁰ ir kt. Tartu universitetas⁷¹ ir Talino technikos universitetas⁷² turi institucines MTD saugyklas. Tartu universiteto valdomą saugyklą „DataDOI“ naudoja ne tik šio universiteto nariai, bet ir dvi kitos Estijos aukštosios mokyklos, kurios joje turi institucines duomenų kolekcijas.

Duomenų valdytojais šalyje centralizuotai nėra ruošiami.

Apibendrinant galima konstatuoti, kad Estija gali pasigirti gerai išvystyta technologine infrastruktūra ir keletu MTD valdymo srityje pažengusių institucijų (Tartu universitetas ir Talino technikos universitetas). Vis dėlto bendros nacionalinės MTD valdysenai skirtos politikos nebuvimas stabdo MTD paslaugų vystymą šalyje ir jos integraciją į EOSC.

⁶³ <https://www.deic.dk/en/data-management/data-management-deic/advisory-boards/eosc-denmark-coordination-forum>

⁶⁴ <https://etag.ee/en/>

⁶⁵ <https://www.hm.ee/en/ministry/ministry/strategic-planning-2021-2035#documents--4>

⁶⁶ <https://etais.ee/en/>

⁶⁷ <https://neic.no>

⁶⁸ <https://www.etis.ee>

⁶⁹ <https://www.skills4eosc.eu>

⁷⁰ <https://fair-impact.eu>

⁷¹ <https://datadoi.ee>

⁷² <https://data.taltech.ee>

2.4.6. Graikija

Graikijoje taikomas centralizuotas-federuotas MTD valdymo ir integracijos į EOSC modelis. Šalyje už AM yra atsakingos trys ministerijos: Plėtros ir investicijų ministerija, Skaitmeninio valdymo ministerija ir Švietimo ir religijos reikalų ministerija.

2019 m. 26 Graikijos aukštojo mokslo ir mokslinių tyrimų institucijos susivienijo į Nacionalinę atvirojo mokslo darbo grupę, pasivadinusią Graikijos AM iniciatyva (angl. *Hellenic Open Science Initiative*, HOSI⁷³), kuri parengė Nacionalinio AM plano pasiūlymą (2020 m.). Šis planas tapo pagrindu Graikijos AM strategijai. Tai įtvirtinta Graikijos skaitmeninės transformacijos strategijoje 2020–2025, kuri nurodo AM kaip vieną iš pagrindinių skaitmeninės transformacijos katalizatorių.

Integracijos į EOSC koordinavimas Graikijoje ir toliau vykdomas per HOSI, kurios veiklą remia Skaitmeninio valdymo ministerija. Į HOSI struktūrą įeina mokslinius tyrimus vykdančios institucijos (pvz., Atėnės mokslinių tyrimų centras ATHENA), EOSC paslaugas teikiančios organizacijos (pvz., GRNET), atsakingos už esminių techninių ir infrastruktūros paslaugų teikimą ir akademinės bibliotekos (per HEAL-Link).

HOSI įkurtos darbo grupės, kurios teikia AM politikos pasiūlymus, mokymus ir edukacinę medžiagą, AM palaikymo programinę įrangą. HOSI nominavo GRNET⁷⁴ kaip įgaliojamą organizaciją (angl. *mandated organisation*) į EOSC. GRNET – Graikijos nacionalinės mokslinių tyrimų ir technologijų infrastruktūra – jungia švietimo ir mokslinių tyrimų e. infrastruktūras ir teikia tinklą, debesų kompiuterijos, HPC, duomenų valdymo paslaugas ir e. infrastruktūrą bei paslaugas aukštojo mokslo ir mokslinių tyrimų institucijoms, visų lygių švietimo įstaigoms ir visoms viešojo sektoriaus agentūroms. Prieigą prie publikacijų Graikijoje užtikrina Nacionalinis dokumentacijos centras (angl. *National Documentation Centre*, EKT⁷⁵), kuris nuo 2012 m. vykdo e. leidybos platformos veiklą ir suteikia AP mokslo e. žurnalams, konferencijų straipsnių rinkiniams, monografijoms, kitiems akademiniais leidiniams.

Oficialiai pateiktos informacijos apie ketinimą įkurti Graikijos nacionalinį EOSC mazgą rasti nepavyko.

Graikija duomenų valdytojų rengimą koordinuoja per Nacionalinę atvirojo mokslo iniciatyvą (HOSI) ir Atėnės mokslinių tyrimų centrą, kurie užtikrina bendrą strategiją ir metodinę pagalbą visoms mokslo institucijoms. Pagrindinis dėmesys skiriamas akademinėms bibliotekininkų perkvalifikavimui per HEAL-Link tinklą ir praktiniam mokymui naudojant išmanųjį „Data Stewardship Wizard“ (DSW) įrankį – skaitmeninę platformą, padedančią tyrėjams ir duomenų specialistams kurti DVP.

Specialistų kompetencijos nuosekliai ugdamos vykdamas mokymus ir tarptautinius projektus, taip užtikrinant sklandžią nacionalinės mokslo sistemos integraciją į EOSC ekosistemą. GRNET yra įkūręs Atvirojo mokslo debesies kompetencijų centrą⁷⁶ įgyvendinant projektą „Sills4EOSC“.

Apibendrinant galima teigti, kad Graikija turi tvirtą teisinį pagrindą skaitmeninės transformacijos strategijoje ir veiksmingą institucinį bendradarbiavimą plėtojant HOSI iniciatyvą. Šalis sėkmingai sujungia išvystytą e. infrastruktūrą (GRNET), mokslo leidybos sistemas (EKT) ir sistemingą duomenų specialistų rengimą, taip užtikrindama techninį ir metodinį suderinamumą su Europos AM ekosistema.

2.4.7. Ispanija

Ispanijos MTD valdymas grindžiamas koordinuotu instituciniu modeliu, kuriame svarbų vaidmenį atlieka nacionalinė atvirosios prieigos turinio ir repozitorijų agregavimo infrastruktūra RECOLECTA⁷⁷, kuri veikia kaip metaduomenų agregavimo centras, o Ispanijos nacionalinė mokslinių tyrimų taryba CSIC⁷⁸ (isp. *Centro Superior de Investigaciones Científicas*) ir Ispanijos mokslo ir technologijų fondas FECYT⁷⁹ (isp. *La Fundación Española para la Ciencia y Tecnología*) vykdo strateginį koordinavimą.

⁷³ <https://www.hellenicopenscience.gr/el/>

⁷⁴ <https://grnet.gr/en/company/>

⁷⁵ <https://www.ekt.gr/en/index>

⁷⁶ https://www.skills4eosc.eu/network/competence-centres/open-science-cloud-competence-center-in-greece?utm_source=chatgpt.com

⁷⁷ <https://www.recolecta.fecyt.es>

⁷⁸ <https://www.csic.es/en/csic>

⁷⁹ <https://www.fecyt.es>

Ispanijos nacionalinė mokslinių tyrimų taryba yra įgaliotoji organizacija EOSC asociacijoje. Ispanija EOSC asociacijoje taip pat turi 10 narių ir 17 stebėtojų teisėmis dalyvaujančių organizacijų, iš kurių 19 yra universitetai ir mokslo institutai, 8 – paslaugų teikėjai.

RECOLECTA yra pagrindinis nacionalinis metaduomenų surinkėjas ir veikia kaip išteklių infrastruktūra, skirta moksliniams tyrimams deponuoti atvirojoje prieigoje. Su RECOLECTA susieta 114 institucinių saugyklų, 191 teminė saugykla ir AP žurnalai. RECOLECTA turi per 33 tūkstančius duomenų rinkinių. Naudoja OAI-PMH (angl. *Open Archives Initiative Protocol for Metadata Harvesting*) protokolą. 2025 m. lapkričio mėnesio duomenimis, RECOLECTA nėra EOSC asociacijos narys.

Ispanija yra patvirtinusi Atvirojo mokslo strategiją⁸⁰ (ENCA 2023–2027), kurios pagrindinės kryptys yra skaitmeninė infrastruktūra, skirta AM, duomenų valdymas pagal FAIR principus, AP prie mokslo publikacijų bei mokslo vertinimas ir skatinimas.

Ispanija yra aktyvi tarptautinių projektų, skirtų duomenų ir metaduomenų integracijai į EOSC, dalyvė. Vienas jų – EOSC-SYNERGY⁸¹. Šis projektas skirtas prisidėti prie EOSC plėtros, diegiant AM paslaugas ir kuriant regionines AM iniciatyvas. Ispanija dalyvauja „OpenAIRE“ tinklo, kuris veikia kaip metaduomenų agregatorius ir teikia paslaugas, siekiant sujungti nacionalines atvirosios prieigos saugyklas su platesne Europos sistema, veikloje. Taip pat dalyvauja projekte, skirtame DI, „EOSC AI4EOSC“; projekte, skirtame infrastruktūrai jūrų ir vidaus vandenų tyrimams, – „EOSC AqualNFRA“; projekte, skirtame COVID duomenų prieinamumui, – „BeYond-COVID“; projekte, skirtame ir FAIR ir atvirų duomenų vandenynų, jūrų, pakrančių ir vidaus vandenų ekosistemai, – „EOSC Blue-Cloud 2026“.

Ispanija EOSC tinklalapyje skelbia, kad kol kas neturi suformuotos vienos nacionalinės EOSC iniciatyvos, tačiau nurodo, kad veikia Ispanijos e. mokslo tinklas (*Red Española de e-Ciencia* – REeC⁸²), kurio tikslai – koordinuoti e. mokslo plėtrą, skatinti MTEPI subjektų tarpusavio bendradarbiavimą, e. mokslinių ir technologinių infrastruktūrų, ypač superkompiuterių tinklo „RedIRIS“, bendradarbiavimą.

Ispanija duomenų valdytojus rengia per Nacionalinę atvirojo mokslo strategiją (ENCA 2023–2027), stiprinama universitetų ir specializuotų kompetencijų centrų (pvz., Madrido Karolio III universiteto) vaidmenį. Mokymų procesas derinamas su tarptautinėmis iniciatyvomis, tokiomis kaip „Skills4EOSC“, užtikrinant specialistų gebėjimus dirbti pagal FAIR principus. Pagrindinis dėmesys skiriamas tyrėjų ir pagalbinio personalo kvalifikacijai kelti organizuojant praktinius mokymus, skirtus duomenų valdymo planavimui, etikai ir atvirajai prieigai.

Taigi Ispanija įgyvendina koordinuotą AM modelį, kuriame strateginis vadovavimas ir RECOLECTA agregavimo infrastruktūra sukuria reikiamas prielaidas nacionalinių išteklių integracijai į europines sistemas pagal FAIR principus. Aktyvus šalies dalyvavimas tarptautiniuose projektuose ir Ispanijos AM strategijos 2023–2027 vykdymas leidžia sėkmingai plėtoti skaitmeninę infrastruktūrą bei rengti kvalifikuotus duomenų specialistus.

2.4.8. Latvija

Latvija 2022 m. patvirtino nacionalinę AM strategiją, paremtą 2020 m. atliktos nacionalinės studijos (Bite et al., 2020) rezultatais. Kartu su strategija patvirtintas ir jos įgyvendinimo planas, kuris pridėtas strategijos dokumento pabaigoje. Strategijoje išskiriamos trys pagrindinės AM vystymo kryptys: atviroji prieiga prie mokslinių publikacijų, FAIR principus atitinkantys MTD ir piliečių mokslas. Į planą įtrauktos duomenims valdyti skirtos veiklos apima reikalavimo visiems iš viešųjų lėšų finansuojamiems projektams pateikti DVP, o duomenis valdyti pagal FAIR principus įvedimą, nacionalinio skaitmeninių paslaugų tiekėjo ir nacionalinės MTD saugyklos sukūrimą ir duomenų valdytojų rengimo koordinavimą nacionaliniu mastu.

Kaip buvo numatyta nacionalinės AM strategijos įgyvendinimo plane, buvo sukurta centralizuotai skaitmeninės paslaugas, tokias kaip virtualūs serveriai, tapatybių valdymas ar didelės apimties duomenų persiuntimas mokslo ir studijų sektoriui teikianti organizacija – Aukštojo mokslo ir mokslo informacinių technologijų bendrų

⁸⁰ <https://www.ciencia.gob.es/en/Estrategias-y-Planes/Estrategias/ENCA.html>

⁸¹ <https://www.eosc-synergy.eu/europe/spain/>

⁸² <https://www.e-ciencia.es>

jų paslaugų centras (VPC⁸³). Šis centras buvo įkurtas bendromis keturių didžiausių Latvijos universitetų (Latvijos universiteto, Rygos technikos universiteto, Latvijos gyvybės mokslų ir technologijų universiteto ir Rygos Stradiņš universiteto) pastangomis. VPC taip pat prižiūri nacionalinę duomenų saugyklą „DataverseLV“, koordinuoja nacionalinio duomenų prižiūrėtojų tinklo kūrimą ir atstovauja šaliai EOSC asociacijoje. VPC bazinės veiklos ir jo narystė EOSC asociacijoje finansuojami iš nacionalinio biudžeto.

Latvija šiuo metu svarsto galimybę integruotis į EOSC, įsteigdamą nacionalinį mazgą. VPC koordinuoja konsultacijas su Latvijos MSI dėl nacionalinio mazgo įgyvendinimo galimybių. Į nacionalinio mazgo kūrimą orientuota ir dabartinė Latvijos MTI vystymo politika. Planuojama, kad reikalingus infrastruktūros pokyčius bus galima finansuoti sujungiant valstybės biudžeto, ES fondų ir privataus sektoriaus lėšas.

Nors nacionalinę AM politiką Latvijoje koordinuoja šalies Švietimo ir mokslo ministerija, atskiri universitetai ir jų vykdomos iniciatyvos atlieka svarbų vaidmenį formuojant AM kultūrą šalyje. Latvijos universiteto biblioteka atstovavo šaliai projekte „OpenAire“, Rygos technikos universitetas dalyvavo projekte „EOSC Nordic“, o Rygos Stradiņš universitetas valdo institucinę MTD saugyklą⁸⁴.

Apibendrinant pažymėtina, kad Latvijos MTD valdymo ekosistema stipriausia politinėje srityje. Šalis turi bendrą nacionalinę AM politiką, kurioje nemažai dėmesio skiriama duomenų valdymui, o duomenų valdymo specialistų rengimas įvardintas kaip strateginis tikslas. Be to, Latvija oficialiai deklaruoja ketinimą kurti nacionalinį EOSC mazgą ir sudarė tam prielaidas įkurdama VPC, šalies organizaciją atstovę EOSC. Deja, nepaisant politinės valios, menkas šalies MSI įsitraukimas lėtina Latvijos nacionalinio mazgo praktinį įgyvendinimą.

2.4.9. Lenkija

Lenkijoje ilgą laiką vyravo decentralizuotas MTD valdymo modelis, grindžiamas stipriomis institucinėmis ir teminėmis infrastruktūromis, kurios aktyviai integruojamos į tarptautines ekosistemas, pvz., „OpenAIRE“. Pastaraisiais metais ši sistema įgijo labiau koordinuotą formą ėmus nacionalines infrastruktūras integruoti į EOSC, 2026 m. buvo įkurtas Lenkijos nacionalinis EOSC mazgas⁸⁵, kuris siekia tapti nacionaliniu MTD valdymo centru.

Pagrindiniai infrastruktūriniai veikėjai yra „PL-Grid“ infrastruktūra, užtikrinanti nacionalinę HPC ir su ja susijusias duomenų paslaugas, bei Tarpdisciplininis matematinio ir kompiuterinio modeliavimo centras Varšuvos universitete⁸⁶ (angl. *Interdisciplinary Centre for Mathematical and Computational Modelling*, ICM), veikiantis kaip nacionalinis kompetencijų ir paslaugų centras MTD skaičiavimų ir skaitmeninių paslaugų srityje. Šios infrastruktūros sudaro techninį duomenų saugojimo, apdorojimo ir prieigos pagrindą.

Mokslinių tyrimų finansavimo lygmeniu Nacionalinis mokslo centras⁸⁷ (angl. *National Science Centre*, NCN) yra įtvirtinęs privalomą DVP reikalavimą visiems finansuojamiems projektams. Tokiu būdu skatinama nuosekli MTD valdymo praktika institucijose, net ir nesant centralizuoto nacionalinio metaduomenų mazgo.

Aukštojo mokslo ir studijų institucijos (ypač universitetai) palaiko nuosavas MTD saugyklas, kurios jungiamos prie „OpenAIRE“ ar kitų tarptautinių metaduomenų agregatorių. Metaduomenų integracija kol kas vykdoma principu „iš apačios į viršų“ – vadovaujantis instituciniais sprendimais ir per infrastruktūrų tinklus, o ne per vieną centrinę nacionalinę platformą, tačiau, atsiradus nacionaliniam EOSC mazgui, situacija ilgainiui gali pasikeisti.

Lenkijos modelis gali būti vertinamas kaip ilgą laiką decentralizuotai vystytas MTD valdymo modelis, kuriame svarbų vaidmenį atliko nacionalinio finansuotojo taikomi duomenų valdymo planavimo reikalavimai ir išplėtos techninės infrastruktūros. Pastaraisiais metais šis modelis įgauna labiau koordinuotą formą, stiprėjant nacionalinei integracijai į EOSC.

Lenkijoje duomenų valdytojų funkcijos daugiausia vystomos instituciniu lygmeniu. Dauguma MTD valdymo kompetencijų telkiamos universitetuose, ypač bibliotekose ir IT padaliniuose, dažnai remiantis NCN nustatytais DVP reikalavimais.

⁸³ <https://vpc.lv>

⁸⁴ <https://dataverse.rsu.lv>

⁸⁵ <https://eosc.pl>

⁸⁶ <https://icm.edu.pl/en/>

⁸⁷ <https://ncn.gov.pl/strona-nieznaleziona>

Metodinę ir techninę paramą MTD srityje teikia infrastruktūros, tokios kaip ICM ir „PL-Grid“, tačiau jų veikla labiau orientuota į skaičiavimo ir techninius aspektus nei į sistemingą duomenų valdytojų kompetencijų ugdymą. Todėl Lenkijos modelį tikslingiau apibūdinti kaip stipriomis institucinėmis ir infrastruktūros kompetencijomis paremtą, šiuo metu vis labiau koordinuojamą MTD valdymo sistemą.

2.4.10. Nyderlandai

Nyderlanduose taikomas institucinis-konsorciumo MTD valdymo modelis, pasižymintis aukštu brandos lygiu ir aiškiu vaidmenų pasidalijimu. Nors Nyderlandai neturi vieno centralizuoto nacionalinio MTD mazgo, nacionalinį koordinavimą užtikrina stiprios konsorciumo organizacijos – SURF ir *Data Archiving and Networked Services*⁸⁸ (DANS).

Nyderlandų mokslinių tyrimų organizacija (angl. *Netherlands Organisation for Scientific Research*⁸⁹, NWO), pagrindinis nacionalinis mokslo finansuotojas, reikalauja, kad finansuojami projektai turėtų DVP, atitinkančius FAIR principus. Tai prisideda prie nuoseklaus MTD valdymo praktikų diegimo institucijose ir stiprina FAIR principais grindžiamą MTD tvarkymą.

DANS veikia kaip pagrindinė nacionalinė duomenų archyvavimo, ilgalaikio išsaugojimo ir kompetencijų stiprinimo organizacija, teikianti repozitorijų ir DOI suteikimo paslaugas. SURF užtikrina tinklų, debesų kompiuterijos, tapatybių ir prieigos valdymo (angl. *Authentication and Authorisation Infrastructure*, AAI) bei kitas technines paslaugas universitetams ir tyrimų institutams. Be to, SURF jau su pirmąja EOSC federacijos plėtros banga jungsis prie EOSC kaip atskiras mazgas.

Nyderlandų infrastruktūra veikia pagal federuoto MTD valdymo (angl. *federated research data management*) principą: universitetai palaiko vietines duomenų saugyklas, o metaduomenys ir paslaugos integruojami per DANS ir SURF⁹⁰ konsorciumus. Tai leidžia išlaikyti institucinį savarankiškumą kartu užtikrinant nacionalinį ir tarptautinį sąveikumą.

Nyderlandai laikomi viena pažangiausių šalių Europoje duomenų kuravimo srityje. Duomenų valdytojų vaidmuo čia yra aiškiai institucionalizuotas, ypač universitetuose, kur duomenų valdytojai veikia kaip atskira profesionalų grupė, glaudžiai bendradarbiaujanti su tyrėjais viso projekto gyvavimo ciklo metu. Nacionaliniu mastu duomenų valdytojų kompetencijos stiprinamos per SURF ir DANS, kurie rengia mokymus, metodines gaires ir palaiko bendruomenes. Taip pat veikia nacionaliniai <https://dans.knaw.nl/en/> duomenų valdytojų tinklai ir kompetencijų centrai, susieti su NWO reikalavimais dėl FAIR principų taikymo ir DVP. Nyderlandų duomenų kuravimo modelis laikytinas brandžiu ir sisteminiu, apimančiu tiek techninius, tiek metodinius, tiek organizacinius aspektus.

2.4.11. Apibendrinimas

Aptartų Europos šalių patirtis rodo, kad MTD valdymo ir integracijos į EOSC sprendimai formuojami labai skirtingais keliais, priklausomai nuo nacionalinės mokslo sistemos brandos, institucinių pajėgumų ir valstybės vaidmens suvokimo. Airijos, Austrijos, Čekijos, Danijos, Estijos, Graikijos, Ispanijos, Latvijos, Lenkijos ir Nyderlandų pavyzdžiai atskleidžia platų spektrą – nuo kryptingai centralizuojamų nacionalinių modelių iki brandžių federuotų ekosistemų, veikiančių be vieno formalaus nacionalinio mazgo.

Centralizuotų arba centralizuotai koordinuojamų modelių pavyzdžiais galima laikyti Čekiją ir Graikiją. Čekijoje integracija į EOSC grindžiama aiškia valstybės investicija ir nacionaliniu projektu EOSC CZ, kurio tikslas – sukurti nacionalinį metaduomenų katalogą ir bendrą paslaugų architektūrą, paremtą jau veikiančia e. infrastruktūra. Graikijoje, nors formaliai nacionalinis EOSC mazgas nėra deklaruotas, praktikoje veikia centralizuota ir federuota sistema, kurioje politikos formavimas, techninė infrastruktūra ir paslaugų teikimas telkiami per nacionalines iniciatyvas ir tinklus. Abiem atvejais matomas stiprus valstybės vaidmuo ir „iš viršaus į apačią“ formuojama architektūra, būdinga šalims, siekiančioms greitai sumažinti fragmentaciją ir sukurti bendrą nacionalinį kontaktinį tašką tarptautinėms iniciatyvoms.

⁸⁸ <https://dans.knaw.nl/en/>

⁸⁹ <https://www.nwo.nl/en/research-data-management>

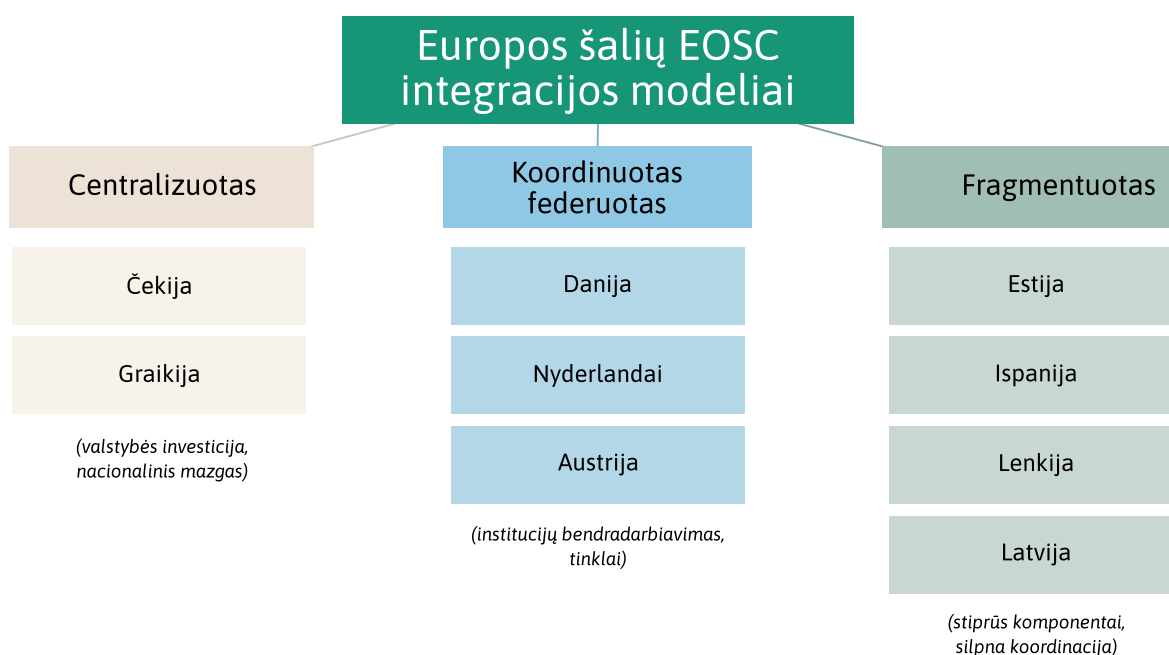
⁹⁰ <https://www.surf.nl/en>

Kita šalių grupė – Danija, Nyderlandai ir Austrija – atstovauja koordinuotam decentralizuotam arba federuotam modeliui. Danijoje egzistuoja nacionalinė duomenų valdymo strategija, tačiau jos įgyvendinimas grindžiamas institucijų bendradarbiavimu, koordinuojamu konsorciumo principu, be siekio kurti vieną centrinę saugyklą ar metaduomenų mazgą. Nyderlandų atvejis išsiskiria aukščiausiu brandos lygiu: čia stiprus nacionalinis finansuotojas, išplėtos nacionalinės paslaugos ir labai aiškiai institucionalizuotas duomenų valdytojų vaidmuo leidžia MTD valdymo sistemai veikti kaip integruotai ekosistemai be centralizuoto valdymo. Austrija užima tarpinę poziciją – čia veikia nacionaliniai koordinavimo mechanizmai, metodiniai centrai ir teminės saugyklos, o EOSC integracija įgyvendinama per koordinuotą federaciją, išlaikant galimybę ateityje prireikus pereiti prie labiau centralizuoto sprendimo.

Trečiajai grupei priskirtinos šalys, kuriose MTD valdymo sistemos išlieka fragmentuotos, nors tam tikri komponentai yra gerai išvystyti. Estijoje stipri techninė infrastruktūra ir aktyvus universitetų vaidmuo neatsveria nacionalinės politikos ir metaduomenų koordinavimo trūkumo. Ispanijoje veikia stiprus metaduomenų agregatorius ir šalis aktyviai dalyvauja tarptautiniuose projektuose, tačiau nacionalinio lygmens koordinacija išlieka fragmentuota. Nepaisant trečiosios grupės šalims būdingos fragmentacijos, joje yra valstybių, inicijavusių nacionalinio EOSC mazgo kūrimą. Lenkija ilgą laiką neturėjo aiškios nacionalinės MTD architektūros vizijos, tačiau nacionalinio EOSC mazgo, kuris pradės veikti jau per pirmąją EOSC federacijos plėtros bangą, sukūrimas turi potencialo situaciją iš esmės pakeisti. Šiuo metu sistemos funkcionavimą užtikrina ryškus finansuotojo spaudimas per privalomus DVP ir išplėtos skaičiavimo infrastruktūros. Latvija, turinti nacionalinę strategiją ir nacionalinę saugyklą, šiuo metu yra pereinamojoje stadijoje ir svarsto nacionalinio EOSC mazgo įgyvendinimą, todėl jos situacija struktūriškai yra artimiausia Lietuvai Baltijos regione.

Bendroji šių šalių analizė leidžia daryti kelias esmines išvagas. Pirmą, formalus nacionalinis EOSC mazgas nėra būtina sąlyga brandžiai MTD valdymo sistemai – Danijos ir Nyderlandų pavyzdžiai rodo, kad aukštas institucinis pajėgumas ir griežti finansuotojų reikalavimai gali užtikrinti efektyvią integraciją ir be centralizuotos architektūros. Antra, mažesnės ir vidutinio dydžio šalys dažniau renkasi koordinuotą centralizaciją, siekdamas sumažinti fragmentaciją ir sukurti aiškų nacionalinį atstovavimą tarptautinėse iniciatyvose. Trečia, visose pažangiausiose sistemose duomenų valdytojų kompetencijos ir duomenų valdytojų tinklai tampa esminiu brandos rodikliu – ten, kur šios kompetencijos stiprinamos sistemiškai, MTD valdymas peržengia formalių reikalavimų lygmenį ir tampa kasdiene tyrimų praktika.

3 pav. Europos šalių EOSC integracijos modeliai



3. Mokslinių tyrimų duomenų atvėrimui, tvarkymui ir priežiūrai reikalingi žmogiškieji ištekliai ir kompetencijos bei jų ugdymas

3.1. EOSC ekosistemos dalyvių vaidmenys ir reikiamos kompetencijos

Siekiant įgyvendinti EOSC viziją, reikalinga nuosekli kompetencijų ugdymo ir mokymų strategija, užtikrinanti FAIR principų ir AM praktikų plėtrą bei inovacijas. Iššūkiu išlieka skaitmeninių kompetencijų specialistų parengimas ir Europos lyderystė AM srityje. 2021 m. parengtoje EOSC kompetencijų ir mokymų darbo grupės ataskaitoje (European Commission, 2021b) įvardijami EOSC ekosistemos dalyviai, atsakingi už mokymus bei pagalbą tyrėjams ir EOSC mazgo atstovams, kurių kompetencijoms ugdyti turi būti skiriamas dėmesys. Tarp jų svarbu paminėti duomenų bibliotekininkus, duomenų tvarkytojus, duomenų valdytojus, EOSC mokymų vykdytojus ir kt.

Atsiradus poreikiui sukurti EOSC kompetencijų ir gebėjimų ugdymo sistemą, 2020 m. buvo įkurta EOSC gebėjimų ir mokymo darbo grupė (angl. *WG Skills and Training*). Darbo grupė daugiausia dėmesio skyrė keturioms prioritetinėms sritims: naujos kartos FAIR ir AM specialistų ugdymui; bendradarbiavimui siekiant stiprinti skaitmeninius įgūdžius FAIR ir AM srityje Europoje; patikimo ir ilgalaikio federacinio žinių centro, kuriame būtų kaupiami mokymosi ir mokymo ištekliai bei susiję įrankiai, kūrimui; kompetencijų ugdymo nacionaliniu lygmeniu skatinimui (European Commission, 2021b).

EOSC gebėjimų ir mokymo darbo grupė parengė EOSC kompetencijų struktūrą, vadovaudamasi įvairiais dokumentais, susijusiais su duomenų tvarkymu ir valdymu reikalingomis kompetencijomis, tokiais kaip Skaitmeninės kompetencijos ir pajėgumų stiprinimas duomenimis grįstame moksle (angl. *Building digital workforce capacity and skills for data-intensive science*) (OECD, 2020), „EOSCpilot FAIR4S“⁹¹ kompetencijų ir gebėjimų struktūra (angl. *EOSCpilot FAIR4S Framework*), Duomenų kompetencijų struktūra (angl. *Data Skills Framework*) (Open Data Institute, 2020), EDISON duomenų mokslo struktūra (angl. *EDISON Data Science Framework*) (Demchenko et al., 2017), „DigComp“ (European Commission Joint Research Centre, 2018). Darbo grupė taip pat išanalizavo ir įvertino kelių konkrečių šalių duomenų valdytojų ir susijusių kompetencijų ugdymo programų reikalavimus (Scholtens et al., 2019; Wildgaard et al., 2020).

EOSC nuolatinės AM ekspertų grupės (angl. *OA Expert Group: Skills, Training, Rewards, Recognition and Upscaling*) veikla taip pat nukreipta į klausimus, susijusius su EOSC aktualių gebėjimų ugdymu, mokymų organizavimu, skatinimu, pripažinimu ir plėtra. Su mokymais susijusios veiklos taip pat vykdomos mazgų koordinavimo komitete. EOSC darbo grupėje diskutuojama dėl EOSC akademijos vaidmens ir indėlio EOSC federacijoje. Tarp prioritetų yra siekis sukurti EOSC akademijos struktūrą ir suformuoti veiklas, įgalinančias dalijimąsi įvairaus pobūdžio ištekliais tarp EOSC mazgų, mokymų integravimas į EOSC federaciją, užtikrinant kompetencijos centrų, kitų mokymų teikėjų ir EOSC mazgų sinergiją, taip pat siekiama pagerinti mokymų medžiagos paiešką. Darbo grupės veiklos taip pat apima EOSC kompetencijų ugdymo ir mokymų gairių kūrimą, ankstesnių rekomendacijų konsolidavimą. Rengiamas naujos kompetencijų ir mokymo darbo grupės veiklos programos projektas.

EOSC ekosistemos dalyviai ir jų kompetencijos

Minėtoje EOSC kompetencijų ir mokymų darbo grupės ataskaitoje EOSC ekosistemos dalyvių modelyje išskiriama dešimt vaidmenų, pasiskirsčiusių tarp keturių skirtingų veiklos sričių: informacinių komunikacinių technologijų (toliau – IKT), bibliotekų ir informacijos mokslo, disciplininių mokslo tyrimų ir plačiosios visuomenės. Vaidmenys sugrupuoti pagal šias veiklos sritis, tačiau ribos tarp plačių kategorijų iš dalies sutampa ir yra lanksčios. EOSC siekiama tarpdiscipliniškumo, todėl dauguma funkcijų yra tarpdisciplininės arba užtikrinamos veikiant bendroms komandoms, atliekančiomis skirtingas funkcijas (2 lentelė). Daugeliui iš minėtų vaidmenų yra būtinos pagrindinės tarpdisciplininės FAIR ir AM kompetencijos, apimančios daugelį mokslinių tyrimų sričių.

⁹¹ <https://eosc-fair4s.github.io/framework>.

2 lentelė. EOSC ekosistemos dalyvių apibūdinimas ir kompetencijos (European Commission, 2021b)

<p>Tyrėjai (angl. <i>researchers</i>)</p>	<p>Apibūdinimas: Tyrėjas yra pagrindinis EOSC ekosistemos naudotojas, siekiant gauti, apdoroti, sukurti, saugoti ir dalytis tyrimų duomenimis.</p> <p>Kompetencijos:</p> <ul style="list-style-type: none"> · Bendros žinios apie EOSC ekosistemą, apimančios „EOSC-Core“ ir „EOSC-Exchange“ tarpdisciplinines paslaugas, skirtas duomenų prieigai, dalijimuisi, pakartotiniam naudojimui ir apdorojimui, bei atitinkamas konkrečioms disciplinoms skirtas paslaugas. · Žinios apie pridėtinę vertę, kurią EOSC paslaugos teikia mokslinių tyrimų ir publikacijų darbo eigai. · Supratimas, kaip įvertinti paslaugų FAIR atitiktį ieškant ir rengiant MTD. · Gebėjimai taikyti EOSC paslaugas visur, kur jos remia mokslinių tyrimų ir inovacijų ciklą ir prisideda prie jų plėtros. · Mokymo ir bendravimo gebėjimai, reikalingi mokyti ir ugdyti kitus tyrėjus ir studentus, kaip atlikti mokslinius tyrimus EOSC teikiamose sąlygose.
<p>EOSC įgalintojai (angl. <i>EOSC enabler</i>)</p>	<p>Apibūdinimas: Tai tam tikros srities mokslininkai, turintys techninių, AM ir FAIR kompetencijų, galintys kurti ir iš dalies įgyvendinti konkrečios srities programas, naudodamiesi EOSC paslaugomis ir duomenimis, ir kurie veikia kaip tiltas tarp mokslo bendruomenės ir techninių paslaugų. EOSC įgalintojas gali būti tyrėjas, turintis techninių įgūdžių, kuris integruoja tam tikros srities duomenų apdoravimo paketą, kad jis veiktų EOSC federacijoje MTI, gaudamas ir saugodamas duomenis, kurie yra randami ir tinkamai suformatuoti, prieinami ir pakartotinai naudojami pasitelkiant „EOSC-Core“ ir „EOSC-Exchange“ paslaugas.</p> <p>Kompetencijos:</p> <ul style="list-style-type: none"> · Įgūdžiai naudoti „EOSC-Core“ ir „EOSC-Exchange“ tarpdisciplinines paslaugas duomenų prieigai, dalijimuisi, pakartotiniam naudojimui ir apdorojimui. · Gilus FAIR principų supratimas. · Įgūdžiai kurti ir koordinuoti naujas aukšto lygio konkrečioms disciplinoms skirtas EOSC paslaugas. · Mokymo ir bendravimo gebėjimai, reikalingi perduoti ir šviesti tyrėjus ir technikus apie AM, FAIR ir EOSC ekosistemos koncepcijas bei paslaugas, taip pat jų pridėtinę vertę mokslinių tyrimų bendruomenėms. · Duomenų įrankių valdymo gebėjimai, reikalingi suprasti paslaugas ir išteklius, reikalingus mokslinių tyrimų grupės ar institucijos veiklai vykdyti AM kontekste.
<p>Duomenų mokslininkai / duomenų analitikai (angl. <i>data scientists / analytics</i>)</p>	<p>Apibūdinimas: Duomenų apdoravimo ekspertai (nebūtinai konkrečios mokslo srities specialistai), gebantys vertinti duomenų kokybę, iš jų išgauti prasmingas išvalgas ir aiškiai jas pateikti. Jie gali kurti universalius sprendimus, pavyzdžiui, mašininio mokymosi algoritmus, veikiančius federuotoje mokslinių infrastruktūrų aplinkoje ir naudojančius iš EOSC paslaugų gaunamus duomenis. Taip pat prisideda prie pažangių duomenų analizės metodų taikymo įvairiose tyrimų srityse. Šis vaidmuo yra esminis užtikrinant kokybišką duomenų naudojimą, FAIR principų įgyvendinimą ir pažangių skaitmeninių sprendimų plėtrą AM ekosistemoje.</p> <p>Kompetencijos: Techniniai gebėjimai naudoti „EOSC-Core“ ir „EOSC-Exchange“ paslaugas duomenų prieigai, dalijimuisi, pakartotiniam naudojimui ir apdorojimui. Gilus FAIR principų supratimas tiek duomenų, tiek programinės įrangos atžvilgiu. Teorinės žinios ir techniniai gebėjimai, reikalingi reikalavimus atitinkančių, registruotų ir per EOSC platformą randamų su duomenimis susijusių paslaugų projektavimui, kūrimui ir diegimui. Mokymo ir bendravimo gebėjimai, reikalingi tyrėjams bei technikams apmokyti ir ugdyti, kaip naudoti įrankius arba perduoti analizės rezultatus.</p>
<p>EOSC edukatoriai (angl. <i>EOSC educators</i>)</p>	<p>Apibūdinimas: Ekspertas, kuris kuria, organizuoja, nustato ugdymo turinį, valdo bei (arba) koordinuoja mokymo veiklas EOSC kontekste, prisideda prie mokymo teikimo. EOSC edukatorius gerai supranta EOSC ekosistemą ir prižiūri mokymo veiklą.</p> <p>Kompetencijos:</p> <ul style="list-style-type: none"> · Gilus EOSC ekosistemos ir principų supratimas. · Gilus FAIR ir AM principų supratimas, siekiant perduoti žinias ir teikti mokymo medžiagą pagal šiuos principus. · Žinios apie „EOSC-Core“ ir „EOSC-Exchange“ paslaugas ir jų taikymo tarpdisciplininiame kontekste privalumus. · Žinios apie mokymosi aplinką, kuri gali būti efektyviai taikoma EOSC kontekste, siekiant tarpdiscipliniškumo. · Geras įvairių suinteresuotųjų šalių poreikių supratimas, siekiant sukurti ir naudoti mokymosi ir mokymo išteklius, tinkamus įvairiems mokymosi būdams. · Bendravimo įgūdžiai.

<p>Duomenų valdytojai / duomenų bibliotekininkai (angl. <i>data stewards / data librarians</i>)</p>	<p>Apibūdinimas: Duomenų valdytojai apibūdinami kaip duomenų rengimo ir tvarkymo ekspertai, įskaitant duomenų atranką, saugojimą, išsaugojimą, kilmės ir kitų metaduomenų rengimą, tvarkymą bei sklaidą. Duomenų bibliotekininkai yra profesionalūs bibliotekos darbuotojai, kurie yra duomenų valdymo ekspertai, naudojantys MTD kaip išteklius arba padedantys tyrėjams, dirbantiems su duomenimis (aprašymas, archyvimas ir sklaida).</p> <p>Kompetencijos: FAIR principų supratimas, siekiant užtikrinti, kad įvairių sričių MTD atitiktų FAIR ir CARE principus. Gebėjimas naudoti „EOSC-Core“ ir „EOSC-Exchange“ paslaugas duomenims skelbti ir saugoti bei palengvinti infrastruktūros ir bibliotekos paslaugų, skirtų duomenų paieškai, tvarkyti, saugoti ir dalintis, plėtojimą pagal šiuos principus. Gebėjimas užtikrinti AM principų laikymąsi „EOSC-Core“ ir „EOSC-Exchange“ paslaugose, susijusiose su duomenimis. Gebėjimas konsultuoti tyrėjus ir studentus duomenų valdymo klausimais pagal FAIR ir CARE principus, įskaitant esamų duomenų rinkinių paiešką ir pakartotinį naudojimą, pasitelkiant EOSC paslaugas ir ekosistemą.</p>
<p>Duomenų kuratoriai (angl. <i>data curator</i>)</p>	<p>Apibūdinimas: Ekspertai yra atsakingi už visų organizacijos duomenų valdymą ir priežiūrą, siekiant užtikrinti atitiktį politikos ir (arba) reguliavimo reikalavimams dėl ilgalaikio saugojimo ir teikti vartotojams kokybiškus duomenis, kurie yra lengvai prieinami suderinant su įvairiais formatais. Duomenų kuratorius galėtų rinkti ir skelbti genominių mėginių, susijusių su gripu, duomenų rinkinį su geografinėmis anotacijomis, naudodamas konkrečiai šričiai skirtus standartinius formatus, užtikrinančius duomenų „FAIRness“.</p> <p>Kompetencijos: Gilus FAIR principų supratimas ir gebėjimas naudotis duomenų skelbimo ir saugojimo paslaugomis. Gebėjimas užtikrinti AM principų laikymąsi „EOSC-Core“ ir „EOSC-Exchange“ paslaugose, susijusiose su duomenimis. Mokymo ir bendravimo įgūdžiai, kad būtų galima apmokyti ir ugdyti kitus dalyvius, kaip teikti ir naudoti kuruojamus duomenis. Žinios ir techniniai įgūdžiai, kad būtų galima kontekstualizuoti mokslinių tyrimų rezultatus, t. y. keistis ir pakartotinai naudoti esamus su moksliniais tyrimais susijusius metaduomenis (apie mokslininkus, projektus, publikacijas, organizacijas, įrangą ir kt.).</p>
<p>Mokslinių tyrimų programų inžinieriai (angl. <i>research software engineer</i>)</p>	<p>Apibūdinimas: IKT ekspertai, kurie projektuoja, diegia, prižiūri ir (arba) integruoja paslaugas ir programinę įrangą EOSC ekosistemoje, kad būtų galima užtikrinti FAIR principus ir AM, garantuojant programinės įrangos kokybę, atkuriamumą ir tvarumą.</p> <p>Kompetencijos: Gebėjimas įgyti žinių, reikalingų „EOSC-Core“ ir „EOSC-Exchange“ paslaugoms diegti, naudoti ir įdiegti. Supratimas apie reikalavimus, keliamus FAIR principams įgyvendinti, duomenų ir programinės įrangos srityje. Mokymo ir bendravimo gebėjimai, reikalingi tyrėjams ir technikams apmokyti ir ugdyti, kaip naudoti mokslinių tyrimų programinę įrangą.</p>
<p>MTD infrastruktūros palaikymo specialistai (angl. <i>data research infrastructure support professional</i>)</p>	<p>Apibūdinimas: IKT ekspertas, kuris valdo ir eksploatuoja tyrimų infrastruktūras bei paslaugas, reikalingas MTD saugoti, ilgalaikiam išsaugojimui ir apdorojimui.</p> <p>Kompetencijos: Žinios apie EOSC ekosistemos principus ir koncepcijas. Techniniai įgūdžiai, reikalingi saugiai diegti ir prižiūrėti „EOSC-Core“ ir „EOSC-Exchange“ paslaugas, susijusias su apdorojimu ir saugojimu, arba duomenų išteklius, pvz., saugyklas ir duomenų bazines. Mokymo ir bendravimo įgūdžiai, reikalingi kitiems dalyviams apmokyti ir ugdyti, kaip naudotis mokslinių tyrimų infrastruktūromis. Įrenginių valdymo įgūdžiai, reikalingi paslaugoms ir ištekliams suprasti, reikalingiems mokslinių tyrimų veiklai palaikyti mokslinių bendruomenių mokslinių tyrimų infrastruktūrose, remiamose EOSC ekosistemos kontekste.</p>
<p>Politikos formuotojai (angl. <i>policy makers</i>)</p>	<p>Apibūdinimas: Politikos formuotojai renka informaciją konsultuodamiesi ir atlikdami tyrimus, ją apibendrina ir iš jos formuoja politiką, mokslo politikos priemones ar strateginius planus, kurie skatina pageidaujamų veiklų vystymą ir gali apimti finansinę paramą moksliniams tyrimams.</p> <p>Kompetencijos: Žinios apie EOSC ekosistemą, ypatingą dėmesį skiriant AM, privatumui ir saugumui bei FAIR principams; gebėjimas suprasti EOSC konkrečių disciplinų ir bendrųjų paslaugų, skirtų EOSC principams, apimčiai ir ribotumams remti. Įrankių valdymo gebėjimai, leidžiantys naudoti paslaugas ir išteklius mokslo bendruomenės veiklai vykdyti EOSC kontekste.</p>

Piliečiai (angl. <i>citizens</i>)	<p>Apibūdinimas: EOSC tikslinė grupė – visuomenės nariai, besidomintys viena ar keliomis mokslo sritimis (įskaitant, bet neapsiribojant atvirojo kodo bendruomene ar mokslinius tyrimus atliekančiomis komercinėmis įmonėmis), kurie nori gauti informacijos ar prisidėti prie piliečių mokslo iniciatyvų ar kitų visuomenės iniciatyvų, arba kurie yra suinteresuoti mokytis ir spręsti konkrečią problemą, nesusijusią su jų profesine veikla. Piliečiai, besidomintys AM, gali sutikti anonimiškai dalytis informacija ir stebėti tam tikrus gyvenimo būdo rodiklius, pvz., sporto veiklą, mainais už grįžtamąją informaciją apie savo sveikatos būklę.</p> <p>Kompetencijos:</p> <ul style="list-style-type: none"> · Pagrindinės žinios apie EOSC principų koncepciją (ypač daug dėmesio skiriant privatumui ir teisiniams klausimams). · Pagrindiniai IKT gebėjimai, reikalingi bendrosioms ir specialiosioms programoms bei priemonėms, skirtoms piliečiams dalyvauti, naudoti.
--	--

EOSC gebėjimų ir mokymų dalyviais ir veikėjais įvardijami institutai (kai kurie jų yra EOSC-A nariai, tačiau daugelis nėra), kompetencijos centrai (įvairių formų ir dydžių), mokslinių tyrimų teminės infrastruktūros, mokymų teikėjai (plačiausia šio termino prasme), EOSC projektai, kuriuose vykdomi mokymai (pvz., „EOSC4Skills“, „EOSC Gravity“); nacionalinės institucijos / projektai, pridedantys prie AM politikos įgyvendinimo, EOSC akademija, EOSC mazgų darbuotojai, atsakingi už mokymus, ir kt.

Ataskaitoje (European Commission, 2021b) aptarti klausimai išlieka aktualūs ir dabartiniame EOSC kūrimo etape. Pateiktose rekomendacijose, padedančiose užtikrinti sklandžius EOSC kūrimo ir plėtros procesus, akcentuojama mokymų ir kompetencijų ugdymo svarba (European Commission, 2021b):

- Sistemingai integruoti FAIR principų ir AM kompetencijas į aukštojo mokslo studijų programas, remiantis suderintu Europos lygmens kompetencijų modeliu ir užtikrinant jų įtraukimą į kvalifikacinius laipsnius bei karjeros vertinimo sistemas.
- Stiprinti koordinuotą mokymų teikimą, plėtojant kompetencijų centrų tinklą ir kuriant sąveikų mokymosi bei mokymų katalogą, integruotą į EOSC sąveikumo sistemą.
- Užtikrinti politikos ir finansavimo suderinamumą, įsteigiant EOSC kompetencijų ir mokymų lyderystės bei ambasadorių programą, kuri skatintų FAIR ir AM kompetencijų įtraukimą į nacionalines ir Europos finansavimo priemones.
- Taikyti EOSC ekosistemos veikėjų sistemą planuojant iniciatyvas, mokymus, skatinimo ir pripažinimo mechanizmus, siekiant ilgalaikio FAIR ir AM integravimo į mokslo praktiką.

Vykdamas „EOSC4Skills“ projektą parengtas „Minimalus būtinų gebėjimų rinkinys“, apibūdinantis AM dalyviams ir duomenų valdytojams reikalingas kompetencijas (Hansen et al., 2025). Duomenų valdysenos ir tvarkysenos specialistai atlieka koordinuojamąjį ir integruojamąjį vaidmenis. Vykdamas koordinuojamąjį vaidmenį duomenų valdytojai teikia pagalbą visose organizacijų mokslinių tyrimų srityse ir padaliniuose, o integruojantis vaidmuo susijęs su mokslinių tyrimų komanda ir mokslinių tyrimų vykdymu (Hansen et al., 2025). Duomenų valdytojai turi bendradarbiauti su visomis suinteresuotosiomis šalimis valdant ir palaikant procesus, skirtus mokslinių tyrimų duomenims rinkti, juos pritaikyti mokslinių tyrimų tikslams, palengvinti jų transformavimą į mokslinių tyrimų rezultatus, kokybės užtikrinimą ir pagrįstų sprendimų priėmimą dėl jų atvėrimo pakartotiniam naudojimui, atsižvelgiant į etinius, teisinius ir socialinius apribojimus.

EOSC ir projektų metu parengtose studijose (European Commission, 2021b; Hansen et al., 2025) akcentuojama, kad sėkmingam EOSC kūrimui esminiai klausimai yra susiję su mokslininkų, mokslinių tyrimų pagalbinių personalo ir EOSC paslaugų teikėjų kompetencijos ugdymu. EOSC kvalifikuotų specialistų tinklas, sudarantis koordinuotą mokymo infrastruktūrą, yra būtinas siekiant pakeisti kultūrą, kad būtų dalijamasi mokslinių tyrimų rezultatais, ir suteikti asmenims bei institucijoms galimybę ugdyti ir išlaikyti EOSC kompetencijas mokslo srityje ir gebėjimus. Nors ataskaitoje pateikiamos rekomendacijos ir nurodomi tolesni veiksmai, kuriais siekiama įveikti esamas spragas ir kliūtis, trukdančias įgyti svarbių kompetencijų, ateityje reikia skirti daug dėmesio tolesnei pažangai šioje srityje.

EOSC dalyvių kompetencijų ugdymo perspektyvos

2026 m. sausio 27–29 d. vykusiose EOSC dirbtuvėse buvo siekiama sukurti bendrą supratimą apie EOSC mokymų kryptis, aptarti su EOSC mokymais susijusius klausimus, liečiančius plėtotinas pagrindines mokymų temas, EOSC akademijos ir kompetencijos centrų vaidmenį ir jų santykį su kitais susijusiais EOSC dalyviais, mokomosios medžiagos paieškos ir katalogo kūrimo aspektus, naujos mokymų darbo grupės įgaliojimus, kiekvienos temos prioritetus ir tolesnius veiksmus, kiekvieno EOSC dalyvio vaidmenis ir įsipareigojimus. Taip pat buvo aptartos 2024/2025 m. EOSC Žiemos mokyklos mokymų (angl. *EOSC Winter School*) rekomendacijos, kuriomis siekiama sustiprinti kompetencijų centrų veiklą ir EOSC veikėjų kompetencijų ugdymą, pripažįstant jų aktualumą ir šiandien (Clare, 2026).

3 lentelė. 2024/2025 m. EOSC Žiemos mokyklos mokymų rekomendacijos

2024	2025
<p>Suderinti mokymosi būdus ir mokymo programas, įskaitant skirtas duomenų valdytojams.</p> <p>Gauti įsipareigojimą patvirtinti minimalius RDA metaduomenų standartus.</p> <p>Dirbti siekiant įtvirtinti suderintą ir sertifikuotą duomenų tvarkytojų mokymo programą.</p> <p>Dalyvauti perduodant „EOSC Future“ mokymo katalogą ES mazgui.</p> <p>Išnagrinėti mokymą kaip paslaugą EOSC naudojant „Galaxy Training Network“.</p> <p>Prisidėti prie „Skills4EOSC“ kompetencijos centro atidarymo renginio.</p> <p>Išnagrinėti galimybę surengti platesnės apimties seminarą apie kompetencijų centrus EOSC simpoziume 2024 m.</p> <p>Didinti informuotumą apie kompetencijų centrus nacionaliniuose trišaliuose renginiuose.</p> <p>Bendradarbiauti su „EOSC Focus“, siekiant remti įtraukties veiklą.</p> <p>Remti EOSC asociaciją ir projektus, siekiant CoARA tikslų.</p> <p>Tęsti diskusijas dėl mokymo sertifikavimo / akreditavimo suderinimo.</p> <p>Sukurti mechanizmą, skirtą strateginei EOSC mokymo elementams prižiūrėti, koordinuoti ir valdyti.</p> <p>Užtikrinti ilgalaikį EOSC mokymo išteklių tvarumą.</p>	<p>Užpildyti komunikacijos spragą tarp EOSC asociacijos / valdybos ir OAE5.</p> <p>Suderinti kompetencijų centrų vaidmenį ir įgyvendinimą, įskaitant jų ryšį su EOSC mazgais.</p> <p>Tęsti bendradarbiavimą su EOSC-A ir „EOSC Focus“ / „Gravity“, siekiant remti vykstančias įtraukties veiklas.</p> <p>Analizuoti EOSC federacijos vadovą ir prisidėti prie jo rengimo.</p> <p>Analizuoti, suderinti ir priimti EOSC atlygio ir pripažinimo mechanizmus.</p> <p>Apibrėžti EOSC mokymo katalogų viziją: kas reikalinga, kam ir kaip tai pasiekti.</p> <p>Sukurti ir suderinti esamus metodus, parengiant mokymosi mechanizmus.</p> <p>Pasinaudoti esamais nacionaliniais kompetencijos centrais ir stiprinti jų dalyvavimą koordinavimo tinkluose ES lygmeniu.</p> <p>Skatinti dalyvavimą duomenų erdvėse, ES misijose ir Europos partnerystėse.</p>

Vienas iš svarbių EOSC dalyvių kompetencijų ugdymo sistemos sudedamųjų dalių yra EOSC akademija. EOSC akademija suprantama kaip bendradarbiavimo iniciatyva, kuria siekiama bendro mokymosi ir žinių dalijimosi tikslų, telkiant EOSC federacijos ekosistemoje sukauptas žinias. Plėtojant EOSC akademijos veiklas siekiama sukurti bendrą supratimą apie EOSC federaciją: žmonių ir techninės infrastruktūros sistemą, padėti esamiems ir kandidatams mazgams ir išteklių teikėjams dalyvauti EOSC kūrimo ir palaikymo veiklose, skatinant EOSC federacijos plėtrą.

Vertinant dabartinę susiformavusią praktiką, EOSC akademijos veikla orientuota į mazgų kompetencijų ugdymą ir mokymo poreikius, o kompetencijų centrų veiklos orientuotos į galutinių vartotojų – tyrėjų – poreikius. Kompetencijos centras EOSC dokumentuose apibūdinamas kaip virtualus centras, skatinantis mokslinių tyrimų kokybę per mokymus ir žinių perdavimą, akcentuojamos bendruomeninės iniciatyvos, kurias remia bendradarbiaujantis žmonių tinklas, teikiantis žinias, geriausią praktiką ir paslaugas, susijusias su AM, ir skatinantis tarpdisciplininį bendradarbiavimą (EOSC Association, 2025). Kompetencijų centrų vaidmuo susijęs su tyrėjų ir vartotojų poreikių identifikavimu, tyrėjų nukreipimu ir susiejimu su kompetencijas turinčiais proceso dalyviais, galinčiais atsakyti į klausimus, kylančius tvarkant ir naudojant MTD. Svarbu užtikrinti komunikaciją tarp bendruomenių ir skirtingų dalyvių (tyrėjų, paramą teikiančių, valdymo). Duomenų valdytojų vaidmuo matomas kaip tarpininkų, suprantančių tyrėjų veiklas ir turinčių techninius duomenų tvarkymo gebėjimus, kurių trūksta tyrėjams. Kaip trūkstamos kompetencijos įvardijami tokie gebėjimai, kaip teikti pagalbą, koordinuoti, informuoti vietos bendruomenę (institucijų, nacionalinę). 2026 m. mokymuose vykusių diskusijų metu akcentuotas poreikis į EOSC federacijos vadovą (EOSC Association, 2025) tarp kitų EOSC mazgo vaidmenų įtraukti gebėjimų ir kompetencijų vadybininkus.

3.2. Duomenų valdytojų bei kitų su mokslo duomenų valdysena ir tvarkysena susijusių specialistų kompetencijų ugdymas

Siekiant įvertinti duomenų valdymo ir tvarkymo specialistams aktualių kompetencijų struktūrą, atsižvelgiama į jų atliekamas funkcijas, įgūdžius ir veiklas. Duomenų valdymo ir tvarkymo specialistai atlieka svarbų vaidmenį užtikrindami MTD vientisumą, kokybę ir naudojimą įvairiuose sektoriuose, įskaitant mokslinius tyrimus, verslą, sveikatos priežiūrą ir kt. sritis. Duomenų tvarkysena (angl. *data stewardship*) apibūdinama kaip atsakomybė už MTD tvarkymo ir dalijimosi jais optimizavimą, siekiant maksimaliai padidinti MTD potencialią vertę mokslo bendruomenei ir visuomenei, kartu gerbiant visų kitų suinteresuotųjų šalių, pavyzdžiui, duomenų subjektų, pažeidžiamų grupių ar bendruomenių, duomenis rengiančių tyrėjų, organizacijų, iš kurių gaunami duomenys, mokslinių tyrimų finansuotojų, mokslo žurnalų ir duomenų saugyklų, teises ir teisėtus interesus (Wendelborn et al., 2023). Duomenų valdytojų atsakomybė – teikti ekspertines konsultacijas MTD gyvavimo ciklo valdymo klausimais, suprasti kiekvieno etapo svarbiausius aspektus ir užtikrinti MTD vientisumą visą procesą. Jie padeda tyrėjams rengti MTD valdymo planus, kurie būtini siekiant užtikrinti atitiktį FAIR principams. Šiuose planuose apibrėžiama, kaip MTD bus tvarkomi visą mokslinių tyrimų projektą ir po jo, skatinant skaidrumą ir atkuriamumą. Duomenų valdytojai taip pat teikia svarbią pagalbą tyrėjams, padėdami jiems susiorientuoti sudėtingoje MTD saugojimo, prieigos ir ilgalaikio išsaugojimo strategijų srityje, siūlydami geriausią praktiką ir turimus išteklius vertingiems MTD apsaugoti.

Atlikus mokslinės literatūros analizę išskirtos pagrindinės duomenų valdytojų kompetencijos (4 lentelė).

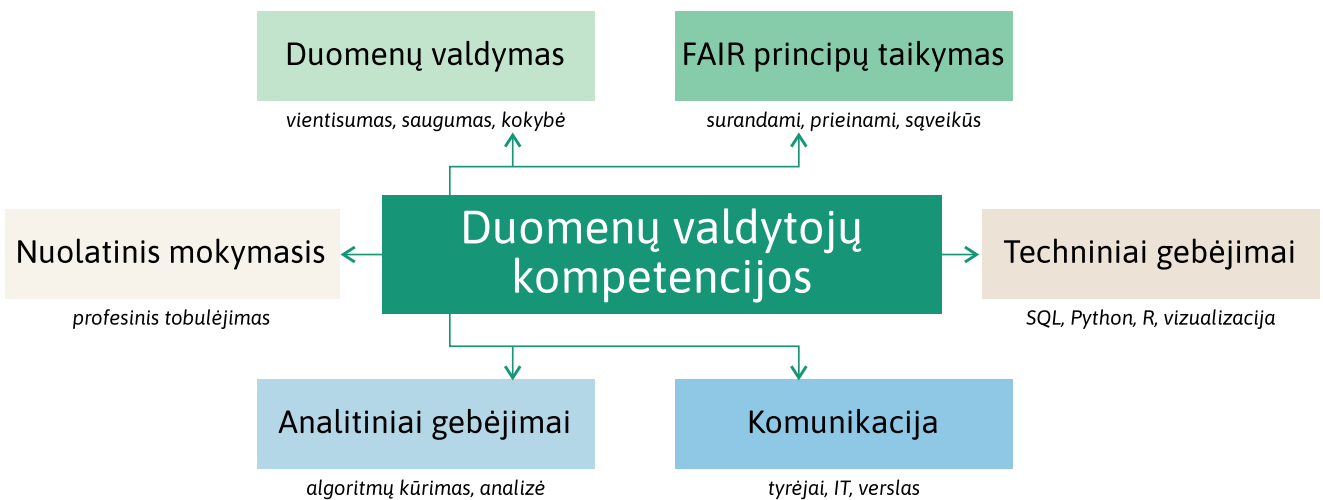
4 lentelė. Pagrindinės duomenų valdytojų kompetencijos

Kompetencijos	Kompetencijų aprašymas	Šaltiniai
Duomenų valdymas	Duomenų valdytojai yra atsakingi už duomenų vientisumą, saugumą, kokybę ir jų tinkamą sutvarkymą laikytis standartų ir taisyklių. Jie atlieka svarbų vaidmenį valdant duomenis, apimant praktikas ir procedūras, skirtas duomenims apsaugoti ir užtikrinti tinkamą jų naudojimą.	Hansen et al., 2025; Kaur et al., 2023; McDowall, 2017; Palmer, 2014; Sylvia & Terhaar, 2023; ES misijose ir Europos partnerystėse.
FAIR principų taikymas	Duomenų valdytojai turi gerai išmanyti FAIR (surandami, prieinami, sąveikūs, pakartotinai naudojami) duomenų valdymo principus. Šie principai būtini veiksmingam mokslinių tyrimų duomenų valdymui.	Aksenova et al., 2024; Demchenko & Stoy, 2021; Pergl et al., 2019; Rinaldi et al., 2025

Techniniai gebėjimai	Gebėjimai apima žinias apie duomenų valdymo įrankius ir technologijas, pvz., SQL, „Python“, R, duomenų vizualizavimo įrankius. Jie taip pat turėtų būti susipažinę su metaduomenų valdymu, duomenų standartais ir duomenų kokybės valdymu.	Demchenko & Stoy, 2021; Hansen et al., 2025; Kaur et al., 2023
Analitiniai gebėjimai	Reikalingi siekiant duomenis transformuoti į prasmingas išvalgas ir praktiškai pritaikomą informaciją. Tai apima algoritmų kūrimą, duomenų analizę ir analitinio turinio patikimumo užtikrinimą.	Demchenko & Stoy, 2021; Palmer, 2014
Komunikaciniai ir bendradarbiavimo gebėjimai	Veiksmingi bendravimo ir bendradarbiavimo gebėjimai yra būtini dirbant su įvairiais suinteresuotaisiais subjektais, įskaitant tyrėjus, IT specialistus ir verslo atstovus, siekiant užtikrinti, kad duomenimis būtų dalijamasi ir jie būtų naudojami veiksmingai.	Demchenko & Stoy, 2021; Kaur et al., 2023; Palmer, 2014
Mokymo ir nuolatinio mokymosi gebėjimai	Nuolatinio profesinio tobulėjimo gebėjimai duomenų valdytojams reikalingi siekiant tobulinti savo kompetencijas ir prisitaikyti prie besikeičiančių duomenų tvarkymo praktikų.	Gruber et al., 2021; Thompson, 2018

Duomenų valdytojai turi turėti įvairių kompetencijų, įskaitant techninius įgūdžius, analitinius gebėjimus, komunikacinius ir bendradarbiavimo gebėjimus. Jie atlieka svarbų vaidmenį valdant duomenis, užtikrina duomenų vientisumą, kokybę ir atitiktį FAIR principams. Nuolatinis mokymas ir tobulėjimas būtini, norint neatsilikti nuo besikeičiančių duomenų valdymo praktikų (žr. 4 pav.).

4 pav. Duomenų valdytojams priskiriamos kompetencijos



Įvairių projektų ir iniciatyvų metu buvo pagrįsta keletas duomenų valdytojų kompetencijų ir struktūros modelių, pavyzdžiui:

- FAIRsFAIR projekto metu sukurta duomenų tvarkymo kompetencijų struktūra, siekiant ją integruoti į universitetų studijų programas. Kompetencijos, įgūdžiai, žinios ir mokymosi rezultatai apibrėžiami naudojant EDISON duomenų mokslo sistemą (Demchenko & Stoy, 2021).
- „FAIR Data Austria“ buvo surengti seminarai, kurių metu buvo identifikuotos ir susietos kompetencijos su iš anksto apibrėžtais duomenų valdytojų modeliais. Šie modeliai padeda pritaikyti mokymo modulius prie konkrečių poreikių ir konteksto (Gruber et al., 2021).

- „Skills4EOSC“ projekto metu buvo siekiama sukurti kompetencijos centrų tinklą, skirtą Europos mokslininkams mokytis ir naujoms pareigybėms, susijusioms su mokslo duomenų valdymu, įsteigti (Corleto et al., 2025; Hansen et al., 2025; Rinaldi et al., 2025).
- Danija ir Nyderlandai yra atlikę duomenų valdytojų rengimo nacionaliniu mastu galimybių studijas (Jetten et al., 2021; Wildgaard et al., 2020).

Siekiant remti duomenų valdytojų mokymą ir kompetencijų ugdymą, parengtos įvairios formalaus ir neformalaus ugdymo programos, užtikrinančios, kad jie būtų gerai pasirengę spręsti sudėtingas duomenų valdymo problemas. Lietuvoje tokių programų šiuo metu nėra siūloma, tačiau tokios programos egzistuoja kitose Europos šalyse.

Didžioji dalis formalių programų vykdomos universitetų kaip dalinės studijos, organizuojamos hibridiniu būdu ir trunkančios du semestrus. Jų turinyje galima identifikuoti nemažai panašumų: visas jas sudaro moduliai, išdėstantys teisinius ir technologinius duomenų valdymo aspektus bei padedantys duomenų valdymui ir tvarkymui pagal AM ir FAIR principus pagrindus. Nemažai programų baigiamos projektais, orientuotais į praktinį mokymų metu įgytų įgūdžių pritaikymą. Vis dėlto esama ir turinio skirtumų, nes vienos programos orientuotos į bibliotekininkus, kitos į mokslininkus, trečios yra atviros įvairioms profesinėms grupėms.

Neformalios programos paprastai būna trumpesnės trukmės nei formalios arba yra prieinamos nuotoliniu mokymo platformose, kurios leidžia mokytis studentui patogiu laiku ir tempu. Atkreiptinas dėmesys, kad neformalaus ugdymo kursai papildo, bet nepakeičia formalaus ugdymo programų. Jie gerai dera kaip praktinis tęsinys po formalių studijų ar kvalifikacijos gilinimo etapas jau dirbantiems duomenų valdytojams.

Toliau apžvelgiami keli formalių ir neformalių užsienyje siūlomų duomenų valdytojų rengimo programų pavyzdžiai.

Formalios programos (programos, kurios vertinamos ECTS kreditais, o jas pabaigus gaunamas diplomą)

Vienos universiteto duomenų valdytojų podiplominės studijos⁹²

Vieną seniausiai egzistuojančių duomenų valdytojų rengimo programų vykdo Vienos universitetas. Ši į studijas baigusius ir mokslinio darbo patirties turinčius asmenis bei universitetų neakademinį darbuotojų orientuota programa sukurta bendradarbiaujant su keliais kitais Austrijos universitetais. Kurso paskaitos vyksta anglų kalba, daugiausia nuotoliniu būdu (kontaktiniu būdu organizuojama tik įvadinė programos savaitė). Programos trukmė – du semestrai. Kontaktinės valandos išdėstytos blokais, po maždaug dvi pilnas darbo dienas kiekvieną mėnesį. Užbaigus kursą ir atlikus baigiamąjį projektinį darbą, išduodamas 15 ECTS kreditų⁹³ vertės duomenų valdytojo sertifikatas. Programą sudaro penki moduliai, kurie apima bendrus duomenų valdymo principus, FAIR principų taikymą, informacinių technologijų ir duomenų mokslo pagrindus ir duomenų valdymo paslaugų organizavimą.

Tamperės universiteto duomenų valdytojų kursai⁹⁴

Tamperės universitetas siūlo į doktorantūros studentus ir jaunuosius mokslininkus orientuotus duomenų valdytojų kursus. Šių anglų kalba vykdomų kursų trukmė – 2 semestrai. Paskaitos, išskyrus du pirmuosius susitikimus, vykstančius Tamperėje, vyksta nuotoliniu būdu. Paskaitos išdėstytos blokais, kiekvieną mėnesį joms skiriamas po kelias dienas arba apie 10 valandų. Programa suskirstyta į penkis modulius, kuriuose supažindinama su AM principais, duomenų gyvavimo ciklu, duomenų valdymu pagal FAIR principus, pagrindiniais kompiuteriniais duomenų apdorojimo ir analizės metodais. Išklausius kursą ir atlikus baigiamąjį projektą išduodamas 15 ECTS kreditų

⁹² <https://postgraduatecenter.at/en/programs/communication-media/data-steward>

⁹³ ECTS (*European Credit Transfer System*) kreditai – Europos Sąjungos sukurta kreditų perkėlimo ir kaupimo sistema, leidžianti suvienodinti studijų krūvio vertinimą ir taip įgalinti vienoje Europos aukštojo mokslo erdvės šalyje baigtų studijų užskaitymą kitose jos šalyse.

⁹⁴ <https://www.tuni.fi/en/study-with-us/tautree-data-steward>

vertės pažymėjimas. Taip pat yra galimybė išklaudyti tik FAIR principams skirtą modulį ir už jį gauti 6 ECTS kreditų vertės pažymėjimą.

Kelno universiteto duomenų bibliotekininkų rengimo programa⁹⁵

Kelno technikos universiteto siūloma duomenų valdymo kompetencijų ugdymo programa, orientuota į mokslinių bibliotekų darbuotojus, kuriose dažniausiai ir dirba su MTD valdymu susijusias paslaugas teikiantys specialistai. Programos trukmė – du semestrai. Bendra kurso apimtis – 240 valandų, išskirstytą į šešis vienos ar dviejų dienų blokus. Programa vykdoma vokiečių kalba, hibridiniu būdu, dalį paskaitų vykdant gyvai, o dalį nuotoliu. Užbaigus kursą ir atlikus baigiamąjį projektą, išduodamas pažymėjimas, kurio vertė – 8 ECTS kreditai. Programa suskirstyta į šešis modulius, kurie orientuoti į praktines duomenų valdymo, analizės ir vizualizacijos žinias. Kadangi programa skirta bibliotekininkams, kurių ne visi gali turėti mokslinio darbo patirties, į kursą įtrauktas ir modulis, skirtas supažindinti su mokslinio darbo specifika.

Graco universiteto duomenų valdytojų kursai⁹⁶

Graco universitetas siūlo duomenų valdytojų rengimo kursą, orientuotą į mokslininkams pagalbą teikiančius specialistus. Kursų trukmė – du semestrai. Kursą sudaro trys trijų dienų trukmės kontaktinio dėstymo blokai ir pora nuotolinių paskaitų kiekvieną savaitę. Paskaitos vyksta vokiečių kalba. Baigus kursus išduodamas 16 ECTS kreditų vertės duomenų valdytojo pažymėjimas. Programos medžiaga suskirstyta į keturis modulius. Ji apima bendruosius duomenų valdymo principus ir bazines duomenų valdytojams reikalingas informacinių technologijų žinias.

Neformalios programos (programos, kurios nevertinamos ECTS kreditais ir nesuteikia diplomo)

OpenAIRE – FAIR mokslinių tyrimų duomenų valdymo intensyvus kursas duomenų valdytojams

Intensyvus 4 dienų FAIR RDM mokymų kursas (*OpenAIRE – FAIR RDM Bootcamp for Data Stewards*)⁹⁷, kurio tikslas – praktinis FAIR principų taikymas ir duomenų valdytojo vaidmens gilinimas. Mokymuose naudojamas „OpenPlato“ platformoje prieinamas nuotolinis mokymo kursas „Mokslinių tyrimų duomenų valdymas duomenų valdytojams“ (angl. *Research Data Management for Data Stewards*)⁹⁸. Mokymo kursas skirtas pradedantiesiems duomenų valdytojams, siekiantiems įgyti praktinių įgūdžių MTD valdymo srityje, taikant FAIR principus, atsižvelgiant į etikos ir teisinius reikalavimus, įgyvendinti AM geriausią praktiką ir įgyti praktinių žinių dirbant su realiais atvejų tyrimais.

EOSC Duomenų valdytojų mokymo kursas⁹⁹

Internetinio EOSC Duomenų valdytojų mokymo kurso (angl. *EOSC Data Steward Training*) tikslas – remti aukštojo mokslo įstaigų ir kitų mokslinius tyrimus atliekančių organizacijų darbuotojų duomenų tvarkymo įgūdžių ugdymą. Kursas apima šias temas: įvadas į FAIR principus, MTD valdymą ir AM; įrankiai FAIR principų žinioms vertinti; atsakingų ir atvirų mokslinių tyrimų supratimas; duomenų valdymo planavimas; duomenų valdymo paslaugų teikimo lygių vertinimas ir planavimas. Tikslinė auditorija: aukštojo mokslo ar mokslinių tyrimų įstaigose duomenų valdymo srityje dirbantys specialistai arba asmenys, turintys mokslinių tyrimų patirties ir svarstantys galimybę pereiti į duomenų valdymo sritį.

⁹⁵ https://www.th-koeln.de/weiterbildung/zertifikatskurs-data-librarian_63393.php

⁹⁶ <https://forschungsdatenmanagement.uni-graz.at/en/data-stewardship-training/> ir <https://www.uniforlife.at/de/data-steward>

⁹⁷ <https://www.openaire.eu/community/blogs/empowering-the-next-generation-of-data-stewards-join-the-fair-rdm-bootcamp>

⁹⁸ <https://openplato.eu/enrol/index.php?id=105>

⁹⁹ <https://learn.eosc-synergy.eu/data-steward-training>

PATTERN projekto mokymai „FAIR mokslinių tyrimų duomenų valdymas: išsamus žvilgsnis į FAIR mokslinių tyrimų duomenų valdymo taikymą praktikoje“¹⁰⁰

PATTERN projekto mokymai „FAIR mokslinių tyrimų duomenų valdymas: išsamus žvilgsnis į FAIR mokslinių tyrimų duomenų valdymo taikymą praktikoje“ (angl. *FAIR Research Data Management: A deeper dive into putting FAIR RDM into practice*) nėra ilgalaikė kvalifikacijos kėlimo programa. Tai intensyvus, praktika grįstas mokymų formatas, skirtas jau turintiems bazinių žinių apie FAIR ir MTD valdymą specialistams, norintiems įgyti konkrečių praktinių žinių: pereiti nuo teorinio FAIR supratimo prie praktinio taikymo, sustiprinti institucines ir projektines duomenų valdytojo veiklos kompetencijas, išmokyti spręsti konkrečius, realius duomenų valdymo atvejus.

Duomenų valdymo ir tvarkymo specialistams skirta metodinė medžiaga

Greta formalių programų, yra nemažai tarptautinių projektų ar konsorciūmų parengtos metodinės medžiagos, kuria galima naudotis savarankiškai tobulinant mokslo duomenų tvarkymo ar valdymo kompetencijas. Du ES programos „Horizontas 2020“ finansuoti FOSTER projektai paliko nemažą įvairius AM aspektus, taip pat ir duomenų valdymą, paliečiančios metodinės medžiagos biblioteką¹⁰¹. Mokslo duomenų valdymo srities metodinės medžiagos sukaupė ir dar vienas „Horizonto 2020“ projektas – FAIRsFAIR¹⁰². Verta atkreipti dėmesį į šiuos kartu su „EOSC Synergy“ projektu 2022 m. „Moodle“ platformoje parengtus nuotolinius duomenų valdytojų mokymus¹⁰³.

MTD valdymu besirūpinanti organizacija *Research Data Alliance* (RDA) taip pat yra sukaupusi nemažai metodinės medžiagos¹⁰⁴. Be to, ši organizacija kasmet pasiūlo įvairios tematikos ir sudėtingumo nuotolinių mokymų mokslo duomenų valdymui ir (ar) tvarkymui aktualiomis temomis¹⁰⁵.

Specialiai socialinių ir humanitarinių mokslų atstovams skirtą duomenų valdymo srities mokymo medžiagą siūlo *Social Sciences and Humanities Open Cloud* (SSHOC) projektas¹⁰⁶, taip pat Europos socialinių mokslų duomenų archyvų konsorciūmas (CESSDA)¹⁰⁷. Gyvybės mokslų srities duomenų valdymo metodinės medžiagos yra sukaupę tarptautinis konsorciūmas ELIXIR¹⁰⁸.

3.3. Apibendrinimas

Įvertinant tarptautinę patirtį, Lietuvai svarbu pereiti nuo projekcinio dalyvavimo prie struktūruotos, ilgalaikės nacionalinės EOSC integracijos politikos, kur kompetencijų ugdymas būtų laikomas strateginiu valstybės prioritetu. Sėkminga integracija į EOSC turėtų būti grindžiama aiškia nacionaline EOSC integracijos ir institucijų duomenų valdymo politika (vaidmenys, atsakomybės, koordinavimas), sistemingu kompetencijų ugdymu (mokymų koordinatoriai, duomenų valdytojai, duomenų tvarkytojai, duomenų vadybininkai), tarptautine integracija (aktyvus dalyvavimas EOSC federacijos veiklose). Siekiant sklandžios šalies integracijos į EOSC būtina įvertinti EOSC federacijos metu kylančius iššūkius. Dėmesys turi būti skiriamas ne tik techninių infrastruktūrų integraciniams ir valdymo sprendimams, integracijai su EOSC ir duomenų valdymui ir tvarkymui svarbių veikėjų ir jiems būdingų vaidmenų įvardijamui, bet ir kompetencijų centro kūrimui ir pozicionavimui nacionaliniu ir instituciniu lygmenimis. Tyrėjų ir kitų su EOSC integracija susijusių dalyvių kompetencijų ugdymas ir ugdymo veiklos koordinavimas turi būti vykdomas sykiu su infrastruktūros ir valdymo sprendimais.

¹⁰⁰ https://www.pattern-openresearch.eu/training_materials/fair-research-data-management-what-is-fair-rdm-and-why-should-we-do-it/

¹⁰¹ <https://www.fosteropenscience.eu/toolkit> ir <https://www.fosteropenscience.eu/courses>

¹⁰² <https://www.fairsfair.eu/competence-centre/training-library>

¹⁰³ <https://doi.org/10.5281/zenodo.6108287>

¹⁰⁴ <https://www.rd-alliance.org/recommendations-and-outputs/catalogue>

¹⁰⁵ <https://www.rd-alliance.org/plenaries-events/webinars/past-webinars>

¹⁰⁶ <https://sshopencloud.eu/training/training-events>

¹⁰⁷ <https://www.cessda.eu/Training-Resources>

¹⁰⁸ <https://elixir-europe.org/about-us/who-we-are>

Nors įvairiuose Europos universitetuose egzistuojančios mokymo programos suteikia galimybę atskiriems specialistams iš Lietuvos įgyti duomenų valdytojų kvalifikaciją, to nepakanka norint nacionaliniu mastu užtikrinti nuoseklų duomenų valdytojų rengimą. Užsienyje organizuojamos programos dažniausiai kainuoja kelis tūkstančius eurų vienam dalyviui, be to, jos neatsižvelgia į Lietuvoje vyraujančią mokslo duomenų ekosistemą, infrastruktūrą ir su mokslo duomenimis susijusį teisinį reguliavimą. Rekomenduojama Lietuvoje įsteigti vietinei situacijai pritaikytą duomenų valdytojų mokymo programą, kurią parengtų ir dėstytų užsienio programose sertifikatus įgiję lektoriai ir (arba) kviestiniai užsienio lektoriai. Rengiant programą gali būti naudingos įžvalgos, sukauptos Danijos ir Nyderlandų parengtose duomenų valdytojų rengimo nacionaliniu mastu temai skirtose studijose (Jetten et al., 2021; Wildgaard et al., 2020). Taip pat galima atsižvelgti į Latvijos, kuri pradėjo duomenų valdytojams rengti skirtą nacionalinio masto projektą¹⁰⁹, patirtimi.

4. Mokslinių tyrimų duomenų ekosistemos Lietuvoje analizė

Rengiant šią GS atlikti empiriniai tyrimai (tyrimų metodika aprašyta GS 1.2 poskyryje) leidžia į šalyje vyraujančią MTD situaciją pažvelgti iš dviejų perspektyvų – kaip ją vertina MSI ir mokslininkai. Šiame skyriuje pirmiausia pateikiamas kiekvienos apklausos rezultatų komentaras, o vėliau aptariamos iš abiejų apklausų bendrosios ir kryžminės analizės.

4.1. Lietuvos mokslo ir studijų institucijų apklausos analizė

Kaip minėta 1.2 poskyryje, pirmiausia, siekiant išsiaiškinti esamą Lietuvos MTD infrastruktūros situaciją bei egzistuojančias Lietuvos MSI duomenų valdymo ir tvarkymo praktikas bei kompetencijas, buvo atliekama žvalgomoji Lietuvos MSI apklausa (apie apklausos vykdymo eigą žr. 1.2.1 poskyrį). Apklausoje sudalyvavo 34 Lietuvos MSI iš 47 (72,3 proc., žr. 5 lentelę).

5 lentelė. Pirmajame tyrime dalyvavusios Lietuvos MSI

Institucija	Dalyvavo	Iš viso šalyje	Proc.
Universitetas	11	17	64,7
Kolegija	11	14	78,6
Institutas	12	16	75,0
Iš viso	34	47	72,3

Visos apklaustos institucijos, išskyrus tris, nurodė, kad jose kuriami ir valdomi (archyvuojami, sisteminami bei išsaugomi pakartotiniam naudojimui) MTD. Dvi institucijos nurodė, kad nevisa apimtimi, tačiau taip pat kuria ir valdo MTD (žr. 6 lentelę).

¹⁰⁹ <https://dataverse.lv/en/par-projektu/>

6 lentelė. Bendroji duomenų valdymo situacija Lietuvos institucijose

Ar jūsų institucijoje kuriami ir valdomi (archyvuojami, sisteminami bei išsaugomi pakartotiniam naudojimui) mokslinių tyrimų duomenys? (N)	
Taip	31
Ne	1
Nevisa apimtimi; tiek, kiek tai reikia publikacijai paskelbti	2

27 institucijos nurodė, kad MTD kuriami ir valdomi individualių mokslininkų ar mokslininkų grupių apimtimi (iš jų 12 – tik tokia apimtimi). 16 institucijų nurodė, kad MTD kuriami ir valdomi visos institucijos ar atskirų jos padalinių apimtimi (iš jų 5 – tik tokia apimtimi). 9 institucijos nurodė, kad MTD kuriami ir valdomi bendrai su kitomis institucijomis, per bendras MTI (iš jų 1 – tik tokia apimtimi). 4 institucijos nurodė, kad duomenys kuriami ir valdomi visokiomis apimtimis (žr. 7 lentelę).

7 lentelė. Duomenų valdymo pobūdis ir apimtys Lietuvos institucijose

K3. Nurodykite, kokia apimtimi kuriami ir valdomi (archyvuojami, sisteminami ir išsaugomi pakartotiniam naudojimui) mokslinių tyrimų duomenys? (N)	
Visos institucijos ar atskirų jos padalinių apimtimi	5
Individualių mokslininkų ar mokslininkų grupių apimtimi	12
Bendrai su kitomis institucijomis, per bendras mokslinių tyrimų infrastruktūras ar kt.	1
Visos institucijos ar atskirų jos padalinių apimtimi ir individualių mokslininkų ar mokslininkų grupių apimtimi	7
Individualių mokslininkų ar mokslininkų grupių apimtimi ir bendrai su kitomis institucijomis, per bendras mokslinių tyrimų infrastruktūras ar kt.	4
Visos institucijos ar atskirų jos padalinių apimtimi ir individualių mokslininkų ar mokslininkų grupių apimtimi, ir bendrai su kitomis institucijomis, per bendras mokslinių tyrimų infrastruktūras ar kt.	4
Nekuria / nevaldo duomenų	1

8 institucijos nurodė, kad nėra reglamentavusios MTD valdymo ir neteikia metodinės pagalbos mokslininkams. Dar dvi nurodė, kad reglamentavimas ir pagalba yra minimalūs arba dar tik rengiamasi tai daryti. 11 institucijų nurodė, kad turi tik vidinius teisės aktus, skirtus MTD valdymo politikai, o dar 7 institucijos nurodė, kad teikia metodinę pagalbą mokslininkams. Tik 6 institucijos nurodė, kad yra ne tik reglamentavusios MTD valdymą, bet ir teikia metodinę pagalbą mokslininkams (žr. 8 lentelę).

8 lentelė. Duomenų valdymo reglamentavimas ir metodinės pagalbos teikimas Lietuvos institucijose.

K4. Ar Jūsų institucija yra reglamentavusi mokslinių tyrimų duomenų valdymą, ar teikiate metodinę pagalbą? (N)	
Taip, turime vidinius teisės aktus, skirtus mokslinių tyrimų duomenų valdymo politikai	11
Taip, teikiame pagalbą mokslininkams, pvz., mokymus apie metaduomenis, duomenų valdymo planus ir kt.	7
Taip, turime vidinius teisės aktus, skirtus mokslinių tyrimų duomenų valdymo politikai ir taip, teikiame pagalbą mokslininkams, pvz., mokymus apie metaduomenis, duomenų valdymo planus ir kt.	6
Planuojame pasitvirtinti vidinius teisės aktus, bet tik mokslinių tyrimų atitikties tyrimų etikai kontekste	2
Ne	8

Tik viena institucija teigė, kad nebūtų suinteresuota prisidėti prie Lietuvos MTD ekosistemos kūrimo ir integracijos į EOSC. 26 institucijos teigė, kad būtų suinteresuotos, o dar 7 – kad būtų suinteresuotos pagal galimybes ir kompetencijas arba tai būtų vienaip ar kitaip aktualu / naudinga (žr. 9 lentelę).

9 lentelė. Duomenų valdymo reglamentavimas ir metodinės pagalbos teikimas Lietuvos institucijose.

K5. Ar jūsų institucija būtų suinteresuota prisidėti prie Lietuvos mokslinių tyrimų duomenų ekosistemos kūrimo ir integracijos į EOSC?	
Taip	26
Pagal galimybes ir kompetencijas; būtų vienaip ar kitaip aktualu / naudinga	7
Ne	1

Antroji išsamioji mokslo ir studijų institucijų apklausa

Antrajame tyrime jau siekta išsamiai išsiaiškinti esamą Lietuvos MTD infrastruktūros situaciją bei egzistuojančias Lietuvos MSI duomenų valdymo ir tvarkymo praktikas bei kompetencijas (apie antrosios apklausos vykdymo eigą žr. 1.2.1 poskyrį).

Apklausoje sudalyvavo 22 Lietuvos MSI iš 33, į kurias kreiptasi (46,8 proc. nuo visų institucijų, į kurias kreiptasi pirmajame tyrime, žr. 10 lentelę). Taigi apklausoje sudalyvavo beveik pusė visų Lietuvos MSI, tarp jų visi didžiausi šalies universitetai.

10 lentelė. Antrajame tyrime dalyvavusios Lietuvos MSI

Institucija	Dalyvavo	Iš viso šalyje	Proc.
Universitetas	9	17	52,9
Kolegija	5	14	35,7
Institutas	8	16	50,0
Iš viso	22	47	46,8

Tyrimo klausimynas suskirstytas į šešias temines sritis, kurios apima technologinę infrastruktūrą, MTD politiką ir reglamentavimą, MTD parametrus, žmogiškuosius išteklius, MTD atvėrimo praktiką bei MSI dalyvavimą tarptautinėse iniciatyvose.

Mokslinių tyrimų duomenų valdymo reglamentavimas

Institucijų pateiktų atsakymų analizė rodo, kad MTD valdymo reglamentavimas Lietuvos MSI sektoriuje jau yra pradėjęs formuotis, tačiau jo institucionalizavimo lygis išlieka nevienodas. Pirminio kodavimo lygmeniu 7 institucijos nurodė turinčios atskirus, aiškiai formalizuotus MTD valdymui skirtus dokumentus, 5 institucijos pažymėjo, kad atitinkamos nuostatos yra integruotos į kitus vidaus teisės aktus, 1 institucija nurodė, kad tokie dokumentai yra rengiami, 2 – kad planuojami parengti, o 6 institucijos pažymėjo, kad tokių dokumentų neturi. Be to, nustatytas 1 mišrus atvejis, kai institucija nurodė, jog MTD valdymo gairės dar rengiamos, tačiau dalis susijusių nuostatų jau yra įtraukta į kitą vidaus dokumentą. Dėl šios priežasties rengiant apibendrintą 1 lentelę detalios kategorijos buvo sujungtos į tris analitines grupes: „turi“, „iš dalies“ ir „neturi“. Pagal šį sujungtą klasifikavimą 7 institucijos priskirtos kategorijai „turi“, 6 – kategorijai „iš dalies“, o 9 – kategorijai „neturi“ (6 priedas 1 lentelė) Analizuojant atsakymus pagal institucijų tipą, matyti, kad MTD valdymo reglamentavimo situacija skirtingose MSI grupėse nėra vienoda. Universitetų ir kolegijų grupėse dažniau nurodoma, kad MTD valdymą reglamentuojantys

dokumentai jau yra parengti arba juos parengti numatyta artimiausiu metu. Mokslinių tyrimų institutų grupėje dažniau fiksuojami atsakymai, kad tokių dokumentų nėra. Šie duomenys rodo, kad MTD valdymo norminis įtvirtinimas Lietuvos MSI sektoriuje pasiskirstęs netolygiai ir skiriasi priklausomai nuo institucinio tipo.

Apklauso rezultatų analizė taip pat rodo, kad institucijose taikomas MTD valdymo reglamentavimas dažniausiai pasižymi ribotu detalumo lygiu. Aiškiausiai normiškai įtvirtinta sritis yra duomenų intelektualinė nuosavybė, tačiau šis klausimas dažniau aptariamas ne MTD valdymui ar AM skituose dokumentuose, bet platesniuose instituciniuose intelektualinės nuosavybės valdymo dokumentuose. Duomenų archyvavimo tvarka yra apibrėžta tik maždaug pusėje apklausoje dalyvavusių institucijų, o aiški duomenų sunaikinimo tvarka nustatyta dar mažesnėje jų dalyje. Duomenų valdymo planų rengimo reikalavimą yra įtvirtinusi mažiau nei pusė institucijų. Be to, tik apie trečdalis institucijų nurodė vykdančios MTD politikos laikymosi stebėseną. Institucijų pateikti atsakymai leidžia matyti, kad šios stebėsenos praktikos nėra vienodos nei savo pobūdžiu, nei formalizavimo lygiu.

MTD valdymo reglamentavimas Lietuvos MSI sektoriuje yra nevienalytis: skiriasi jo formalizavimo lygis, dokumentų turinys ir jų įgyvendinimo stebėseną.

Infrastruktūros pajėgumas ir technologinis brandumas

Duomenų saugojimui ir mainams naudojamos infrastruktūros analizė rodo, kad jos išvystymo lygis ir funkcinių sprendimų įvairovė paprastai siejasi su institucijos dydžiu. Beveik visi universitetai nurodo turintys specialiai duomenų prieglobai ir archyvavimui skirtas saugyklas ar serverius, 5 universitetai nurodo naudojantys „Dspace“ technologija paremtas institucines saugyklas, kurios įgalina ne tik duomenų saugojimą, bet ir jų aprašymą bendrojo pobūdžio metaduomenimis ir atvėrimą. Dar du universitetai nurodo turintys atvirosios prieigos MTD saugykla, skirtas platesnei nei vien institucijos bendruomenė. Institutai ir kolegijos neturi institucinių saugyklų – jos dažniau naudojasi serveriais ar debesijos saugyklomis. Kelios institucijos (1 kolegija ir 1 mokslinių tyrimų institutas) nurodo naudojančios ne centralizuotą institucinę infrastruktūrą, o asmeninius darbuotojų išteklius. Kalbant apie duomenų mainus, institucijos, turinčios dalinimąsi duomenimis įgalinančias saugyklas, naudoja būtent jas. Institucijos, kurios tokių saugyklų neturi, duomenų mainams naudoja interneto svetaines su paieškos galimybe ir debesijos tiekėjų siūlomas bendradarbiavimo funkcijas.

Institucijų pateikti duomenys rodo, kad deklaruojama duomenų saugojimo infrastruktūra tik iš dalies sutampa su faktinėmis MTD saugojimo praktikomis. Dalis universitetų nurodė duomenis saugantys institucinėse saugyklose arba tarptautinėse disciplininėse ir tarpdisciplininėse duomenų saugyklose, tačiau bendra rezultatų struktūra rodo didelę saugojimo vietų įvairovę. Pagal institucijų tipus matyti, kad universitetai ir kolegijos dažniau sieja duomenų saugojimą su instituciniais ar išoriniais sprendimais, o mokslinių tyrimų institutai dažniau nurodo asmeninio lygmens duomenų saugojimo praktikas. Duomenų saugojimo praktikų pasiskirstymas rodo, kad Lietuvos MSI sektoriuje dominuoja decentralizuoti, ir ne visais atvejais tarpusavyje suderinti duomenų saugojimo sprendimai.

Tik nedidelė dalis institucijų šiuo metu turi integracijai į EOSC būtiną infrastruktūrą – duomenų saugykla, leidžiančias išsaugoti ne tik duomenis, bet ir jų metaduomenų aprašus.

Iki šiol sukaupti duomenys ir jų atvirumas

Institucijų apklauso duomenys rodo, kad didžiausi sukauptų MTD kiekiai fiksuojami mokslinių tyrimų institutuose. Šios institucijos dažniau nei kiti institucijų tipai nurodė didžiausią sukauptų duomenų rinkinių skaičių ir dideles bendras duomenų apimtis. Gamtos mokslų srityje institutai nurodė kuriantys daugiau kaip 10 000 duomenų rinkinių per metus, o bendras sukauptų duomenų kiekis viršija 10 TB (6 priedas, 9 lentelė). Didelės duomenų apimtys taip pat fiksuojamos technologijų, medicinos ir žemės ūkio mokslų srityse. Per metus sukuriama duomenų rinkinių skaičius skirtingose srityse svyruoja nuo 100 iki daugiau kaip 10 000 (6 priedas, 10 lentelė), todėl matyti, kad duomenų kūrimo intensyvumas ir galimos jų apimtys skirtingose mokslo srityse nėra vienodi. Be to, mokslinių tyrimų institutai dažniau nei kiti institucijų tipai nurodė kuriantys tarpdisciplininius duomenis.

Universitetų grupėje labai didelės duomenų apimtys, viršijančios 10 TB, ir didelis duomenų rinkinių skaičius, viršijantis 10 000 vienetų (6 priedas, 10 lentelė), daugiausia fiksuojami gamtos ir technologijų mokslų srityse.

Kitose srityse universitetai dažniau nurodo vidutinius arba mažus sukauptų duomenų kiekius, o bendros apimtys dažniausiai neviršija 100 GB. Apklauso duomenys taip pat rodo, kad universitetuose tarpdisciplininių duomenų kūrimas nurodomas rečiau arba apie jį nepateikiama tiksli informacija. Kolegijos tarpdisciplininius duomenis mini kiek dažniau nei universitetai, tačiau jų nurodomi mastai išlieka mažesni nei mokslinių tyrimų institutuose. Kolegijų grupėje visose mokslo srityse vyrauja nedideli sukauptų duomenų kiekiai – iki 100 duomenų rinkinių, o jų bendros apimtys dažniausiai neviršija 100 GB (6 priedas, 10 lentelė). Svarbu ir tai, kad daugelyje mokslo sričių reikšminga dalis institucijų apskritai nurodo tokių duomenų neturinčios. Tai ypač ryšku scenos ir ekrano menų, vaizduojamųjų menų, o kai kuriais atvejais ir humanitarinių bei socialinių mokslų srityse.

Institucijų apklauso duomenys rodo, kad MTD surandamumas pagal metaduomenis skirtingose mokslo srityse yra nevienodas, tačiau daugeliu atvejų išlieka ribotas. Tarp institucijų, nurodžiusių turinčias atitinkamos srities duomenų, medicinos ir sveikatos moksluose 50,0 proc. pažymėjo, kad didžioji dalis arba visi duomenys randami pagal metaduomenis. Socialiniuose moksluose toks rodiklis siekia 46,2 proc., gamtos moksluose – 30,8 proc., technologijų moksluose – 27,3 proc., o humanitariniuose moksluose – 25,0 proc. Žemės ūkio moksluose, scenos ir ekrano menuose bei vaizduojamuosiuose menuose nebuvo nė vieno atvejo, kad didžioji dalis arba visi duomenys būtų randami pagal metaduomenis. Kartu matyti, kad daugelyje sričių vis dar vyrauja ribotas surandamumas: gamtos moksluose 61,5 proc. institucijų nurodė, kad surandama tik maža dalis arba nė vieni duomenys, technologijų moksluose – 72,7 proc., humanitariniuose moksluose – 66,7 proc., o medicinos ir sveikatos moksluose bei socialiniuose moksluose – po 50,0 ir 46,2 proc. atitinkamai.

MTD metaduomenų atviro prieinamumo rodikliai taip pat rodo nevienodą situaciją tarp mokslo sričių. Tarp institucijų, nurodžiusių turinčias atitinkamos srities duomenų, socialiniuose moksluose 38,5 proc. pažymėjo, kad didžioji dalis arba visi metaduomenys yra atvirai prieinami. Medicinos ir sveikatos bei humanitariniuose moksluose toks rodiklis siekia po 33,3 proc., technologijų moksluose – 27,3 proc., o gamtos moksluose – 23,1 proc. Žemės ūkio moksluose, scenos ir ekrano menuose bei vaizduojamuosiuose menuose nebuvo nė vieno atvejo, kad didžioji dalis arba visi metaduomenys būtų atvirai prieinami. Daugelyje sričių dažniau nurodoma, kad atvirai prieinama tik maža dalis metaduomenų arba neprieinamas nė vienas: gamtos moksluose taip atsakė 76,9 proc. institucijų, technologijų moksluose – 72,7 proc., medicinos ir sveikatos bei humanitariniuose moksluose – po 66,7 proc., socialiniuose moksluose – 53,8 proc., o žemės ūkio moksluose ir menų srityse – 100 proc. institucijų, turinčių atitinkamų sričių duomenų.

Žmogiškieji ištekliai, pagalba ir paskatos

MSI duomenų valdymui skiriami žmogiškieji ištekliai. Institucijų apklausa parodė, kad jų pajėgumai šioje srityje labai skiriasi. Didieji universitetai turi dedikuotą personalą. Institucijų atsakymai rodo netolygų žmogiškųjų išteklių pasiskirstymą MTD valdymo ir susijusių funkcijų srityje. Viena institucija aiškiai išsiskiria turėdama apie 7,8 etato, paskirstytą tarp kelių padalinių. Dar dvi institucijos nurodė šioms funkcijoms skiriančios apie 3 etatus, o pavienėse institucijose identifikuojama apie 2 etatus arba apytikriai 1–2 etatus. Vis dėlto daugumoje kitų institucijų šioms veikloms skiriamos tik labai nedidelės etatų dalys – nuo 0,01 iki 0,3 etato – arba jos vykdomos greta kitų bibliotekos ar IT darbuotojų funkcijų. Be to, dalis institucijų nurodė apskritai neturinčios specialiai šioms funkcijoms paskirto personalo. Tai leidžia teigti, kad daugelyje institucijų MTD valdymo veiklos tebėra fragmentiškai integruotos į esamas pareigybės, o dedikuoti pajėgumai išlieka labiau išimtimi nei įprasta praktika.

11 lentelė. MSI duomenų valdymui skiriami žmogiškieji ištekliai

Etatų apimtis	Institucijų skaičius
Apie 7–8 etatus	1 institucija
Apie 3 etatus	2 institucijos
Apie 2 etatus	1 institucija
Apie 1–2 etatus	1 institucija
Nuo 0,01 iki 0,3 etato	Dauguma kitų institucijų
Specialiai nepaskirta personalo	4 institucijos

Mokslo ir studijų institucijose siūlomi mokymai ir informaciniai renginiai duomenų valdymo tema yra nevienodai intensyvūs. Aktyviausios institucijos reguliariai organizuoja mokymus apie duomenų valdymo planų rengimą, FAIR principų taikymą, duomenų saugyklų naudojimą, citavimo įrankius ir tyrimų etiką. Dalis institucijų duomenų valdymo mokymų pačios neorganizuoja, tačiau skatina tyrėjus dalyvauti kitų institucijų ar išorinių teikėjų siūlomuose mokymuose. Didžioji dalis tyrimų institutų nurodo šiuo metu tokių mokymų nesilūtantys.

Keliose institucijose duomenų valdymo temos yra integruojamos į doktorantūros studijas, o kai kur doktorantams numatyta pareiga ankstyvame studijų etape parengti savo tyrimo ar doktorantūros projekto duomenų valdymo planą. Pavienėse institucijose MTD temos įtraukiamos ir į bakalauro bei magistrantūros studijų programas, dažniausiai su atskirų studijų modulių ar tyrimų metodologijos dalykų turiniu.

Be reguliarių mokymų ar į studijų programas integruotų veiklų, institucijos taip pat organizuoja kompetencijoms stiprinti skirtus renginius, tokius kaip teminės duomenų dienos, atvirosios prieigos savaitės renginiai, seminarai ir periodiniai susitikimai. Informacija apie šias galimybes dažniausiai skleidžiama vidiniais naujienlaiškiais arba viešinant kitų institucijų organizuojamus renginius.

12 lentelė. Duomenų valdymo kompetencijų stiprinimo praktikos MSI

Mokymų/ veiklų pobūdis	Institucijų grupė
Reguliariai organizuoja duomenų valdymo mokymus ir informacinius renginius	4 institucijos
Pačios mokymų neorganizuoja, bet skatina dalyvauti kitų institucijų ar išoriniuose mokymuose	2 institucijos
Šiuo metu duomenų valdymo mokymų nesilūdo	4 institucijos
Duomenų valdymo temas integruoja į doktorantūros studijas	5 institucijos
Duomenų valdymo temas integruoja į pirmosios ir antrosios pakopos studijas	2 institucijos
Kompetencijoms stiprinti organizuoja papildomus renginius, seminarus ar temines iniciatyvas	3 institucijos
Informaciją apie mokymus ir renginius platina vidiniais komunikacijos kanalais	4 institucijos

Konsultacinė pagalba duomenų valdymo klausimais mokslo ir studijų institucijose dažniausiai teikiama bibliotekose, IT padaliniuose ir etikos priežiūros funkcijas vykdančiuose padaliniuose. Bibliotekos paprastai konsultuoja dėl duomenų valdymo planų, metaduomenų ir atvirosios prieigos klausimų, IT padaliniai – dėl techninių sprendimų, o su tyrimų etika susiję klausimai sprendžiami pasitelkiant atitinkamus institucinius mechanizmus. Daugumoje institucijų konsultacijos teikiamos pagal poreikį, o tik pavienėse jos vykdomos nuolat ir derinamos su reguliariais mokymais.

Vertinant praktinę pagalbą, dalis institucijų nurodo, kad ji šiuo metu neteikiama arba yra labai ribota. Tais atvejais, kai praktinės paslaugos teikiamos, jos dažniausiai apima duomenų saugojimo sprendimus ir techninę pagalbą. Pavienėse institucijose taip pat teikiama platesnė pagalba, susijusi su metaduomenų rengimu, prieigos organizavimu ir duomenų valdymo procesu stebėseną.

13 lentelė. Apibendrinta konsultacinės ir praktinės pagalbos duomenų valdymo klausimais situacija MSI

Sritis	Apibendrinta situacija
Konsultacijų teikėjai	Dažniausiai bibliotekos, IT padaliniai ir etikos priežiūros funkcijas vykdančios padaliniai
Konsultacijų pobūdis	Dažniausiai teikiamos pagal poreikį
Nuolatinė pagalba	Būdinga pavienėms institucijoms
Praktinės pagalbos prieinamumas	Dalis institucijų jos neteikia arba teikia ribotai
Dažniausios praktinės paslaugos	Duomenų saugojimas ir techninė pagalba
Platesnės praktinės paslaugos	Metaduomenų rengimas, prieigos organizavimas, stebėseną

Lietuvos AP saugyklų situacijos analizė

Apklausa tikslas nebuvo tik inventorizacija, apklausa buvo struktūrinis įrankis, leidęs susieti teorinius ES reikalavimus (AGA ir EOSC-IF) su praktine Lietuvos mokslo situacija.

Tiesioginė koreliacija su AGA kriterijais

Apklausa buvo suformuota taip, kad surinkti duomenys leistų identifikuoti saugyklų atitiktį patikimos saugyklos statusui pagal tris AGA apibrėžtas kategorijas:

- **Sertifikavimo indikacija.** Klausimai apie „CoreTrustSeal“, „nestor Seal“ ar ISO standartus leido nustatyti aukščiausią patikimumo lygį (sertifikuotos saugyklos) arba institucijų ambiciją jo siekti (pvz., KTU, VDU, LSMU planai).
- **Pripažintų disciplininių centrų identifikavimas.** Teiraudamasi apie naudojamas išorines saugyklas („Zenodo“, PDB, NCBI), apklausa patvirtino, kad Lietuvos tyrėjai integruojasi į tarptautiniu mastu pripažintą mokslinių bendruomenių saugyklas.
- **Techninės kokybės charakteristikos.** Institucijų atsakymai apie PID, tokius kaip DOI ar „Handle“, bei metaduomenų išsaugojimą po duomenų sunaikinimo yra esminiai įrodymai, kad net ir nesertifikuotos institucinės sistemos pasižymi AGA reikalaujamomis patikimumo savybėmis.

Sąveikumo (EOSC-IF) pamatai

Institucijos pateikė ne tik bendrus aprašymus, bet ir specifinius duomenis, kurie leidžia įvertinti jų sąveikumą keturiais EOSC-IF lygmenimis:

- **Techninis sąveikumas.** Pateikta informacija apie naudojamas technologijas („DSpace-CRIS“, „Dataverse“) ir sąsajas (REST API, OAI-PMH) įrodo sistemų gebėjimą automatiškai keistis duomenimis, o tai yra nacionalinio EOSC mazgo veikimo sąlyga.
- **Semantinis sąveikumas.** Nuorodos į naudojamus metaduomenų standartus („Dublin Core“, CERIF, „OpenAIRE“ gaires) rodo, kad institucijos užtikrina, jog jų duomenys būtų vienareikšmiškai suprantami ir randami globalioje ekosistemoje.
- **Teisinis ir organizacinis sąveikumas.** Pateiktos nuorodos į vidinius norminius dokumentus, intelektinės nuosavybės nuostatus ir CC licencijų taikymą leido suprasti teisinį saugyklų pamatą, be kurio saugykla negalėtų būti laikoma EOSC ekosistemos dalimi.

Institucinis vs. disciplininis požiūris

Institucijų apklausos rezultatai rodo, kad atsakymai apie naudojamą aparatinę ir programinę įrangą daugiausia atspindi bendrąjį institucinio lygmens infrastruktūros vaizdą. Tai sietina su pačia klausimų logika: institucijų buvo prašoma apibūdinti naudojamus technologinius sprendimus bendrai, todėl atsakymai natūraliai telkėsi į centralizuotai valdomus ir administraciniu lygmeniu geriausiai matomus elementus – serverius, debesijos paslaugas, institucines saugyklas, CRIS sistemas ir kitus bendruosius IT sprendimus. Tokiu būdu apklausa pirmiausia atskleidė ne tiek visą MTD valdymo praktikų įvairovę, kiek institucijos bendrą technologinį ir organizacinį pagrindą.

Apklauskos rezultatai leidžia daryti prielaidą, kad institucijų pateikti atsakymai apie naudojamą aparatinę ir programinę įrangą pirmiausia fiksuoja ne visą MTD valdymo infrastruktūrų visumą, bet tą jos dalį, kuri institucijoje centralizuotai matoma, administruojama ir apibrėžiama kaip bendra organizacinė infrastruktūra. Kitaip tariant, klausimų formuluotė ir atsakymų pobūdis sudarė palankesnes sąlygas identifikuoti bendruosius institucinio lygmens sprendimus – serverius, debesijos paslaugas, CRIS sistemas ar institucines saugyklas – nei specializuotas, konkrečių mokslo sričių ar tyrimų bendruomenių naudojamas duomenų infrastruktūras

Toks atsakymų pobūdis rodo metodologiškai svarbų skirtumą tarp institucinio ir disciplininio pjūvio. Institucinis pjūvis atskleidžia organizacijos bendrą technologinį ir valdymo pagrindą, o disciplininis ar funkcinis pjūvis

leidžia matyti, kaip duomenų valdymas faktiškai organizuojamas skirtingose mokslo srityse, kokios specializuotos saugyklos naudojamos ir su kokiomis teminėmis ar tarptautinėmis infrastruktūromis institucija yra susijusi. Dėl šios priežasties institucijų apklausos duomenys ir kituose šaltiniuose nustatytas infrastruktūrų spektras nesutampa ne todėl, kad vieni iš jų būtų netikslūs, bet todėl, kad jie aprašo skirtingus analizės lygmenis.

Šį neatitikimą iliustruoja atvejai, kai institucija apklausoje akcentuoja vieną centralizuotą sprendimą, pavyzdžiui, institucinę saugyklą ar CRIS sistemą, tačiau papildomi šaltiniai rodo jos dalyvavimą gerokai platesniame specializuotų infrastruktūrų tinkle. Tai matyti universitetų pavyzdžiuose, kur greta bendrųjų infrastruktūros sprendimų identifikuojamos ir tokios teminės ar tarptautiniu mastu reikšmingos infrastruktūros kaip MAO, COD, „CossyBio“, „LitGrid-HPC“, RILAD, CLARIN-LT, EST LT, GGP LT ar FOREST 4.0. Tokie atvejai rodo, kad apklausa labiau fiksuoja instituciškai konsoliduotą infrastruktūros vaizdą, o ne visą realiai naudojamų ar su institucija siejamų duomenų aplinkų įvairovę.

Koordinavimo trūkumas institucijos viduje

Institucijų atsakymai rodo, kad reikšminga dalis specializuotų ir išorinių duomenų saugyklų nėra sistemingai integruotos į bendrą institucinį duomenų valdymo vaizdą. Kitaip tariant, instituciniu lygmeniu ne visuomet yra vientisa apskaita ar koordinavimo mechanizmas, leidžiantis identifikuoti visas mokslininkų ar tyrimų grupių naudojamas duomenų saugojimo ir deponavimo aplinkas. Dėl šios priežasties dalis realiai naudojamų infrastruktūrų apklausos atsakymuose neatsispindi kaip institucijos infrastruktūros dalis, nors praktikoje jos atlieka svarbų vaidmenį tyrimų duomenų valdymo procese.

Bendrosios ir funkcinės infrastruktūros vaizdo skirtis

Kita reikšminga apklausos duomenų interpretavimo aplinkybė yra ta, kad dalis institucijų savo turimą infrastruktūrą apibūdina daugiausia naudodamos bendruosius technologinius ar fizinius išteklius, nepakankamai išskirdamos jų funkcinę paskirtį MTD valdymo požiūriu. Dėl to apklausos atsakymuose kai kuriais atvejais susidaro supaprastintas vaizdas, kuriame matomi techniniai ištekliai, bendrojo pobūdžio saugojimo sprendimai ar publikacijų prieigos priemonės, tačiau neišryškėja konkrečių infrastruktūrų vaidmuo duomenų kaupimo, kuravimo, deponavimo ir sklaidos procese.

Toks pateikimas riboja galimybę vien iš apklausos duomenų įvertinti faktinį institucijų dalyvavimą specializuotose duomenų infrastruktūrose. Kitaip tariant, apklausa geriau atskleidžia bendrą infrastruktūros aprūpinimą ir administraciniu požiūriu matomus sprendimus negu funkcinio požiūriu diferencijuotą duomenų valdymo aplinką. Dėl to išsamesniam vertinimui būtina apklausos rezultatus derinti su kitais šaltiniais, leidžiančiais identifikuoti institucijų sąsajas su teminėmis, disciplininėmis ir tarptautiniu mastu reikšmingomis infrastruktūromis.

Institucijų apklausos duomenys, pateikti kartu su GS teikiamoje rinkmenoje „rezultatas.csv“, pirmiausia atspindi bendrąjį institucinės infrastruktūros ir IT ūkio pjūvį, apimančį serverius, debesijos sprendimus, CRIS sistemas ir kitus centralizuotai matomus bei administruojamus technologinius sprendimus. „saugyklos3.csv“ rinkmena leidžia rekonstruoti funkcinio požiūriu išsamesnį vaizdą, rodantį, kad ta pati institucija gali būti susijusi su keliomis specializuotomis, teminėmis ar tarptautiniu mastu pripažintomis duomenų deponavimo ir valdymo infrastruktūromis. Todėl šie du duomenų pjūviai neturėtų būti vertinami kaip vienas kitam prieštaraujantys, bet veikiau kaip papildantys skirtingus analizės lygmenis: „rezultatas.csv“ labiau atskleidžia bendrąją institucinę brandą ir centralizuotai valdomą infrastruktūrinį pagrindą, o „saugyklos3.csv“ – faktinę MTD infrastruktūrų įvairovę ir funkcinę diferenciaciją. Būtent šių šaltinių sugretinimas leidžia tiksliau įvertinti realų institucijų vaidmenį Lietuvos MTD ekosistemoje.

Lietuvos institucijų saugyklų brandos ir pasirengimo formuoti EOSC LT mazgą apibendrinimas

Institucijų atsakymai daugiausia atspindi bendrą administracinį ir infrastruktūrinį jų veiklos pjūvį, todėl leidžia vertinti ne pavienes mokslo infrastruktūras, o bendrąjį institucijų IT aplinkos, duomenų valdymo organizavimo ir koordinavimo brandos lygį. Toks holistinis vaizdas yra svarbus todėl, kad būtent bendrieji instituciniai pajėgumai – technologinė bazė, valdymo sprendimai, atsakomybių pasiskirstymas ir gebėjimas koordinuoti skirtingas duomenų valdymo praktikas – sudaro prielaidas veikti ir specializuotoms, atskirų mokslo sričių poreikiams pritaikytoms saugykloms.

Ne mažiau reikšminga ir tai, kad institucijų pateikta informacija atskleidžia ne tik esamus pajėgumus, bet ir struktūrinius ribotumus. Atvirai įvardijami koordinavimo trūkumai, specializuotų kompetencijų stoka ar poreikis stiprinti institucinį duomenų valdymo organizavimą suteikia svarbų empirinį pagrindą nacionalinio lygmens sprendimams. Tai leidžia tiksliau nustatyti sritis, kuriose reikalingos viešosios investicijos, metodinė pagalba ar koordinavimo mechanizmai, siekiant nuosekliai plėtoti nacionalinę EOSC LT mazgo ekosistemą.

Institucijų pateikta informacija, nors ir suformuluota per bendrą infrastruktūrinę prizmę, suteikė objektyvų pagrindą klasifikuoti Lietuvos AP saugyklas pagal tarptautinius brandos modelius. Tai leido nustatyti, kad nacionalinė ekosistema disponuoja kritiniais patikimumo ir techninio sąveikumo komponentais (API, PIDs, standartizuoti metaduomenys), kurie būtini siekiant visavertės integracijos į EOSC.

Remiantis atliktos apklausos rezultatais bei jų analize mokslo politikos kontekste, pateikiamas apibendrinantis vertinimas dėl Lietuvos institucijų pasirengimo atitikti patikimos ir sąveikios saugyklos statusą bei formuoti nacionalinį EOSC mazgą.

Atlikta mokslo ir studijų institucijų apklausa atskleidė netolygų, tačiau techniškai perspektyvų vaizdą. Nors institucijos savo infrastruktūrą dažnai vertina per bendrą IT ūkio prizmę, gauti duomenys leidžia identifikuoti jų faktinę atitiktį tarptautiniams standartams.

- **Atitiktis patikimos saugyklos statusui (pagal AGA).** Esminius patikimumo kriterijus – nuolatinių identifikatorių (PIDs) naudojimą ir standartizuotų metaduomenų taikymą – jau dabar atitinka dalis didžiausių Lietuvos universitetų (VU, KTU, VDU, LSMU, VILNIUS TECH, KU, MRU), kuriuose naudojama specializuota programinė įranga („DSpace-CRIS“, „Datavers“). Vis dėlto formalaus sertifikavimo srityje institucijų praktikos šioje yra nevienodos: dalis jų yra inicijavusios arba numato inicijuoti patikimos saugyklos sertifikavimo procesus, įskaitant atitiktį „CoreTrustSeal“ reikalavimams, o kitos prioritetą teikia ISO standartais grindžiamoms kokybės ar informacijos valdymo sistemoms diegti. Tai leidžia teigti, kad formalaus patikimumo užtikrinimo kryptis Lietuvos mokslo ir studijų sistemoje jau ryškėja, tačiau jos institucinė sklaida ir brandos lygis tebėra netolygūs. Kolegijose duomenų saugojimo praktikos dažniau grindžiamos bendrosios paskirties debesijos sprendimais, kurie pagal AGA nuostatas nėra laikomi patikimomis mokslinių tyrimų duomenų saugyklomis.

- **Atitiktis sąveikios saugyklos statusui (pagal EOSC-IF).** Techninio sąveikumo pamatas Lietuvoje yra tvirtas dėl LITNET infrastruktūros ir plačiai naudojamų atvirų protokolų (pvz., OAI-PMH, REST API). Tačiau semantinis sąveikumas išlieka didžiausiu iššūkiu: nors institucijos naudoja bazinius standartus („Dublin Core“), trūksta bendro požiūrio į specifinius mokslo sričių žodynus ir ontologijas. Organizaciniu lygmeniu pastebimas ryškus fragmentiškumas – institucijos pripažįsta, kad mokslininkų naudojamų tarptautinių saugyklų („Zenodo“, PDB, NCBI) integravimas į vidinę apskaitą dažnai nėra koordinuojamas.

- **Pasirengimas formuoti EOSC LT NODE.** MSI pasirengimą formuoti nacionalinį EOSC mazgą galima vertinti kaip techniškai pažengusį, bet organizaciškai nebrandų. **Branduolio lygiu:** MIDAS (VU) ir LiDA (KTU) jau dabar veikia kaip natūralūs agregatoriai, turintys didžiausią technologinį potencialą tapti mazgo ašimi. **Teminiu lygiu:** nacionalinės MTI, įtrauktos į LMT kelrodį, pasižymi aukšta moksline branda ir tarptautine integracija, tačiau joms būtinas papildomas „vertimo“ sluoksnis (*Mapping Repositories*) sujungti su centriniu mazgu. **Instituciniu lygiu:** universitetai demonstruoja norą tapti federacijos nariais, tačiau kolegijoms ir smulkiems institutams EOSC LT NODE būtų esminis išsigelbėjimas nuo nepatikimų debesijos paslaugų naudojimo.

Instituciniu lygmeniu duomenų valdymo pajėgumai yra labai nevienodi. Daugelis Lietuvos aukštųjų mokyklų ir mokslo institucijų MTD valdymo ir tvarkymo funkcijas vykdo nevienodai. Jas dažniausiai atlieka bibliotekų, IT padalinių, projektų administravimo ar etikos komitetų darbuotojai, o aiškiai identifikuojamos ir dedikuotos MTD valdymo specialistų pareigybės egzistuoja retai. Didieji universitetai šioje srityje yra pažengę labiausiai – juose MTD funkcijas vykdo įvairių pareigybių specialistai, dirbantys daliniu arba visu etatu, daugiausia bibliotekose ir su mokslu ar IT susijusiuose padaliniuose. Kolegijose ir mokslo institutuose dedikuotų MTD specialistų dažniausiai nėra, o šios funkcijos vykdomos fragmentiškai, derinant jas su kitomis pareigomis.

MTD valdymo mokymai Lietuvos aukštosiose mokyklose ir mokslo institucijose taip pat nėra sistemaiškai

išplėtoti. Tik nedidelė dalis institucijų siūlo formalizuotus, kreditais grįstus MTD valdymo kursus. Dažniausiai mokymai yra neformalūs arba fragmentiški, o MTD temos integruojamos netiesiogiai – per tyrimų metodologijos, tyrimų etikos, baigiamųjų darbų rengimo dalykus, ypač doktorantūros lygmeniu. Tyrėjų mokymuose aktyviausi yra keli universitetai.

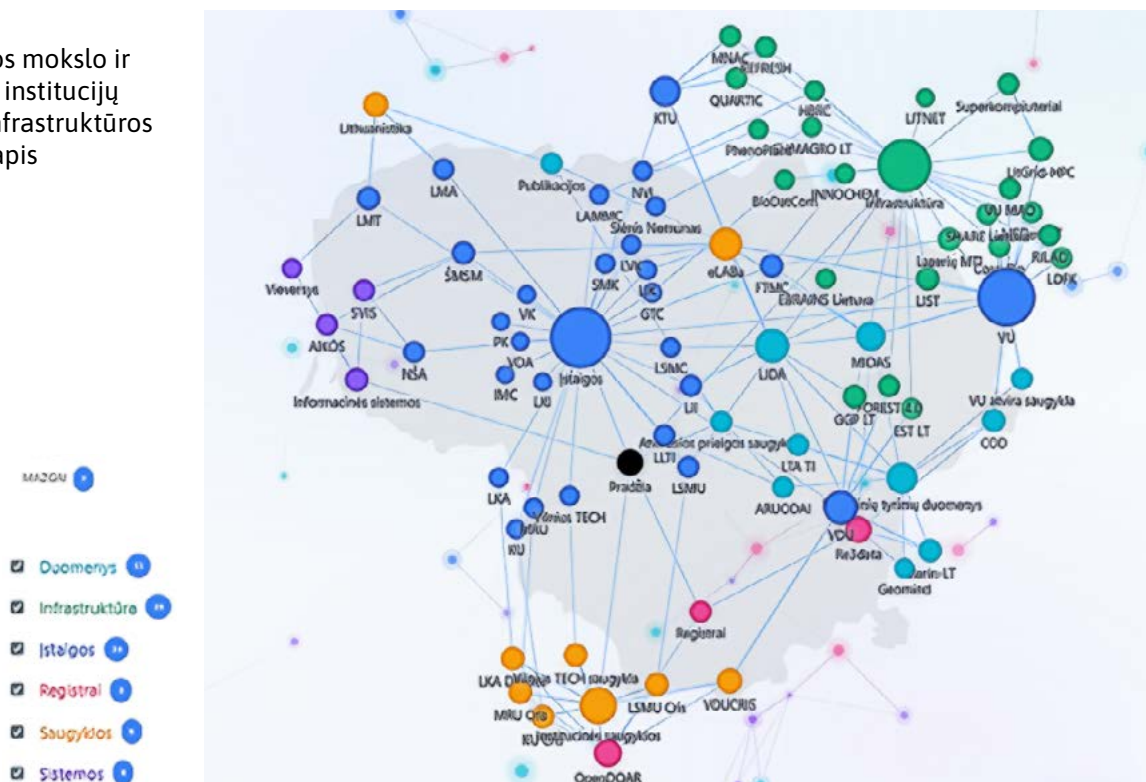
Bibliotekos institucijose atlieka esminį vaidmenį ugdant MTD valdymo kompetencijas: jos organizuoja mokymus ir seminarus, teikia individualias konsultacijas, konsultuoja dėl duomenų valdymo planų (DVP), FAIR principų taikymo, duomenų saugyklų ir atvirosios prieigos. Individuali ir praktinė pagalba daugumoje institucijų teikiama pagal poreikį, tačiau ji ne visur yra institucionalizuota ir dažnai priklauso nuo konkrečių specialistų kompetencijų. Ne visose institucijose aktyviai fiksuojamas tyrėjų poreikis tokioms konsultacijoms, o MTD valdymo atsakomybė dažnai paliekama pačių tyrėjų ar tyrimų grupių sprendimui.

MTD valdymo praktikoms Lietuvoje reikšmingą įtaką daro išoriniai veiksniai – projektų reikalavimai, tarptautinės narystės, institucijų strategijos ir finansavimo modeliai, kurie skatina laikytis duomenų valdymo planų ir FAIR principų. Tai lemia, kad MTD valdymo brandos lygis institucijose yra labai įvairus, o pažanga daugiausia priklauso nuo turimų žmogiškųjų išteklių, infrastruktūros ir institucinio prioritetizavimo.

MTD valdymo brandos ir pajėgumų netolygumai institucijose taip pat atskleidžia platesnę sistemine problemą – Lietuvoje ilgą laiką stigo vieningo, vienoje vietoje prieinamo MTD ekosistemos vaizdo. Informacija apie institucinius pajėgumus, infrastruktūrinius sprendimus, teikiamas paslaugas ir sąsajas su specializuotomis duomenų valdymo infrastruktūromis buvo fragmentiška ir išskaidyta, todėl riboja galimybes sistemiškai vertinti ekosistemos būklę, identifikuoti struktūrinius netolygumus ir pagrįsti tolesnes intervencijas. Siekiant prisidėti prie šios spragos mažinimo, studijos rengimo metu sukurtas Lietuvos mokslo ir studijų MTD infrastruktūros žemėlapis. Jis yra ne tik informacinė, bet ir analitinė priemonė, leidžianti struktūruoti fragmentišką lauką, integruoti skirtingus duomenų pjūvius ir didinti visos sistemos matomumą. Žemėlapis sudaro prielaidas tiksliau identifikuoti institucinių pajėgumų pasiskirstymą, infrastruktūrinius trūkumus, funkcinės diferenciacijos mastą ir galimas prioritėtines stiprinimo kryptis. Su žemėlapiu galima susipažinti Lietuvos mokslo tarybos interneto svetainėje <https://lmt.lv.lt/veiklos-sritys/mokslo-politika/atvirasis-mokslas/lmsi-zemelapis/> jo fragmentas pateikiamas 5 paveiksle.

5 pav.

Lietuvos mokslo ir studijų institucijų MTD infrastruktūros žemėlapis



4.2. Mokslininkų apklausos analizė

Trečiajame tyrime siekta išsiaiškinti Lietuvos mokslininkų patirtis dirbant su MTD ir naudojantis MTI. Tam taip pat buvo atliekama internetinė apklausa (apie trečiosios apklausos vykdymo eigą žr. 1.2 poskyrį).

Apklausoje sudalyvavo 232 Lietuvos mokslininkai iš 1568, į kuriuos kreiptasi (14,8 proc.). Didžioji dalis (76,2 proc.) kontaktuočių mokslininkų dirbo universitetuose (žr. 14 lentelę).

14 lentelė. Trečiajame tyrime dalyvauti kvieisti mokslininkai pagal instituciją

Institucija	N	Proc.
Universitetas	1195	76,2
Kolegija	10	0,6
Institutas	363	23,2
Iš viso	1568	100

Didžioji dalis apklausoje dalyvavusių mokslininkų buvo pripažinti tyrėjai (54,3 proc.). Mažiausiai buvo pradedančiųjų ir patvirtintų tyrėjų (9 proc.), o dar kiek daugiau nei trečdalis (35,8 proc., žr. 15 lentelę) buvo pirmaujantys tyrėjai.

15 lentelė. Trečiajame tyrime dalyvavę mokslininkai pagal tyrėjų karjeros pakopas

Tyrėjų karjeros pakopa	N	Proc.
Pradedantysis tyrėjas	1	0,4
Patvirtintas tyrėjas	20	8,6
Pripažintas tyrėjas	126	54,3
Pirmaujantysis tyrėjas	83	35,8
Neatsakė	2	0,9
Iš viso	232	100

Pagal mokslinės veiklos patirtį metais apklausoje dalyvavusių mokslininkų pasiskirstymas gana proporcingas. Dvi didžiausios grupės – mokslininkai, turintys 11–20 (38,4 proc.) arba 21–30 metų (31,8 proc.) mokslinės veiklos patirtį (žr. 16 lentelę).

16 lentelė. Trečiajame tyrime dalyvavę mokslininkai pagal mokslinės veiklos patirtį metais

Patirtis metais	N	Proc.
≤10	32	13,8
11–20	90	38,78
21–30	73	31,5
≥31	35	15,1
Neatsakė	2	0,9
Iš viso	232	100

Pagal mokslo (meno) sritis apklaustų mokslininkų pasiskirstymas gana tolygiai atspindi mokslininkų, kurie buvo kviešti dalyvauti tyrime, pasiskirstymą. Iš dalyvavusiųjų 64,7 proc. dirbo gamtos ir technikos mokslų (toliau – GTM) srityje, o 35,3 proc. – humanitarinių ir socialinių mokslų (toliau – HSM) srityje. Tarp mokslininkų, kurie buvo kviešti dalyvauti apklausoje, 67,0 proc. dirbo GTM srityje, o 33,0 proc. – HSM srityje. Daugiausia mokslininkų apklausoje atstovauja gamtos ir socialiniams mokslams (žr. 17 lentelę).

17 lentelė. Trečiajame tyrime dalyvavę mokslininkai pagal mokslo (meno) sritis.

Mokslo (meno) sritis	N	Proc.
Gamtos mokslai	77	33,2
Technologijos mokslai	36	15,5
Medicinos ir sveikatos mokslai	18	7,7
Žemės ūkio mokslai	17	7,4
Socialiniai mokslai	61	26,3
Humanitariniai mokslai	18	7,7
Scenos ir ekrano menai	0	0,0
Vaizduojamieji menai	0	0,0
Tarpdisciplininė	5	2,2
Iš viso	232	100

Trečiojo tyrimo klausimynas suskirstytas į šešias temines sritis, kurios apima mokslininkų santykį su MTD, jų naudojamą infrastruktūrą, institucines sąlygas, teisinius reikalavimus, motyvaciją, kompetencijų poreikį ir duomenų valdymo praktikas.

Santykis su mokslinių tyrimų duomenimis ir duomenų kraštovaizdis

Mokslininkų apklausos rezultatai rodo, kad MTD kūrimas Lietuvos mokslo sistemoje yra plačiai paplitusi praktika, tačiau kiti duomenų gyvavimo ciklo etapai yra gerokai mažiau įsitvirtinę. Net 91,8 proc. respondentų nurodė kuriantys MTD, tačiau tik 51,7 proc. teigė juos tvarkantys ir saugantys, o 40,5 proc. – validuojantys (7 priedas, 1 lentelė). Tai leidžia daryti išvadą, kad duomenų generavimas daugeliui tyrėjų jau yra natūrali mokslinės veiklos dalis, tačiau sistemingas duomenų valdymas vis dar nėra vienodai integruotas į kasdienę tyrimų praktiką. Kitaip tariant, duomenų kūrimas sistemoje yra gerokai labiau išplėtotas nei jų nuoseklus parengimas ilgalaikiam saugojimui, atvėrimui ar pakartotiniam naudojimui.

Analizuojant duomenų tipus, išryškėja aiškūs skirtumai tarp gamtos, technologijos, medicinos ir sveikatos, žemės ūkio mokslų sričių (toliau – GTM) bei humanitarinių ir socialinių mokslų sričių (toliau – HSM). GTM srityse statistiškai reikšmingai dažniau dirbama su eksperimentiniais duomenimis, kuriuos nurodė 84 proc. šios grupės respondentų, taip pat su skaičiavimų ir modeliavimo duomenimis, vaizdiniais duomenimis, programinės įrangos ir kodo duomenimis, dirbtinio intelekto bei prognozavimo modeliais. HSM srityse ryškiau dominuoja apklausų duomenys, tekstiniai ir dokumentiniai duomenys, garso ir vaizdo medžiaga bei skaitmenizuoti archyviniai šaltiniai. Tai rodo, kad Lietuvos tyrėjų duomenų profiliai yra labai nevienalyčiai, o skirtingų mokslo sričių poreikiai duomenų valdymo, aprašymo, saugojimo ir atvėrimo požiūriu negali būti vertinami kaip vienodi (7 priedas, 2 lentelė). Svarbu ir tai, kad skiriasi ne tik duomenų tipai, bet ir jų kiekiai. GTM tyrėjai gerokai dažniau nurodo kuriantys didesnę eksperimentinių, modeliavimo ir vaizdinių duomenų rinkinių skaičių per metus. Pavyzdžiui, 14,7 proc. GTM tyrėjų teigė per metus sukuriantys daugiau kaip 100 eksperimentinių duomenų rinkinių, o HSM grupėje tokių buvo 2,4 proc. Panašus skirtumas matomas ir vaizdinių duomenų atveju – daugiau kaip 100 tokių rinkinių per metus nurodė sukuriantys 14 proc. GTM respondentų ir 4,9 proc. – HSM respondentų. HSM grupėje dažniau pasitaiko

mažesni tekstinių, dokumentinių, garso ir vaizdo ar archyvinių duomenų rinkiniai, nors dalyje kategorijų matyti ir pavienių didesnių apimčių atvejų. Taigi apklausa leidžia teigti, kad tyrėjų duomenų kūrimo intensyvumas yra tiesiogiai susijęs su mokslo srities specifika (7 priedas, 4 lentelė).

Analizuojant duomenų apimtis matyti, kad daugeliu atvejų mokslininkai nurodo sukuriantys iki 50 GB apimties rinkinius, tačiau tam tikrose srityse išryškėja ir didesnių apimčių segmentai. GTM tyrėjai dažniau dirba su didesnės apimties eksperimentiniais, modeliavimo ir vaizdiniais duomenimis: pavyzdžiui, 13,3 proc. GTM respondentų eksperimentinių duomenų rinkinius įvertino kaip 51–100 GB, o 8,7 proc. – kaip 101 GB–10 TB apimties. HSM tyrėjų grupėje didesnės apimtys labiau išryškėja tekstinių, dokumentinių, garso ir vaizdo bei skaitmenizuotų archyvinių duomenų kategorijose. Tai rodo, kad net ir tose srityse, kurios įprastai laikomos mažesnės duomeninės apkrovos laukais, yra reikšmingų infrastruktūrinių poreikių. Apklausa taip pat leidžia matyti, kad didelių apimčių duomenys nėra išskirtinai vienos mokslo srities bruožas, tačiau jų pobūdis ir tipai tarp sričių labai skiriasi (7 priedas, 6 lentelė).

MTD kūrimas Lietuvoje jau yra beveik universali praktika, tačiau jų sistemingas kuravimas ir parengimas ilgalaikiam saugojimui, atvėrimui bei pakartotiniam naudojimui vis dar išlieka silpniausia duomenų gyvavimo ciklo grandimi.

Duomenų tvarkymo, saugojimo ir dalijimosi praktikos

Tyrimas rodo, kad duomenų dalijimosi praktikos Lietuvoje yra nevienalytės, o atviro dalijimosi modelis dar nėra tapęs vyraujančia norma. Dažniausiai respondentai nurodė dalijantys duomenis kaip priedus prie publikacijų – taip atsakė 55,2 proc. apklaustųjų. Kartu 51,3 proc. respondentų nurodė duomenimis besidalijantys tik gavę asmeninę užklausą, 37,1 proc. – institucijos viduje, o atvirose duomenų saugyklose duomenis teikia 27,2 proc. apklaustųjų. 15,5 proc. respondentų nurodė duomenimis apskritai nesidalijantys. Tokia struktūra leidžia daryti išvadą, kad tyrėjų praktikoje kol kas labiau paplitęs riboto ar selektyvaus prieinamumo modelis, o ne iki galo institucionalizuotas atvirų duomenų modelis (7 priedas, 7 lentelė).

Tarp mokslo sričių identifikuota skirtumų. GTM respondentai dažniau dalijasi duomenimis kaip publikacijų priedais, pagal bendradarbiavimo sutartis ir institucijos viduje, o HSM grupėje santykinai daugiau tyrėjų duomenimis dalijasi tik gavę asmeninę užklausą arba iš viso jais nesidalija. Šie skirtumai leidžia manyti, kad dalijimosi modelius veikia ne tik individualūs tyrėjų pasirinkimai, bet ir duomenų tipų jautrumas, jų naudojimo kontekstas bei susiformavusios konkrečių sričių tyrimų praktikos. Kitaip tariant, duomenų dalijimasis Lietuvoje vyksta ne vien pagal bendrą atvirojo mokslo logiką, bet ir pagal disciplinines veiklos normas.

Koreliacija matoma ir tarp duomenų valdymo planų rengimo ir dalijimosi praktikų. Tyrėjai, kurie rengia duomenų valdymo planus todėl, kad to reikalauja finansuojanti organizacija, gerokai dažniau nurodo besidalijantys duomenimis per atviras saugyklas nei tie, kurie tokių planų nerengia. Tarp planų nerengiančių respondentų gerokai didesnė dalis apskritai nesidalija duomenimis. Tai leidžia teigti, kad formalūs reikalavimai, ypač kylantys iš finansavimo sistemos, turi apčiuopiamą poveikį praktinei duomenų atvėrimo elgsenai. Kita vertus, vien duomenų valdymo planų egzistavimas automatiškai negarantuoja pažangesnės praktikos, tačiau apklausa rodo, kad jis bent iš dalies susijęs su brandesniu požiūriu į duomenų tvarkymą ir dalijimąsi (7 priedas, 9 lentelė).

Pažangesnės MTD praktikos Lietuvoje labiau siejasi su aiškiais reikalavimais, institucinėmis paskatomis ir palankesne aplinka nei su vien individualiu savanorišku mokslininkų angažavimusi.

Pažymėtina ir tai, kad naudojimasis specializuota pagalba, pavyzdžiui, duomenų valdytojų paslaugomis, apklausoje beveik nefiksuojamas. Kadangi tokiomis paslaugomis naudojasi tik 5 respondentai, šis kanalas kol kas negali būti laikomas realiai išplėtotu tyrėjų praktikos elementu. Tai leidžia manyti, kad specializuota pagalba duomenų valdymo srityje Lietuvos mokslo sistemoje tebėra menkai prieinama arba dar nėra plačiau institucionalizuota.

Nesidalijimo duomenimis priežastys ir pagrindiniai sunkumai

Tyrėjų atsakymai rodo, kad duomenų nesidalijimą lemia ne viena pagrindinė kliūtis, o kelių skirtingų veiksnių derinys. Dažniausiai nurodomos priežastys yra duomenų konfidencialumas, nesutvarkyti duomenys, baimė,

kad jais pasinaudos kiti, paskatų ar pripažinimo sistemos stoka, taip pat autorių teisių, asmens duomenų apsaugos ir etinių apribojimų klausimai. Tai leidžia manyti, kad duomenų nesidalijimas nėra vien paprastas tyrėjų nenoras atverti duomenis. Daugiausia jis susijęs su objektyviomis teisinėmis, etinėmis, techninėmis ir institucinėmis kliūtėmis, kurios įvairiose mokslo srityse pasireiškia nevienodai.

HSM srityse statistiškai reikšmingai dažniau nei GTM respondentai nurodo etinius apribojimus ir asmens duomenų apsaugos klausimus. GTM srityse dažniau išryškėja infrastruktūros ir techninio parengimo problemų, susijusių su didesniais duomenų kiekiais, jų apdorojimu ar saugojimu. Taigi duomenų atvėrimo kliūtys taip pat turi aiškiają disciplininę dimensiją.

Skirtingos mokslo sritys susiduria ne su tomis pačiomis MTD atvėrimo kliūtėmis: vienur dominuoja infrastruktūros ir technologiniai poreikiai, kitur – etiniai, teisiniai ir kompetencijų iššūkiai.

Kalbant apie bendruosius sunkumus kuriant mokslinių tyrimų duomenis, aiškiausiai išsiskiria laiko trūkumas – jį nurodė 69 proc. respondentų. Be to, 44,8 proc. respondentų pažymėjo kompetencijų trūkumą, 44 proc. – neaiškius reikalavimus, 34,5 proc. – infrastruktūros trūkumą, o 33,6 proc. – finansavimo stoką. Šie rezultatai rodo, kad duomenų valdymas daugeliui tyrėjų vis dar suvokiamas kaip papildoma, šalia pagrindinės mokslinės veiklos egzistuojanti atsakomybė, kuriai dažnai trūksta aiškaus laiko, išteklių ir metodinio palaikymo. Kitaip tariant, apklausa leidžia kalbėti ne tik apie individualaus pasirengimo, bet ir apie sisteminės aplinkos ribotumą (7 priedas, 11 lentelė).

Skirtumai pagal tyrėjų karjeros pakopos stadiją taip pat reikšmingi. Pradedantieji ir patvirtinti tyrėjai dažniau nei kitos grupės nurodo kompetencijų ir reikalavimų aiškumo trūkumą, o pirmaujantieji tyrėjai dažniau mini laiko ir finansavimo stoką. Tai rodo, kad skirtinguose karjeros etapuose tyrėjai susiduria su nevienodais duomenų valdymo iššūkiais: ankstesniuose tyrėjo karjeros etapuose labiau atsispindi gebėjimų ir reikalavimų aiškumo stygius, o vėlesniuose – ribotas laikas ir poreikis didesniems ištekliams. Tokia išvada pagrindžia poreikį kurti diferencijuotas pagalbos priemones skirtingoms tyrėjų grupėms (7 priedas, 12 lentelė).

Svarbu ir tai, kad individualūs sunkumai gana aiškiai siejasi su tuo, kaip tyrėjai vertina institucines sąlygas. Tarp tų, kurie institucines MTD tvarkymo sąlygas vertina kaip netinkamas, 76,9 proc. nurodo infrastruktūros trūkumą, 57,7 proc. – kompetencijų stoką, ir tiek pat neaiškius reikalavimus. Tarp respondentų, kurie institucines sąlygas vertina kaip tinkamas, infrastruktūros trūkumą mini tik 12 proc., kompetencijų trūkumą – 30,1 proc. Tai aiškiai rodo, kad nemaža dalis tyrėjų kasdienių sunkumų yra susiję ne tik su individualiais gebėjimais, bet ir su platesniu instituciniu kontekstu (7 priedas, 13 lentelė).

FAIR principų supratimas ir taikymas

Tyrimo metu išryškėjo, kad FAIR principai Lietuvos tyrėjų bendruomenėje dar nėra tapę universaliai aiškiai ir iki galo įgyvendinamu praktiniu standartu. 31,9 proc. respondentų nurodė, kad FAIR principų reikalavimai jiems yra aiškūs ir suprantami. Dar 34,1 proc. pažymėjo, kad šie reikalavimai jiems aiškūs tik iš dalies, o tokia pati dalis – 34,1 proc. – teigė, kad jie apskritai nėra aiškūs. Tai rodo, kad bent du trečdaliai apklaustų tyrėjų veikia riboto aiškumo aplinkoje, kurioje FAIR principai nėra iki galo įsivertinti kaip praktiniai mokslinės veiklos standartai.

Tarp mokslo sričių taip pat matomi tam tikri skirtumai: GTM grupėje FAIR reikalavimus aiškiais laiko 35,3 proc. respondentų, HSM grupėje – 25,6 proc. Nors skirtumas nėra drastiškas, jis rodo, kad kai kuriose srityse šie principai šiek tiek labiau įsitvirtinę. Vis dėlto abiejose grupėse bendra situacija rodo, kad metodinio aiškumo lygis yra nepakankamas, jei siekiama nuoseklaus FAIR principų taikymo visoje sistemoje (7 priedas, 14 lentelė).

Ypač reikšmingas yra ryšys tarp FAIR supratimo ir realaus jų taikymo. Tarp tyrėjų, kuriems FAIR principai aiškūs, 70,3 proc. nurodo, kad jų duomenys yra surandami, 74,3 proc. – kad prieinami, 66,2 proc. – kad sąveikūs, o 54,1 proc. – kad pakartotinai naudojami. Tarp tų, kurie FAIR supranta tik iš dalies, visi šie rodikliai yra kur kas mažesni, o FAIR apskritai netaikančių dalis siekia 36,7 proc. (7 priedas, 16 lentelė).

Ne mažiau svarbus ir ryšys tarp FAIR supratimo bei realių dalijimosi praktikų. Tyrėjai, kuriems FAIR principai aiškūs, gerokai dažniau dalijasi duomenimis per atviras saugyklas – tokių yra 48,6 proc. Tarp tų, kuriems FAIR principai neaiškūs, per atviras saugyklas dalijasi tik 7,6 proc. respondentų. Be to, tarp FAIR principų nesuprantančių tyrėjų gerokai didesnė dalis apskritai nesidalija duomenimis. Taigi apklausa aiškiai rodo, kad FAIR supratimas yra

vienas reikšmingiausių pažangesnės duomenų valdymo elgsenos veiksnių (7 priedas, 17 lentelė).

Dar kitoks vaizdas susidaro pažvelgus į atskirų FAIR principų taikymo paplitimą. Išsamų arba dalinį susipažinimą su FAIR principais deklaravusius respondentus (N = 153) paklausus, ar ir kuriuos iš FAIR principų taiko savo pačių duomenims, tik ketvirtadalis iš jų nurodė, kad jų duomenys atitinka visus keturis principus, maždaug pusė nurodė, kad jų duomenys atitinka vieną ar kelis iš keturių principų, o likęs ketvirtadalis nurodė, kad jų duomenys FAIR principų neatitinka (7 priedas, 18 lentelė).

Institucinė aplinka ir jos poveikis tyrėjų praktikoms

Tyrimo metu analizuota ne tik individualios mokslininkų praktikos, bet ir tai, kaip jos siejasi su institucine aplinka. Rezultatai rodo, kad geresnės institucinės sąlygos dažniau siejasi su brandesnėmis dalijimosi ir FAIR taikymo praktikomis. Tarp tyrėjų, kurie institucines MTD tvarkymo sąlygas vertina kaip tinkamas, 34,9 proc. dalijasi duomenimis per atviras saugyklas, kai tarp sąlygas kaip netinkamas vertinančių respondentų tokių yra 19,2 proc. Be to, tarp sąlygas kaip netinkamas vertinančių tyrėjų gerokai dažniau pasitaiko ir nesidalijimo atvejų. Tai leidžia spręsti, kad duomenų valdymo praktikos priklauso ne vien nuo asmeninės motyvacijos ar tyrėjo kompetencijų, bet ir nuo to, kiek palanki ir funkcionuojanti yra institucijos sukurta aplinka (7 priedas, 19 lentelė).

Panaši tendencija matoma ir kalbant apie paskatas. Tose institucijose, kuriose tyrėjai nurodo jaučiantys skatinimą kaupti, archyvuoti ir dalintis duomenimis, dažniau pasitaiko duomenų valdymo planų rengimas, dalijimasis pagal bendradarbiavimo sutartis ar institucijos viduje, o nesidalijimo rodikliai yra mažesni. Ten, kur paskatų nėra arba jos neaiškios, dažniau vyrauja pasyvesnės praktikos. Nors apklausa neleidžia vienareikšmiškai nustatyti priežastinio ryšio, ji gana aiškiai rodo, kad institucinė paskatų sistema yra vienas iš veiksnių, susijusių su aktyvesniu ir brandesniu duomenų valdymu (7 priedas, 21 lentelė).

Tyrėjų patiriami sunkumai MTD srityje yra ne vien individualūs – jie susiję su tuo, kaip vertinamos institucijų sudarytos sąlygos tvarkyti duomenis.

Tarptautiškumas kaip pažangos veiksnys (?)

Tyrimas atskleidžia, kad tarptautiškumas yra reikšmingas pažangesnių duomenų valdymo praktikų veiksnys. Tyrėjai, kurie dalyvauja tarptautinių mokslinių tyrimų infrastruktūrų ar užsienio mokslo institucijų veikloje vykdant bendrus projektus, dažniau dalijasi duomenimis atvirose saugyklose, pagal bendradarbiavimo sutartis ir institucijos viduje. Tarp šios grupės rečiau pasitaiko apskritai nesidalijančių duomenimis respondentų. Tai leidžia daryti išvadą, kad tarptautiniai bendradarbiavimo tinklai veikia kaip savotiškas pažangesnių standartų ir praktikų perdavimo kanalas (7 priedas, 23 lentelė).

Panaši tendencija matoma ir FAIR principų taikymo srityje. Tyrėjai, dalyvaujantys tarptautinėje veikloje vykdant bendrus projektus, dažniau nurodo taikantys surandamumo, prieinamumo, sąveikumo ir pakartotinio naudojimo principus, o FAIR netaikančių dalis tarp jų yra mažesnė. Taigi apklausos duomenys leidžia teigti, kad tarptautiškumas gali būti laikomas vienu iš veiksnių, susijusių su brandesniu duomenų valdymu ir didesne orientacija į atviresnius standartus (7 priedas, 25 lentelė).

Tyrėjai, dalyvaujantys tarptautinėse veiklose, dažniau dalijasi duomenimis, dažniau taiko FAIR principus ir rečiau apskritai nesidalija MTD, todėl tarptautiškumą galima laikyti vienu iš veiksnių, susijusių su brandesnėmis duomenų valdymo praktikomis.

Apibendrinimas

Apibendrinant mokslininkų apklausos rezultatus teigtina, kad Lietuvos mokslo sistemoje MTD kūrimas jau yra tapęs plačiai paplitusia ir daugeliu atvejų įprasta mokslo veiklos dalimi, tačiau šių duomenų valdymo, kuravimo, atvėrimo ir pakartotinio naudojimo brandumas tebėra netolygus, fragmentiškas ir nepakankamai institucionalizuotas. Tyrimo duomenys rodo, kad sistema šiandien stipresnė duomenų generavimo nei jų viso gyvavimo ciklo valdymo grandyje: nors tyrėjai duomenis kuria beveik universaliai, gerokai mažesnė jų dalis juos sistemingai tvarko, validuoja, atveria ar parengia ilgalaikiam naudojimui. Kartu išryškėja, kad pažangesnės MTD praktikos

nėra atsitiktinės ar vien individualios geros valios rezultatas – jos nuosekliai siejasi su aiškesniais reikalavimais, FAIR principų supratimu, palankesne institucine aplinka, paskatomis, duomenų valdymo planų rengimu ir tarptautinio bendradarbiavimo patirtimi. Ne mažiau svarbu tai, kad tyrimas paneigia prielaidą apie vienalytę MTD situaciją visose mokslo srityse: skirtingose mokslo srityse ar sričių grupėse susiklosto nevienodi duomenų profiliai, apimtys, rizikos ir kliūtys, todėl vienodas reguliacinis ar infrastruktūrinis sprendimas visai sistemai nebūtų pakankamas. Lietuvoje MTD pažanga negali būti grindžiama vien formalių nuostatų egzistavimu ar pavienių tyrėjų iniciatyva. Jai būtina nuosekli, diferencijuota ir sistemiškai veikianti ekosistema, apimanti metodinį aiškumą, kompetencijų stiprinimą, specializuotą tarpinę pagalbą, adekvačią infrastruktūrą, institucines paskatas ir realiai veikiančius duomenų atvėrimo bei pakartotinio naudojimo mechanizmus.

Apklauso duomenys leidžia teigti, kad tarptautiškumas yra susijęs su brandesniu duomenų valdymu ir didesne orientacija į atviresnius standartus, tačiau šis ryšys tikėtina bent iš dalies atspindi ir griežtesnius projektinius bei institucinius reikalavimus, būdingus tarptautinio bendradarbiavimo aplinkai.

4.3. Fokusuotosios grupinės diskusijos

Siekiant Lietuvos MSI pristatyti atliktų tyrimų rezultatus ir papildomai patikslinti ar pagrįsti preliminarias įžvalgas ir išvadas dėl Lietuvos jungimosi į EOSC planavimo, buvo įgyvendintos ir dvi fokusuotosios grupinės diskusijos su MSI atstovaujančiais atsakingais asmenimis (plačiau apie fokusuotųjų grupinių diskusijų metodiką žr. 1.2.2 poskyrį). Diskusijos leido ne tik pristatyti tarpinius rezultatus, bet ir patikrinti jų interpretaciją, išgryninti svarbiausias sisteminės problemas bei geriau suprasti skirtingų institucijų poreikius ir lūkesčius.

Diskusijos patvirtino, kad Lietuvos MSI noras jungtis į EOSC yra didelis, tačiau faktinis pasirengimas tam išlieka nepakankamas ir netolygus. Dalyviai pabrėžė, kad nacionalinė MTD sistema išlieka fragmentuota, o tarp duomenų kūrimo masto ir jų valdymo brandumo išlieka ryškus atotrūkis. Net ir didžiausios šalies mokslo ir studijų institucijos akcentavo papildomų investicijų poreikį infrastruktūrai, metaduomenų kokybei, FAIR principų įgyvendinimui ir duomenų valdymo kompetencijoms stiprinti.

Kita svarbi diskusijose išryškėjusi tema buvo nacionalinio veikimo modelio aiškumo stoka. Dalyviai akcentavo poreikį aiškiau apibrėžti nacionalinius tikslus, stebėsenos logiką, duomenų atvėrimo prioritetus ir institucinių bei nacionalinių sprendimų santykį. Buvo pabrėžta, kad Lietuvai reikalinga ne visiška MTD centralizacija, bet nuoseklesnė koordinavimo sistema, ypač metaduomenų, standartų, konsultavimo ir bendrų paslaugų lygmeniu, kartu išlaikant galimybę pačius duomenis saugoti juos sukūrusiose institucijose arba specializuotose saugyklose.

Diskusijos taip pat parodė, kad institucijų poreikiai skiriasi priklausomai nuo jų dydžio, profilio ir mokslo sričių specifikos. Mažesnės institucijos ir kolegijos labiau akcentavo praktinių rekomendacijų, konsultavimo, kooperacijos ir bendrų sprendimų poreikį, nes savarankiškai plėtoti visus MTD valdymo procesus joms dažnai stinga žmogiškųjų ir technologinių išteklių. Tuo tarpu menų, humanitarinių ir sveikatos mokslų atstovai išskyrė (tarp) disciplininius ypatumus – autorių teisių ribojimus, jautrių duomenų apsaugą, nuasmeninimo poreikį, specifinius metaduomenų reikalavimus ir ribotas galimybes taikyti universalius sprendimus.

Ne mažiau svarbi diskusijų tema buvo tyrėjų bendruomenės pasirengimas. Dalyviai pažymėjo, kad viena iš esminių kliūčių išlieka ne tik infrastruktūros ar finansavimo stygius, bet ir pasitikėjimo, motyvacijos bei praktinio supratimo trūkumas. Buvo akcentuota, kad MTD valdymo pažanga priklausys ne vien nuo techninių sprendimų, bet ir nuo nuoseklaus kompetencijų stiprinimo, konsultacinės pagalbos bei atvirojo mokslo kultūros įtvirtinimo institucijose.

Galiausiai diskusijose išskirta doktorantų grupė kaip svarbi potencialių duomenų valdymo praktikų stiprinimo grandis. Pabrėžta, kad doktorantai gali būti vertinami ne tik kaip būsima mokymų auditorija, bet ir kaip aktyvūs atsakingo duomenų tvarkymo ir atvirojo mokslo kultūros formavimo dalyviai.

4.4. Apibendrinimas

Apibendrinant institucijų ir mokslininkų apklausų rezultatus, matyti, kad Lietuvos mokslo ir studijų sistemoje MTD kūrimas jau yra plačiai įsitvirtinusi mokslinės veiklos dalis, tačiau jų valdymo brandos lygis viso duomenų gyvavimo ciklo požiūriu išlieka netolygus. Institucijų lygmeniu MTD valdymas daugumoje institucijų turi bent tam tikrą norminį pagrindą, arba MTD klausimai yra įtraukti į kitus vidaus dokumentus, o pirmojoje apklausoje beveik visos institucijos nurodė, kad jose MTD kuriami ir valdomi. Vis dėlto šis institucinės brandos (vertinant plačiai per MTD prizmę) pagrindas nėra vienodai formalizuotas, jo turinys nevienodas, o praktinis įgyvendinimas ir stebėseną daugelyje vietų išlieka riboti. MTD Lietuvoje jau yra pripažįstami kaip svarbi mokslinės veiklos dalis, tačiau ne visur jie dar yra valdomi kaip savarankiška, nuosekliai institucionalizuota politikos ir mokslinės praktikos sritis.

Institucijų apklausoje ir atliktame tyrime išryškėjo diferenciacija pagal institucijų tipus. Universitetuose dažniau pastebimas brandesnis reguliacinis pagrindas, labiau išplėtotą infrastruktūrą, gausnesni žmogiškieji išteklių ir dažnesnės ir taip pat gausesnės metodinės pagalbos formos. Kolegijose ir dalyje mokslinių tyrimų institutų MTD funkcijos dažniau išlieka paskirstytos tarp bibliotekų, IT ar kitų padalinių darbuotojų ir jiems priskirtinų funkcijų, specialiai šioms veikloms skirtų pareigybių yra mažiau, o institucinė pagalba dažniau teikiama pagal poreikį, o ne kaip savalaikiai ir nuolat.

Technologinės infrastruktūros požiūriu Lietuvos MTD sistema pasižymi decentralizuota ir nevienodo brandumo sprendimų struktūra, kurioje skirtingų institucijų naudojami saugojimo ir prieigos modeliai ne visais atvejais yra pakankamai suderinti tarpusavyje. Dalis universitetų jau turi institucines saugyklas ir infrastruktūrą, leidžiančią ne tik saugoti duomenis, bet ir juos aprašyti metaduomenimis bei tam tikru mastu atverti. Tačiau bendra saugojimo praktikų struktūra rodo didelę įvairovę: greta institucinių saugyklų naudojami serveriai, debesijos sprendimai, išorinės ar tarptautinės saugyklos, o kai kuriais atvejais ir tyrėjų asmeniniai išteklių. Mokslo politikos požiūriu tokia situacija riboja galimybes valstybės ir pačių mokslo ir studijų institucijų lygmeniu užtikrinti patikimą MTD apskaitą, įvertinti jų mastą, pasiskirstymą, saugojimo vietas, ilgalaikio išsaugojimo poreikius ir atsarginių kopijų kūrimą bei su tuo susijusias sąnaudas.

Iki šiol sukauptų MTD masto analizė rodo, kad Lietuvos MTD kraštovaizdis yra labai nevienodas tiek pagal mokslo sritis, tiek pagal institucijų tipus. Didžiausi duomenų kiekiai ir apimtys fiksuojami mokslinių tyrimų institutuose gamtos, technologijų, medicinos ir žemės ūkio mokslų srityse; universitetai labai didelius MTD kiekius dažniau deklaruoja gamtos ir technologijų mokslų srityse, o kolegijose vyrauja palyginti nedideli duomenų mastai. Kartu svarbu pažymėti, kad dalyje mokslo (meno) sričių reikšminga dalis institucijų apskritai nurodo tokių duomenų neturinčios. Šie duomenys rodo ne tik nevienodą MTD intensyvumą, bet ir ribotas institucijų galimybes nuosekliai tvarkyti savo turimų duomenų išteklių apskaitą. Apklausos rezultatai leidžia matyti, kad institucijos ne visada gali tiksliai įvardyti, kiek ir kokių MTD yra sukaupusios, kokios jų bendros apimtys ir kur jie saugomi, o dalis duomenų tebėra laikoma mokslininkų kompiuteriuose ar pan. Tokia situacija apsunkina tiek institucinio, tiek nacionalinio lygmens MTD inventorizavimą ir riboja galimybes patikimai planuoti ilgalaikio išsaugojimo poreikius ypač tų duomenų, kurie pasižymi išliekamąja moksline verte. Papildomai pažymėtina, kad duomenys apie ilgalaikio saugojimo sąnaudas taip pat labai fragmentiški, todėl jų nepakanka platesniam nacionalinio masto vertinimui. Tai reiškia, kad Lietuvos MTD sistema negali būti vertinama kaip vienalytė nei pagal duomenų intensyvumą, nei pagal infrastruktūros poreikį, nei pagal pačių duomenų matomumą ir apskaitos brandą. Vienose srityse pagrindinis klausimas yra didelių mastų duomenų saugojimas, kitose – pats duomenų valdymo institucionalizavimas, jų apskaita, matomumas ir ilgalaikio išsaugojimo planavimas.

Mokslininkų apklausos rezultatai papildoma institucinį vaizdą parodydami, kad pagrindinė problema Lietuvos MTD sistemoje nėra susijusi su duomenų kūrimo apimtimi, bet su silpniau išplėtotomis vėlesnėmis duomenų gyvavimo ciklo grandimis. Nors didžioji dauguma tyrėjų nurodo kuriantys MTD, gerokai mažesnė jų dalis juos sistemingai tvarko, saugo, validuoja ar pateikia atviresnėmis dalijimosi formomis. Praktikoje dažniau vyrauja duomenų teikimas kaip publikacijų priedai, dalijimasis institucijos viduje arba tik gavus individualią užklausa, o naudojimasis atviromis duomenų saugyklomis dar nėra tapęs dominuojančia praktika. Kartu apklausos duomenys rodo, kad pažangesnės MTD valdymo praktikos tyrėjų lygmeniu dažniau siejasi su geresniu FAIR principų supratimu, duomenų valdymo planų rengimu, palankesne institucine aplinka, aiškesniu reguliavimu ir tarptautiškumu

patirtimi. Tai leidžia daryti išvadą, kad tyrėjų elgsena MTD valdymo srityje labai priklauso ne vien nuo individualių pasirinkimų, bet ir nuo institucinio bei platesnio mokslo politikos konteksto.

Ypač svarbus MTD metaduomenų matomumas. Institucijų apklausa rodo, kad daugelyje mokslo sričių MTD surandamumas pagal metaduomenis ir atviras metaduomenų prieinamumas išlieka riboti. Net tose srityse, kur situacija palyginti geresnė, brandesnės praktikos nėra dominuojančios. Tai reiškia, kad Lietuvos problema šia diena yra ne tai, kiek duomenų sukuriama, bet tai, kiek tie duomenys matomi, aprašyti, atsekami ir parengti būti pakartotinai naudojami.

Žmogiškųjų išteklių ir paskatų analizės pjūvis leidžia daryti dar vieną svarbią išvadą: Lietuvoje MTD valdymo motyvacinis ir organizacinis pagrindas kol kas daugiausia remiasi pavieniais stipresniais instituciniais branduoliais ir išoriniais reikalavimais, o ne visoje sistemoje nuosekliai išplėtotais pajėgumais. Tik nedidelėje dalyje institucijų MTD valdymas yra patikėtas atskiroms pareigybėms, o daug kur šios veiklos lieka fragmentiškai paskirstytos tarp kitų funkcijų. Be to, dauguma institucijų nenurodo finansinių ar karjeros paskatų MTD kuruoti, o informaciniai ar sklaidos renginiai apie MTD valdymą reguliariai vyksta tik dalyje dažniausiai universitetinio tipo institucijų. Šį institucinį vaizdą papildo ir mokslininkų apklausos duomenys: pradedantieji ir patvirtinti tyrėjai dažniau nurodo kompetencijų trūkumą, o pirmaujantieji tyrėjai – laiko ir finansavimo stoką.

Būtent šiame kontekste ypač svarbu vertinti Lietuvos jungimosi į EOSC perspektyvą. Pirmosios institucijų apklausos rezultatai rodo labai aukštą bendrą institucijų suinteresuotumą prisidėti prie Lietuvos MTD ekosistemos kūrimo ir integracijos į EOSC. Tik viena institucija tai vertino neigiamai, o likusios arba išreiškė aiškų norą prisidėti, arba tai laikė aktualiu pagal savo galimybes ir kompetencijas. Taigi politinės ir institucinės valios elementas iš esmės egzistuoja. Tačiau išsamesnė antrosios apklausos analizė rodo, kad ši valia dar nereiškia vienodo pasirengimo. Institucijos labai nevienodai disponuoja norminiu pagrindu, saugyklomis, metaduomenų praktika, personalu, paskatomis ir praktine pagalba. Todėl EOSC integracija Lietuvos atveju pirmiausia nėra vien techninis „prisijungimo“ klausimas ir juo labiau ne vien klausimas, ar kiekviena institucija turi savo atskirą saugyklą. Ji veikia kaip gebėjimo užtikrinti matomus, aprašytus, valdomus, sąveikius ir patikimai inventorizuojamus MTD.

Iš atlikto tyrimo ryškėja tai, kad Lietuvos jungimasis į EOSC turėtų būti suprantamas kaip kelių tarpusavyje susijusių lygmenų procesas. Pirmiausia, reikia stiprinti nacionalinį MTD matomumą: žinoti, kur ir kokie duomenys saugomi, kokių sričių duomenys dominuoja, kokios jų apimtys ir kokių metaduomenų apie juos yra. Antra, reikia stiprinti institucinį pagrindą: nuoseklesnį reglamentavimą, aiškesnius duomenų valdymo planų reikalavimus, archyvavimo ir sunaikinimo tvarkas, stebėseną, konsultacinę pagalbą ir dedikuotus žmogiškuosius išteklius. Trečia, reikia stiprinti atvirumo ir FAIR dimensiją: ne tik skelbti duomenis, kai tai įmanoma, bet pirmiausia sistemingai aprašyti juos metaduomenimis, padaryti surandamus ir sąveikius. Ketvirta, svarbus tarptautiškumo matmuo: tyrėjų apklausa rodo, kad dalyvaujantieji tarptautiniuose projektuose ir infrastruktūrose dažniau naudoja pažangesnes saugyklas ir dažniau taiko FAIR principus, taigi tarptautinė įtrauktis veikia kaip vienas iš brandesnių praktikų katalizatorių.

Todėl strateginiu požiūriu Lietuvos įsijungimas į EOSC neturėtų būti suprantamas kaip vienkartinis infrastruktūros projektas ar vienos technologinės platformos sukūrimas. Labiau tikėtina ir dokumento duomenimis pagrindžiama kryptis yra laipsniškas, federuotos logikos principais paremtas procesas, kuriame derinami skirtingi institucinių, disciplininių ir nacionalinio koordinavimo lygmenų sprendimai. Tokiam modeliui būtina ne absoliuti centralizacija, o koordinuota decentralizacija: bendri reikalavimai ir matomumas, bet skirtingi techniniai įgyvendinimo keliai pagal institucijų brandą, mokslo sričių specifiką ir jau turimą infrastruktūrą.

Apibendrinant galima teigti, kad Lietuvos MTD sistema šiuo metu nėra vienybė nei pagal institucinį pasirengimą, nei pagal valdymo brandą, nei pagal praktinį MTD tvarkymo, saugojimo ir atvėrimo išplėtojimą. Nors duomenų kūrimas Lietuvos mokslo ir studijų sistemoje jau yra plačiai įsitvirtinęs, jų matomumas, aprašomumas, atsekamumas, atvirumas, valdymo nuoseklumas ir ilgalaikio išsaugojimo planavimas vis dar turi ryškių skirtumų tarp institucijų ir mokslo sričių. Ypač svarbu tai, kad iki šiol MTD Lietuvoje tvarkomi taikant skirtingus metaduomenų standartus ir nevienodas aprašymo praktikas, todėl duomenų surandamumas, sąveikumas ir galimybės juos sistemingai integruoti platesnėje ekosistemoje išlieka ribotos. Būtent šios aplinkybės yra esminės vertinant Lietuvos pasirengimą jungtis į Europos atvirojo mokslo debesį (EOSC), nes toks jungimasis reikalauja ne vien pavienių technologinių sprendimų, bet ir semantiškai bei architektūriškai suderintos aplinkos, leidžiančios duomenis aprašyti, susieti, rasti ir panaudoti tarp skirtingų sistemų bei institucijų.

Atsižvelgiant į empirinio tyrimo rezultatus, šios studijos užsakymo pagrindu sukurti Lietuvai skirti architektūriniai, semantiniai, jungimosi į EOSC ir institucinio valdymo modeliai bei scenarijai, orientuoti į identifikuotų MTD netolygumo, riboto matomumo, nevienodų metaduomenų aprašymo praktikų, infrastruktūros išvystymo skirtumą ir nevienodo institucinio pasirengimo problemų sprendimą. Šie modeliai ir scenarijai apima ne tik technines jungimosi į EOSC prielaidas, bet ir nacionalinio modelio, sudarančio sąlygas derinti skirtingų institucijų kuriamų ir saugomų MTD architektūrinius, semantinius ir valdymo principus, formavimą. Lietuvai siekiant visavertiško įsijungimo į EOSC, esminiu uždaviniu tampa ne vien MTD kūrimas, bet ir jau kuriamų duomenų aprašymo suderintais principais, jų matomumo, sąveikumo ir tinkamumo veikti federuotoje Europos MTD erdvėje užtikrinimas. Išsamiau architektūriniai modeliai, jungimosi į EOSC ir institucinio valdymo scenarijai pristatomi studijos 5 skyriuje.

5. SSGG analizė ir Lietuvos įsijungimo į EOSC alternatyvos ir jų scenarijai

Šiame skyriuje analizuojamos Lietuvos įsijungimo į EOSC stiprybės, silpnybės, galimybės ir grėsmės (SSGG) bei trys galimos Lietuvos įsijungimo į EOSC alternatyvos: decentralizuotas *status quo* modelis, nacionalinio mazgo sukūrimo alternatyva ir regioninio mazgo sukūrimo alternatyva. Toks analitinis pjūvis leidžia ne tik palyginti skirtingo masto intervencijas, bet ir įvertinti, koku mastu kiekviena iš jų prisidėtų prie nacionalinės mokslinių duomenų erdvės formavimo, Lietuvos mokslo duomenų matomumo didinimo, duomenų pakartotinio naudojimo skatinimo bei pasirengimo visavertei integracijai į EOSC federaciją. Pasirinktas alternatyvų išdėstymas atitinka klasikinę galimybių studijos logiką, pagal kurią *status quo* alternatyva naudojama kaip analitinis atskaitos taškas, leidžiantis nuosekliai įvertinti kitų alternatyvų kuriamą pridėtinę vertę tiek Lietuvos mokslo sistemai, tiek jos integracijai į Europos atvirojo mokslo infrastruktūrą, jų institucinį ir technologinį įgyvendinamumą bei ilgalaikį strateginį poveikį. Alternatyvos diferencijuojamos pagal intervencijos apimtį, koordinavimo sudėtingumą ir jų potencialą kurti nacionalinę pridėtinę vertę mokslo duomenų valdymo ir atvirojo mokslo ekosistemos plėtros požiūriu.

5.1. SSGG analizė

Remiantis išsamia esamos nacionalinės situacijos analize, tarptautine gerąja praktika, ES MTD valdymo strateginėmis nuostatomis ir teisiniu reguliavimu, surinktais MSI ir mokslininkų apklausų duomenimis ir jų palyginimu bei dviejų fokusuotųjų grupių diskusijomis, atliktas sisteminis Lietuvos situacijos įvertinimas ir palyginimas su EOSC koncepcija, FAIR principais ir EOSC-IF idealiojo modelio reikalavimais. Remiantis šia medžiaga identifikuotos Lietuvos įsijungimo į EOSC stiprybės, silpnybės, galimybės ir grėsmės.

18 lentelė. Lietuvos įsijungimo į EOSC stiprybės, silpnybės, galimybės ir grėsmės

Stiprybės	Silpnybės
<ul style="list-style-type: none"> Lietuvos MTI dalyvavimas tarptautinėse ESFRI ERIC infrastruktūrose, skatinantis augimą. 	<ul style="list-style-type: none"> Nepakankamas ir fragmentuotas finansavimas.
<ul style="list-style-type: none"> Geroji praktika, kylanti iš ES programų, apimanti MTD valdyseną ir tvarkyseną. 	<ul style="list-style-type: none"> Nepakankamas atvertų MTD kiekis.
<ul style="list-style-type: none"> Pavienės iniciatyvos skaitmeninimo srityje, stiprinančios MTD valdymo kultūrą ir siekias atverti MTD. 	<ul style="list-style-type: none"> Koordinuoto MTD valdymo modelio ir politikos nebuvimas.

• Plačiai įsitvirtinęs MTD kūrimas.	• Nepakankamai išvystyta MSI saugyklų sistema ir fragmentuoti bei nesąveikūs MTD.
• Sukuriami dideli MTD kiekiai.	• MTD valdymo kompetencijų ir kvalifikuotų specialistų trūkumas.
• Didelė MSI motyvacija integruotis į EOSC.	• Netolygiai išvystyta atskirų MSI infrastruktūra ir MTD valdymo praktika.
• FAIR principų taikymas kaip ryškus MTD atvėrimo pažangos veiksnys.	• FAIR principai dar nėra tapę visų MSI praktiniu MTD valdymo standartu.
• Palanki Lietuvos teisinė bazė.	• Teisinis poįstatyminių aktų neapibrėžtumas.
Galimybės	Grėsmės
• Galimybė Lietuvai dalyvauti EOSC veikloje.	• Nestabili geopolitinė situacija ir jos keliamos grėsmės valstybės saugumui.
• Prieiga prie platesnės Europos MTD ir paslaugų ekosistemos ir MTD mainai.	• Tvaraus finansavimo neapibrėžtumas.
• Galimybė vystyti AM praktikas Lietuvoje.	• Prastėjanti demografinė situacija dėl gimstamumo ir protų nutekėjimo.
• DI raidose pažanga.	• Nepakeliama administracinė, atitikties ir teisinio reguliavimo našta mažesnėms MSI.
• AM pagrindu plėtojama mokslinių tyrimų kultūra, etika, piliečių mokslas, pasitikėjimas mokslu.	• Fragmentuoti, neapskaityti, nevaldomi, nerandami ir todėl pakartotinai nepanaudojami Lietuvos MSI rengti MTD.
• Ekonominė nauda valstybei, verslui ir MSI.	• Sąveikumo reikalavimų sudėtingumas ir jų dažna kaita.
• Platesnės mokslinio bendradarbiavimo galimybės.	• Didėjantis atotrūkis tarp MTD kūrimo masto ir jų valdymo brandumo nacionaliniu ir instituciniu lygmenimis.
• Nacionalinio duomenų saugumo ir suvereniteto stiprinimas.	• Nacionalinė politika ir teisinė bazė, nemotyvuojanti mokslininkų kurti ir atverti savo MTD taip dalyvaujant AM ir MTD ekonomikos ekosistemoje.

Remiantis šiuo SSGG vertinimu suformuluotos trys Lietuvos integracijos į EOSC alternatyvos ir jų scenarijai (žr. 5.2– 5.4 poskyrius).

5.2. Pirmoji alternatyva: status quo – decentralizuotas modelis

Analizuojant Lietuvos įsiliejimo į EOSC ekosistemą galimybes, pirmiausia būtina įvertinti decentralizuotą *status quo* modelį. Kaip jau buvo identifikuota GS kontekste, EOSC architektūra numato tris fundamentalius integracijos kelius: per nacionalinį, regioninį arba specifinį teminį mazgą. Pastaroji galimybė – teminis mazgas – gali būti suprantamas dvejopai: kaip individuali Europos MTI (e. infrastruktūra) arba kaip jungtinis teminių infrastruktūrų klasteris. Be tiesioginio dalyvavimo formų, federacijos valdysenos modelis taip pat numato lanksčią galimybę prie jau funkcionuojančių mazgų jungtis netiesiogiai, įgyjant trečiosios šalies statusą. Visi šie teminių mazgų konstravimo atvejai reprezentuoja decentralizuotą požiūrį, kuriam būdingas institucinis autonomiškumas nuo centrinių nacionalinių ar regioninių valdymo struktūrų. Pagrindiniu šio modelio techniniu ir organizaciniu pagrindu laikomos ESFRI ERIC tipo mokslinių tyrimų infrastruktūros. Pirmojo EOSC plėtros etapo patirtis rodo, kad į šią federaciją sėkmingai integravosi tokios tarptautinės infrastruktūros, kaip ELIXIR, „Euro-BioImaging“, „Instruct - ERIC“ ir EMBL, o CLARIN ERIC bei keletas kitų subjektų šiuo metu baigia pasirengimo procesus ateinantiems integracijos etapams.

Lietuvos įsitraukimas į šiuos procesus pasižymi istoriniu nuoseklumu – šalies mokslo bendruomenė ir atsakingos institucijos dar nuo XXI a. pirmojo dešimtmečio aktyviai dalyvauja pasirengimo etapuose įsiliejant į ES MTI veiklą. Šis sistemingas įdirbis atsispindi 2011–2024 metų laikotarpiu parengtuose keturiuose nacionaliniuose MTI kelrodžiuose. Strateginių dokumentų raida indikuoja ne tik dinamišką Lietuvos infrastruktūrų skaičiaus augimą, bet ir kokybinę jų integraciją į bendrą ES e. infrastruktūrų erdvę. Šią tendenciją patvirtina ir atlikta kiekybinė analizė: nacionaliniuose Kelrodžiuose duomenys mokslo infrastruktūrų kontekste minimi net 192 kartus, o šie paminėjimai apima visas mokslo sritis bei skirtingus infrastruktūrų tipologinius segmentus. Tai liudija apie suformuotą platų teminį tinklą, kuris yra kaip pagrindas decentralizuotam patekimo į EOSC keliui.

Vertinant šios alternatyvos įgyvendinamumo potencialą, išskiriami keli esminiai privalumai, identifikuojami esamoje MTEPI aplinkoje. Visų pirma, fiksuojamas palankus strateginis kontekstas: šalyje yra parengta MTI raidai tinkama teisinė bazė, patvirtinti plėtrai būtini strateginiai dokumentai, stebimas nuoseklus mokslinio personalo kompetencijų augimas mokslo institucijose (žr. Kelrodis, p. 33 (Lietuvos mokslo taryba, 2024b)). Atskiru sėkmės rodikliu laikomas „labai geras kai kurių MTI bendradarbiavimas su verslo sektoriumi bei išvystyti tvarūs globalūs ryšiai su tarptautinėmis mokslo infrastruktūromis“ (ten pat). Taip pat svarbiu privalumu laikytina Kelrodyje (p. 21) numatyta finansavimo priemonių sinergijos galimybė, kuri leidžia skirtingas MTI funkcijas ir veiklas finansuoti derinant įvairius šaltinius.

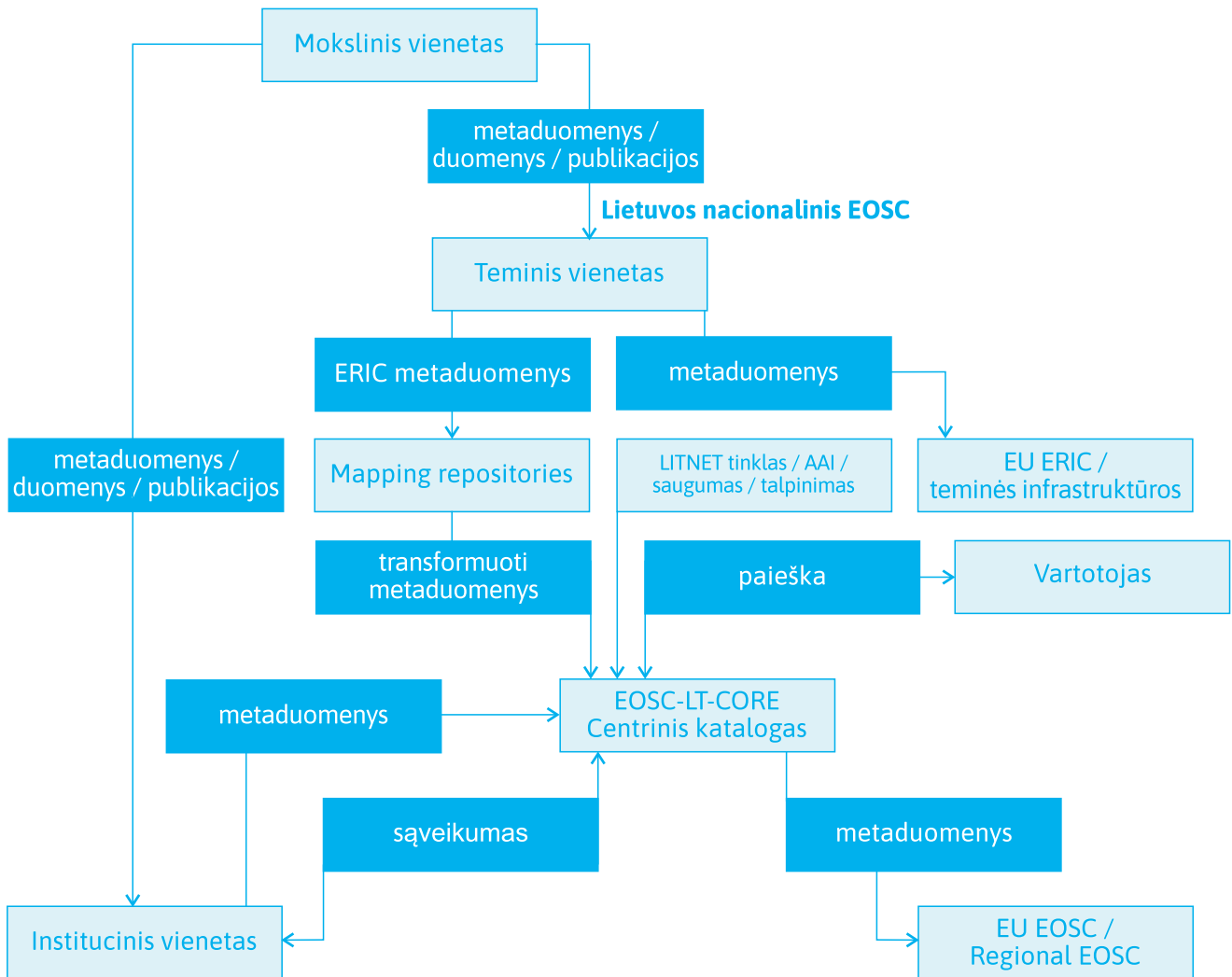
Visgi, analizuojant *status quo* scenarijų, būtina įvertinti ir egzistuojančias ekosistemos spragas (žr. Kelrodis, p. 40). Šiuo metu identifikuojamas palyginti lėtas atvirosios prieigos prie duomenų centrų formavimasis, o duomenų bei metaduomenų apdorojimo standartizavimo, sistemų suderinamumo ir dalijimosi duomenimis praktikos išlieka fragmentuotos. Šio modelio logika diktuoja, kad šie iššūkiai galėtų būti sprendžiami ne kuriant naujus institucinius darinius, o pasitelkiant papildomą tikslinį finansavimą esamoms infrastruktūroms, ypač jų duomenų saugykloms. Tokiu būdu racionaliausia būtų koncentruotis į egzistuojančių MTI veiklos sričių stiprinimą, užtikrinant efektyvų Lietuvoje sukuriamų duomenų srautą į EOSC bei teikiant metodinę ir organizacinę paramą nacionalinėms MTI jų integracijos procese.

5.3. Antroji alternatyva: nacionalinis EOSC mazgas

Aprašas

Nacionalinio EOSC mazgo sukūrimo alternatyva vertintina kaip strategiškai nuoseklus ir su dabartine EOSC plėtros logika suderintas kelias, nes pati EOSC vystoma kaip federacinė, iš tarpusavyje sąveikių mazgų sudaryta infrastruktūra. Oficialiuose EOSC dokumentuose nurodoma, kad būsimoji EOSC federacija bus grindžiama daugelio mazgų sąveika, o šie mazgai atliks pagrindinių prieigos taškų funkciją naudotojams, paslaugoms ir skaitmeniniams ištekliams. EOSC EU mazgas šiuo požiūriu laikytinas referenciniu įgyvendinimo modeliu, apibrėžiančiu sąveikumo principus nacionaliniams, regioniniams ir teminiams mazgams. Todėl nacionalinio mazgo kūrimas suprastinas ne kaip nuo Europos atskirta nacionalinė iniciatyva, bet kaip struktūriškai suderintas žingsnis, leidžiantis Lietuvai kartu stiprinti nacionalinę MTD erdvę ir pasirengti visavertei integracijai į EOSC federaciją. Šioje alternatyvoje išskiriami trys organizacinio valdymo scenarijai, kuriems bendra esminė mazgo architektūra, tačiau jie skiriasi valdymo, koordinavimo ir institucinės atsakomybės pasiskirstymu.

Nacionalinio EOSC-LT mazgo architektūros logika ir struktūriniai komponentai



6 pav. Nacionalinio EOSC-LT mazgo architektūros struktūriniai komponentai

Nacionalinio EOSC-LT mazgo architektūra grindžiama federuotu modeliu, kuriame skirtingi MTD kūrimo, saugojimo ir sklaidos komponentai jungiami į vientisą nacionalinę sąveikumo aplinką. Šio modelio paskirtis – sudaryti prielaidas Lietuvos mastu koordinuotai MTD erdvei, kuri leistų ne tik užtikrinti efektyvesnį duomenų matomumą, paiešką ir pakartotinį naudojimą nacionaliniu lygmeniu, bet ir pasirengti nuosekliai integracijai į besiformuojančią EOSC federaciją. Architektūra remiasi principu, kad duomenys ir metaduomenys išlieka kuo arčiau jų kilmės vietos, o nacionalinis mazgas užtikrina jų atrandamumą, sąveikumą ir koordinuotą prieigą.

Pagrindinis šios architektūros elementas yra EOSC-LT-Core – centrinis nacionalinio mazgo komponentas, kurio branduolį sudaro nacionalinis centrinis MTD metaduomenų katalogas. Šis komponentas atlieka metaduomenų agregavimo, katalogavimo, sąveikumo koordinavimo ir centralizuotų paslaugų teikimo funkcijas. Jo paskirtis – sudaryti bendrą nacionalinį prieigos tašką prie Lietuvos MTD išteklių, užtikrinti skirtingų saugyklų ir paslaugų tarpusavio suderinamumą, taip pat sudaryti technines ir organizacines prielaidas autentifikavimo, prieigos valdymo ir kitoms horizontaliojo pobūdžio paslaugoms veikti. Kartu EOSC-LT-Core veikia kaip pagrindinis Lietuvos sąveikos su EU EOSC vartų taškas, per kurį į europinę federaciją perduodami nacionaliniu lygmeniu agreguoti ir semantiškai sąveikūs metaduomenys. Būtent šis komponentas sudaro nacionalinio mazgo branduolį, nuo kurio priklauso visa tolesnė sąveikumo architektūra. Architektūroje išskiriami trys pagrindiniai duomenų ir metaduomenų šaltinių tipai. Pirmasis yra mokslinis vienetas, kuris suprantamas kaip pirminis MTD ir susijusių metaduomenų generatorius – atskiras mokslininkas, tyrimų grupė, projektas ar kita mokslo veiklą vykdanči struktūra. Šiuo lygmeniu sukuriama duomenys, publikacijos ir juos apibūdinantys metaduomenys, kurie vėliau perduodami į vieną iš nacionalinio mazgo priimančiųjų komponentų, priklausomai nuo jų pobūdžio, institucinės aplinkos ir taikytino domeninio standarto.

Antrasis šaltinių tipas yra institucinis vienetas, apimantis institucijų valdomas ar centralizuotai palaikomas MTD saugyklas. Šis komponentas skirtas toms mokslo ir studijų institucijoms, kurios jau turi arba gali turėti savo institucinę saugojimo ir metaduomenų valdymo infrastruktūrą. Instituciniai vienetai nacionalinio mazgo architektūroje veikia kaip federuoti partneriai, kurie išlaiko tam tikrą autonomiją, tačiau savo metaduomenis teikia į centrinį katalogą ir taip tampa nacionalinės mokslinių duomenų erdvės dalimi.

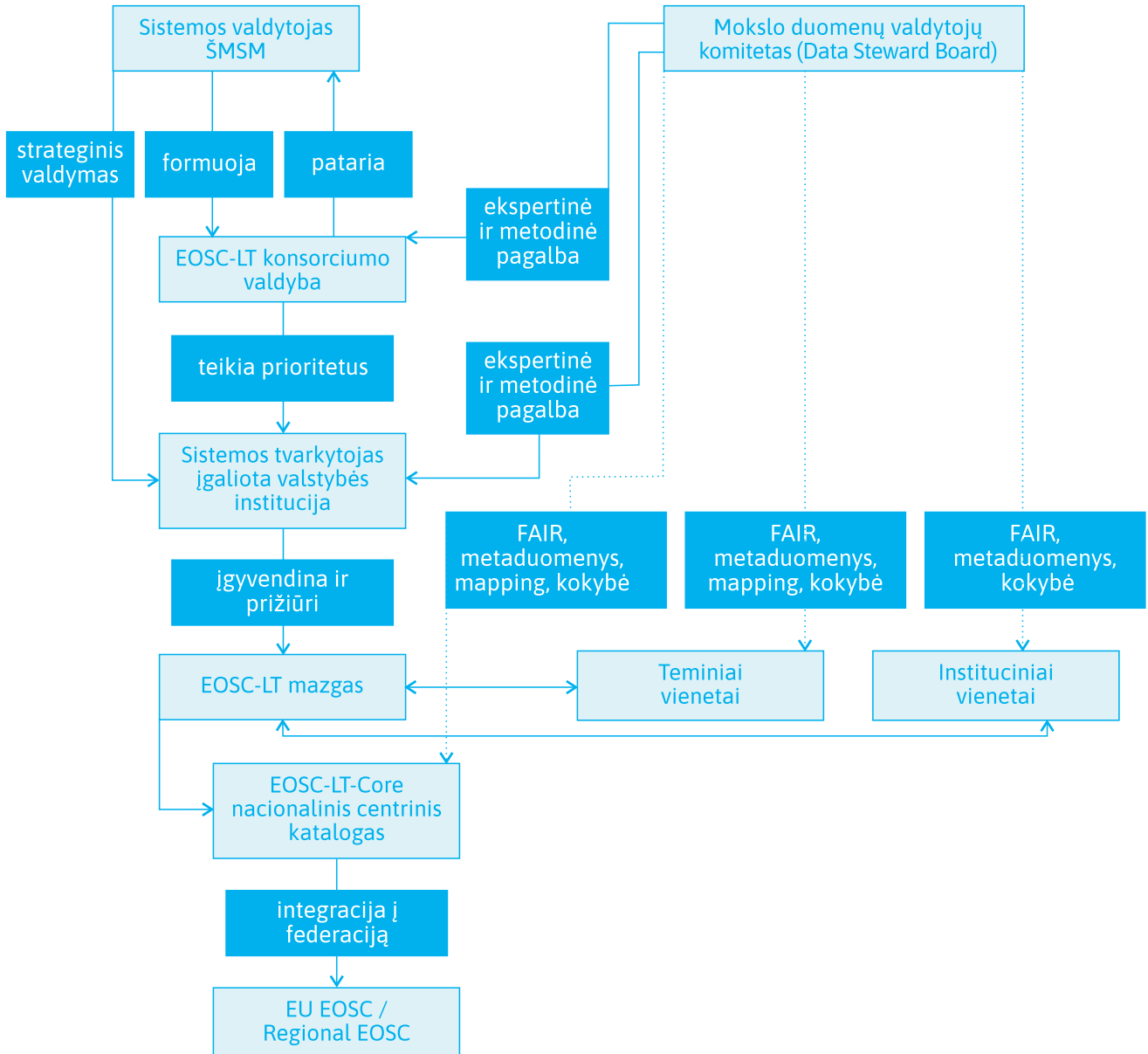
Trečiasis šaltinių tipas yra teminis vienetas, apimantis domenesines ar temines saugyklas, kurios paprastai jau yra įsijungusios į europines mokslinių infrastruktūrų ekosistemas, pavyzdžiui, ERIC tipo infrastruktūras. Šie vienetai dažniausiai yra technologiškai brandesni ir veikia pagal konkrečiai mokslo sričiai pritaikytus metaduomenų modelius bei standartus. Dėl šios priežasties jie atlieka svarbų vaidmenį tarptautinėje MTD apykaitoje, tačiau kartu kelia didesnius sąveikumo reikalavimus nacionaliniu lygmeniu, nes jų naudojami metaduomenų modeliai ne visada tiesiogiai suderinami su bendru nacionalinio mazgo katalogu.

Būtent šiam iššūkiui spręsti architektūroje numatomi metaduomenų susiejimo ir transformavimo sprendiniai (angl. *mapping repositories*). Jų paskirtis – automatiškai ar pusiau automatiškai transformuoti teminių vienetų naudojamus metaduomenų modelius į bendrą nacionaliniam mazgui taikomą sąveikumo modelį. Tokiu būdu užtikrinama, kad teminiai vienetai galėtų išlaikyti suderinamumą su jų tarptautinėmis domeninėmis infrastruktūromis, tačiau kartu taptų visaverte nacionalinės MTD erdvės dalimi. Šis komponentas yra esminis EOSC-IF principais grindžiamo sąveikumo architektūros elementas.

Horizontalų techninį pagrindą šiai architektūrai gali sudaryti LITNET infrastruktūra, kuri užtikrina akademinio tinklo junglumą, saugumo ir autentifikavimo paslaugas be prireikus technines prielaidas atskiriems mazgo komponentams pateikti. Tokiu būdu LITNET veikia ne kaip savarankiškas duomenų šaltinis, bet kaip techninis įgalintojas, palaikantis nacionalinio mazgo veikimą ir saugų skirtingų institucinių infrastruktūrų sujungimą į bendrą akademinį tinklą.

Architektūros logika remiasi tuo, kad MTD ir metaduomenys į nacionalinį mazgą patenka keliais kanalais – per institucinius vienetus arba teminius vienetus. EOSC-LT-Core šiuos srautus agreguoja, transformuoja į sąveikią metaduomenų visumą ir formuoja nacionalinį centrinį katalogą, kuris tampa pagrindiniu prieigos, paieškos ir matomumo tašku nacionaliniams naudotojams. Per šį komponentą taip pat vyksta Lietuvos įsijungimas į EU EOSC, nes būtent jis sudaro technines ir organizacines prielaidas perduoti suderintus metaduomenis į aukštesnio lygmens europinius ar regioninius mazgus. Tokia architektūra leidžia vienu metu spręsti dvi strategines užduotis: kurti veikiančią nacionalinę MTD erdvę Lietuvoje ir užtikrinti struktūriškai suderintą pasirengimą visavertei integracijai į EOSC federaciją.

5.3.1. 1 scenarijus. Nacionalinio EOSC-LT mazgo kaip valstybės informacinės sistemos scenarijus



7 pav. Nacionalinis EOSC-LT mazgas kaip valstybės informacinė sistema

Šio scenarijaus esmė – nacionalinė MTD erdvė ir ją įgalinantis EOSC-LT mazgas organizuojami vadovaujantis valstybės informacinių išteklių valdymo logika, o pagrindinis mazgo branduolys – EOSC-LT-Core – realizuojamas kaip nacionalinės svarbos valstybės informacinės infrastruktūros komponentas. Šiame modelyje EOSC-LT-Core kartu veikia kaip nacionalinis centrinis MTD metaduomenų katalogas ir pagrindinis Lietuvos integracijos į EU EOSC taškas.

Scenarijus numato aiškų funkcijų pasiskirstymą tarp strateginio ir operacinio lygmens. Politinio ir strateginio valdymo funkcija tenka ŠMSM, o techninio įgyvendinimo, priežiūros ir eksploataavimo funkcija pavedama sistemos tvarkytojui – LMT, Valstybės skaitmeninių sprendimų agentūrai ar kitai tam tinkamą kompetenciją ir mandatą turinčiai valstybės institucijai. Toks modelis sudaro prielaidas EOSC-LT mazgą institucionalizuoti kaip ilgalaikę valstybės skaitmeninę infrastruktūrą, užtikrinant aiškią atsakomybę už jo veikimą ir tęstinumą po projekto pabaigos.

Strateginiam, ekspertiniam ir bendruomeniniam dalyvavimui užtikrinti prie šio modelio formuojama EOSC-LT konsorciumo valdyba, vienijanti į nacionalinį mazgą įsijungusių MSI atstovus. Ši taryba veikia kaip patariamasis ir koordinuojantis organas, padedantis derinti nacionalinius prioritetus, institucinius poreikius ir mazgo plėtros kryptis.

Toks modelis leistų nacionalinę MTD erdvę įtvirtinti ne kaip laikiną projektinę iniciatyvą, bet kaip ilgalaikę valstybės informacinės infrastruktūros dalį. Tai reikšminga keliais aspektais. Pirma, toks institucinės valdytos modelis didintų tęstinumą ir tvarumą po projekcinio finansavimo pabaigos. Antra, jis sudarytų prielaidas nacionalinę MTD erdvę susieti su kitais valstybės informaciniais ištekliais ir duomenų erdvėmis, taip stiprinant horizontalų sąveikumą nacionaliniu mastu. Trečia, šis modelis leistų EOSC-LT mazgą traktuoti kaip nacionalinės svarbos skaitmeninę infrastruktūrą, o ne tik kaip atskiro sektoriaus iniciatyvą.

Mokslo duomenų valdytojų taryba pagal šį scenarijų veikia ne tik kaip centrinis patariamasis organas, bet ir kaip metodinis tinklas, kurio praktinės funkcijos realizuojamos EOSC-LT-Core, teminiuose vienetuose, instituciniuose vienetuose ir bendroje nacionalinėje MTD saugykloje. Ji apima tiek centrinį mokslo duomenų valdytojų forumą / komitetą (angl. *Data Steward Board*), tiek paskirtus ar funkcinio pagrindu veikiančius duomenų valdytojus EOSC-LT-Core, teminiuose ir instituciniuose vienetuose. Šis sluoksnis atsakingas už praktinį FAIR principų, metaduomenų standartų, duomenų aprašymo, kokybės užtikrinimo ir sąveikumo reikalavimų įgyvendinimą. Kartu duomenų valdytojų taryba veikia kaip ekspertinis patariamasis mechanizmas sistemos valdytoji ir EOSC-LT konsorciumo tarybai, užtikrinantis, kad strateginiai sprendimai būtų grindžiami realia duomenų valdymo praktika ir instituciniu grįžtamoju ryšiu.

Valdymo požiūriu šis scenarijus remiasi trimis tarpusavyje susietais lygmenimis: politiniu-strateginiu, instituciniu-operaciniu ir ekspertiniu-praktiniu. Politinį ir strateginį lygmenį užtikrina sistemos valdytojas, operacinį ir techninį – sistemos tvarkytojas, o ekspertinį bei metodinį – EOSC-LT konsorciumo taryba ir mokslo duomenų valdytojų tinklas. Tokia daugiasluoksnė valdymo schema sudaro prielaidas ne tik technologiškai įveikinti nacionalinį EOSC mazgą, bet ir užtikrinti jo kokybišką veikimą, priimtinumą mokslo bendruomenei bei ilgalaikį institucionalizavimą.

Papildomą šio scenarijaus stiprybę sudaro galimybė EOSC-LT-Core komponentą pateikti valstybės lygmens debesų kompiuterijos infrastruktūroje – „GovCloud“ aplinkoje. Toks sprendinys sudarytų prielaidas užtikrinti tvarų techninį palaikymą, nuolatinę stebėseną, saugų atsarginių kopijų kūrimą, patikimą incidentų valdymą ir saugius tarpinius ryšius su kitomis valstybės bei mokslo infrastruktūromis. Tai sustiprintų EOSC-LT mazgo, kaip nacionalinės svarbos skaitmeninės infrastruktūros, institucionalizavimą ir sumažintų techninio tęstinumo rizikas po projekcinio finansavimo pabaigos.

Vis dėlto tokio modelio įgyvendinimas suponuoja aiškią institucinę prielaidą: institucija, kuriai ateityje būtų pavestos sistemos tvarkytojo funkcijos, projekte turėtų dalyvauti partnerio teisėmis ir būti atsakinga už EOSC-LT-Core komponento realizavimą. Tik toks sprendinys sudarytų prielaidas užtikrinti sklandų perėjimą iš projekto į nuolatinę eksploataciją, aiškią atsakomybę už techninę priežiūrą ir realų sistemos tvarumą po projekto pabaigos.

Taigi šis scenarijus yra labiausiai orientuotas į nacionalinį tvarumą, formalų institucinį įtvirtinimą ir integraciją su kitomis valstybės duomenų valdymo iniciatyvomis. Jo stiprybė – aiški atsakomybių struktūra, tęstinumo garantijos ir galimybė nacionalinę mokslo duomenų erdvę traktuoti kaip strateginę nacionalinės svarbos skaitmeninę infrastruktūrą.

1 scenarijaus įgyvendinimo planas, etapai ir darbų seka

Pirmojo scenarijaus įgyvendinimas turėtų būti organizuojamas kaip nuoseklus trejų metų procesas, kuriame nacionalinis EOSC-LT mazgas būtų kuriamas ne kaip vienkartinis techninis sprendinys, o kaip palaipsniui institucionalizuojama nacionalinės svarbos skaitmeninė infrastruktūra. Šio proceso logika turėtų remtis trimis nuosekliais etapais: pirmaisiais metais sukuriama teisinė, organizacinė, architektūrinė ir metodinė prielaidos, antraisiais metais vyksta techninis ir organizacinis mazgo įveiklinimas bei bandomasis veikimas, o trečiaisiais metais užtikrinama plėtra, eksploatacinė branda ir tvarus perėjimas į nuolatinę veiklą.

PIRMAISIAIS METAIS pagrindinis dėmesys turėtų būti skiriamas pasirengti, pamatinėms struktūroms sukurti ir metodiniam sluoksniui suformuoti. Šiame etape turi būti galutinai apibrėžta valdysenos schema: sistemos valdytojo ir tvarkytojo vaidmenys, EOSC-LT konsorciumo tarybos funkcijos, mokslo duomenų valdytojų tarybos vieta bendroje valdymo struktūroje, partnerių atsakomybės ir sprendimų priėmimo tvarka. Kartu turi būti parengtas detalus EOSC-LT-Core architektūrinis modelis, apimantis nacionalinį centrinį MTD metaduomenų katalogą, autentifikavimo ir prieigos valdymo mechanizmus, sąveikumo sprendinius, stebėsenos ir saugumo sluoksnius, taip pat principinius ryšius su instituciniais, teminiais vienetais.

Tuo pat metu pirmaisiais metais turi būti atlikta esamų institucijų ir teminių saugyklų, metaduomenų modelių, FAIR brandos ir sąveikumo situacijos analizė, identifikuoti potencialūs federacijos dalyviai, prioritetiniai integravimo atvejai ir svarbiausi metodiniai skirtumai tarp skirtingų duomenų šaltinių. Ypač svarbu, kad būtent šiame etape būtų sukurtas mokslo duomenų valdytojų kompetencijų modelis, apibrėžiantis jų vaidmenį, atsakomybes ir reikalingas kompetencijas tiek centrinio mazgo lygmeniu, tiek instituciniuose ir teminiuose vienetuose. Remiantis šiuo modeliu, turi būti parengtos mokymo programos, metodinės gairės ir praktiniai darbo standartai FAIR principų, metaduomenų kokybės, duomenų aprašymo, atitikčių sudarymo (angl. *mapping*), duomenų publikavimo ir sąveikumo srityse. Per pirmuosius metus taip pat turi būti įgyvendinta pirmoji duomenų valdytojų mokymų iteracija, kad nuo antrųjų metų pradžios apmokyti duomenų valdytojų tinklas jau galėtų realiai įsijungti į projekto veiklas institucijose ir centriniame mazge. Pirmųjų metų pabaigoje turi būti parengti techniniai ir organizaciniai specifikacijų dokumentai, suderintas bazinis nacionalinis metaduomenų modelis, apibrėžti atitikčių sudarymo principai ir pradėtas EOSC-LT-Core realizavimas.

ANTRAISIAIS METAIS projektas turėtų pereiti į aktyvų įgyvendinimo ir bandomojo įveiklinimo etapą. Šiuo laikotarpiu turi būti sukurtas ir įdiegtas EOSC-LT-Core branduolys, įskaitant nacionalinį centrinį MTD metaduomenų katalogą, pagrindines paieškos ir atrandamumo funkcijas, sąveikumo mechanizmus ir pirmąsias technines integracijas su institucinėmis bei teminėmis saugyklomis. Turi būti kuriami, testuojami ir tobulinami metaduomenų transformavimo ir susiejimo sprendiniai, leidžiantys skirtingus duomenų šaltinius įtraukti į bendrą nacionalinį katalogą išlaikant jų domeninę specifiką.

Šiame etape jau visavertiškai turi pradėti veikti mokslo duomenų valdytojų tinklas. Pirmaisiais metais parengti ir apmokyti duomenų valdytojai nuo antrųjų metų turi įsijungti į praktinį projekto vykdymą EOSC-LT-Core, instituciniuose vienetuose, teminiuose vienetuose ir bendroje nacionalinėje mokslinių duomenų saugykloje. Jie turėtų vykdyti FAIR principų diegimo, metaduomenų kokybės, duomenų aprašymo, sąveikumo, atitikčių nustatymo, konsultavimo ir metodinės pagalbos funkcijas, kartu užtikrindami, kad techninis mazgo kūrimas būtų lydimas realios duomenų valdymo praktikos. Antraisiais metais taip pat turi vykti papildomi tiksliniai mokymai, mentorystė ir kompetencijų stiprinimas, grindžiami pirmųjų bandinių patirtimi. Antrųjų metų pabaigoje nacionalinis mazgas jau turėtų veikti bandomuoju režimu, apimdamas bent dalį institucinių, teminių ir bendros nacionalinės saugyklos srautų, o duomenų valdytojų tinklas turėtų būti tapęs realiai veikiančiu nacionalinio mazgo metodiniu sluoksniu.

TREČIAISIAIS METAIS prioritetas turėtų būti teikiamas apimties plėtrai, veiklos stabilizavimui ir pasirengimui pereiti į nuolatinę eksploataciją. Šiame etape turi būti plečiamas integruotų institucinių ir teminių vienetų skaičius, tobulinamas nacionalinio katalogo funkcionalumas, stiprinamos saugumo, stebėsenos, incidentų valdymo, atsarginių kopijų ir veiklos tęstinumo priemonės. Kartu turi būti toliau brandinamas duomenų valdytojų tinklas kaip nuolatinė nacionalinio mazgo veikimo dalis: plečiama jo aprėptis, tikslinamos metodinės gairės, atnaujinamos mokymo programos ir užtikrinamas naujų institucijų įtraukimas į bendrą duomenų valdymo praktiką.

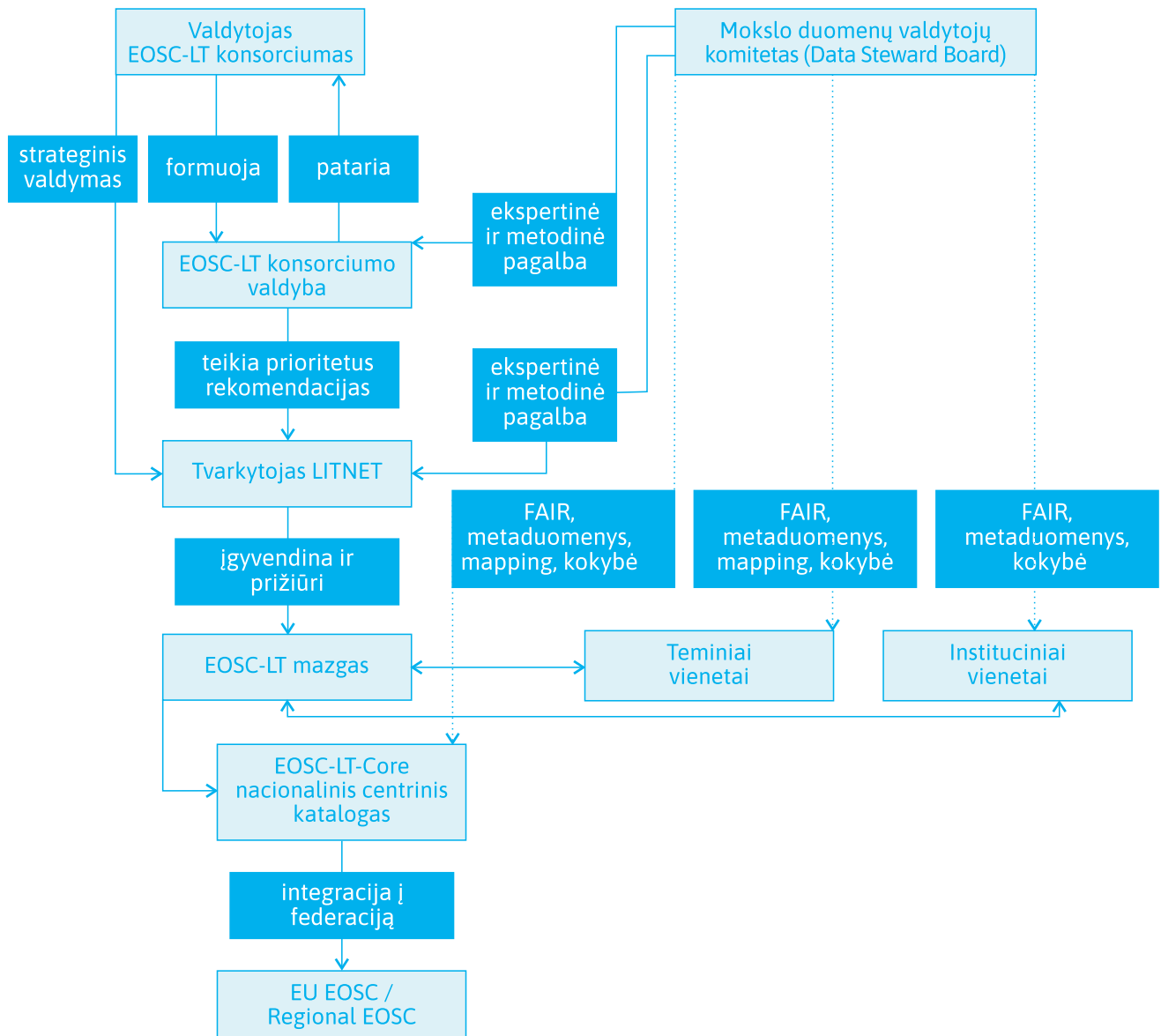
Trečiaisiais metais taip pat turi būti užtikrintas visapusiškesnis Lietuvos pasirengimas integracijai į EU EOSC vartų logiką, sudarant prielaidas nuosekliam nacionaliniu lygmeniu agreguotų ir semantiškai suderintų metaduomenų perdavimui į aukštesnio lygmens europinę federaciją. Kartu turi būti įtvirtintas ilgalaikis veiklos modelis po projekto pabaigos: patvirtintos eksploataavimo procedūros, atsakomybės, finansavimo ir techninės priežiūros schema, taip pat užtikrintas perėjimas iš projekto režimo į nuolatinės valstybės informacinės infrastruktūros veikimo režimą. Šio etapo pabaigoje EOSC-LT mazgas jau turėtų veikti ne kaip bandomasis sprendinys, o kaip realiai eksploatuojamas nacionalinis MTD prieigos, atrandamumo ir sąveikumo mazgas, turintis ir techninį branduolį, ir veikiantį institucinį bei metodinį duomenų valdytojų sluoksnį.

Bendra darbų seka šiame scenarijuje turėtų būti tokia: pirmiausia sukuriama valdysena, architektūrinis ir metodinis pagrindas, kartu su duomenų valdytojų kompetencijų modeliu ir pirmąja mokymų iteracija; paskui įgyvendinamas techninis branduolys ir pirmosios integracijos, o apmokytas duomenų valdytojų tinklas įtraukiamas į praktinį mazgo veikimą; galiausiai plečiamas federacinis tinklas, stiprinamas eksploatacinis stabilumas ir užtikrinamas ilgalaikis nacionalinio mazgo institucionalizavimas po projekto pabaigos. Tokia dinamika atitinka šio scenarijaus esmę: nacionalinis EOSC-LT mazgas kuriamas ne kaip laikina projektinė platforma, bet kaip ilgalaikė, valstybės lygmeniu įtvirtinta ir metodiniu, techniniu bei organizaciniu požiūriu brandi nacionalinės svarbos skaitmeninė infrastruktūra.

19 lentelė. 1 scenarijaus įgyvendinimo planas, etapai ir darbų seka

Laikotarpis	Pagrindiniai darbai	Pagrindinis rezultatas
1 metai	Sukuriama valdysenos schema; paskiriami valdytojo ir tvarkytojo vaidmenys; parengiama EOSC-LT-Core architektūra; atliekama institucinių ir teminių vienetų analizė; suderinamas bazinis metaduomenų ir sąveikumo modelis; sukuriamas mokslo duomenų valdytojų kompetencijų modelis; parengiamos mokymo programos ir įgyvendinama pirmoji mokymų iteracija	Sukurtas organizacinis, architektūrinis ir metodinis pagrindas; parengtas valdytojų tinklas praktiniam įsitraukimui nuo 2 metų
2 metai	Sukuriamas ir bandomuoju režimu įveiklinamas EOSC-LT-Core; diegiamos pirmosios integracijos su instituciniais ir teminiais vienetais; kuriami ir testuojami metaduomenų atitikčių sprendiniai; valdytojai įsijungia į FAIR, metaduomenų kokybės, duomenų aprašymo, sąveikumo ir konsultavimo veiklas; vyksta papildomi tiksliniai mokymai	Veikiantis bandomasis nacionalinis mazgas ir praktikoje veikiantis valdytojų tinklas
3 metai	Plečiamas integruotų institucinių ir teminių vienetų skaičius; tobulinamas katalogo funkcionalumas; stiprinamos stebėsenos, saugumo ir tęstinumo priemonės; brandinamas valdytojų tinklas; įtvirtinamos eksploataavimo, finansavimo ir atsakomybės procedūros; užtikrinamas perėjimas iš projekto į nuolatinę veiklą	Brandus, eksploatuojamas nacionalinis EOSC-LT mazgas, pasirengęs ilgalaikiai veiklai ir integracijai į EOSC federaciją

5.3.2. 2 scenarijus. Konsorciumo ir akademinio tinklo pagrindu organizuojamo EOSC-LT mazgo scenarijus (sektorinė mokslo duomenų erdvė)



8 pav. Konsorciumo ir akademinio tinklo pagrindu organizuojamas EOSC-LT mazgas

Šio scenarijaus esmė – nacionalinė MTD erdvė ir ją įgalinantis EOSC-LT mazgas organizuojami ne kaip valstybės informacinės sistemos tipo infrastruktūra, bet kaip sektorinė MTD erdvė, grindžiama EOSC-LT konsorciumo ir Lietuvos akademinio tinklo LITNET partneryste. Šiame modelyje pagrindinis mazgo branduolys – EOSC-LT-Core – kartu veikia kaip nacionalinis centrinis MTD metaduomenų katalogas ir pagrindinis Lietuvos integracijos į EU

EOSC taškas, tačiau jo valdymo ir eksploataavimo logika pirmiausia siejama su akademinio sektoriaus saviorganizacija ir infrastruktūriniu bendradarbiavimu.

Scenarijus numato, kad strateginio valdymo funkciją vykdytų EOSC-LT konsorciumas, vienijantis į nacionalinį mazgą įsitraukusias MSI, įskaitant ir pagrindinius viešojo sektoriaus mokslo politikos veikėjus. Operacinio ir techninio lygmens atsakomybė šiame modelyje būtų pavedama LITNET, kuris yra Lietuvos MSI kompiuterių tinklas, kurio paslaugomis naudojasi daugiau kaip 1000 institucijų, o tinklo bei paslaugų palaikymą ir plėtrą vykdo techniniai centrai penkiuose Lietuvos universitetuose. LITNET taip pat yra pasaulinės mokslo ir studijų tinklų infrastruktūros GÉANT dalis ir užtikrina tarptautinius sujungimus su kitų šalių akademiniais tinklais bei MTEPI infrastruktūromis.

Tokia organizacinė schema turi aiškią logiką akademinio sektoriaus požiūriu. Kadangi LITNET jau teikia akademinėi bendruomenei tinklo, junglumo, tapatybių ir kitų bendrųjų skaitmeninių paslaugų sluoksnį, EOSC-LT-Core ir bendra nacionalinė MTD saugykla šiame scenarijuje galėtų būti natūraliau integruojami į esamą akademinio tinklo infrastruktūrą. Tai potencialiai supaprastintų techninį diegimą, palengvintų saugių ryšių organizavimą ir leistų efektyviau panaudoti LITNET sukauptą infrastruktūrinę bei organizacinę patirtį.

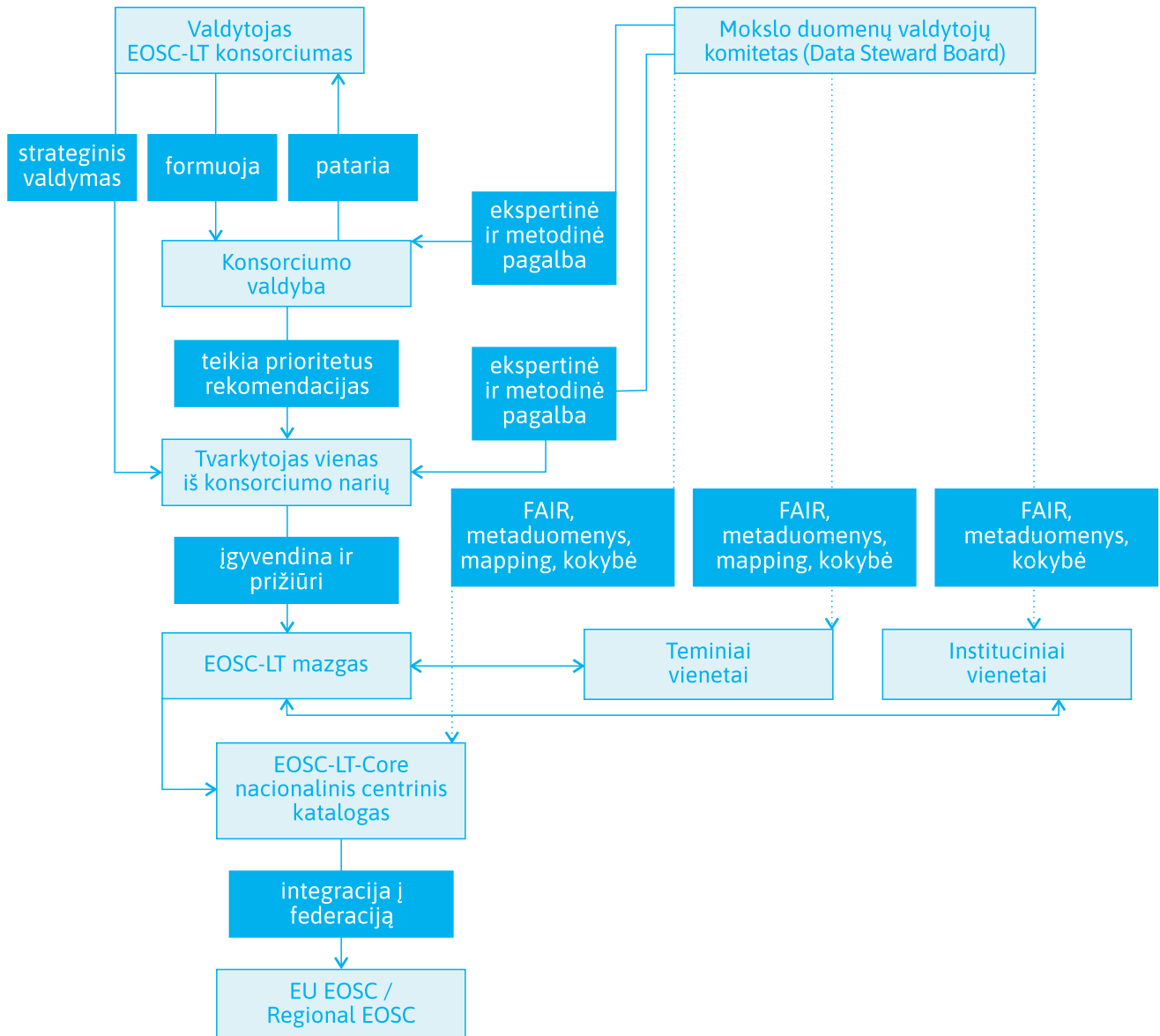
Mokslo duomenų valdytojų tinklas šiame scenarijuje išlieka ne mažiau svarbus negu pirmajame variante ir iš esmės atlieka tą pačią funkciją. Praktinės FAIR principų, metaduomenų kokybės, duomenų aprašymo, atitikčių sudarymo ir sąveikumo palaikymo funkcijos turėtų būti realizuojamos EOSC-LT-Core, teminiuose vienetuose, instituciniuose vienetuose ir bendroje nacionalinėje MTD saugykloje. Kaip ir pirmajame scenarijuje, centrinis mokslo duomenų valdytojų forumas / komitetas šiuo atveju veiktų kaip ekspertinis ir metodinis tinklas, jungiantis paskirtus ar funkciniu pagrindu veikiančius duomenų valdytojus atskiruose mazgo komponentuose. Skirtumas būtų ne funkcinis šio sluoksnio vaidmuo, bet jo koordinavimo logika. Šiame modelyje duomenų valdytojų tinklas būtų labiau siejamas su konsorciumo ir akademinio tinklo organizacine logika, o ne su valstybės informacinės sistemos valdysena.

Vis dėlto, lyginant su pirmuoju scenarijumi, šis modelis pasižymi silpnesniu formaliu instituciniu įtvirtinimu. Pirmasis scenarijus remiasi valstybės informacinių išteklių valdymo logika, kuri įstatymu apibrėžia valstybės informacinių išteklių rūšis, jų politiką formuojančias ir įgyvendinančias institucijas bei valdytojo ir tvarkytojo vaidmenis. LITNET, nors ir labai svarbus nacionalinės akademinės infrastruktūros operatorius, savo prigimtimi yra orientuotas į akademinio sektoriaus tinklo ir paslaugų veikimą, o ne į nacionalinės svarbos valstybės informacinės sistemos tipo valdyseną. Dėl to šis scenarijus būtų stiprus techninės integracijos ir sektoriaus autonomijos požiūriu, tačiau silpnesnis ilgalaikio formalizuoto tęstinumo, valstybės lygmens atsakomybių įtvirtinimo ir horizontalios integracijos su kitomis nacionalinėmis duomenų erdvėmis požiūriu. Papildomai pažymėtina, kad šiame scenarijuje ne taip aiškiai apibrėžtas ilgalaikio finansavimo modelis po projekto pabaigos. Tai reiškia, kad mazgo veiklos tęstinumas didesniu mastu priklausytų nuo vėlesnių konsorciumo susitarimų ir finansavimo sprendimų, o ne nuo iš anksto labiau institucionalizuoto valstybės lygmens tęstinumo modelio kaip pirmajame scenarijuje.

Svarbi šio scenarijaus stiprybė yra būtent didesnė akademinio sektoriaus autonomija. Valdymas čia būtų arčiau pačios mokslo ir studijų bendruomenės, o infrastruktūriniai sprendimai – glaudžiau susieti su jau veikiančiu akademinio tinklo ekosistemos modeliu. Tokia schema gali būti patraukli tuo, kad leidžia greičiau mobilizuoti sektorines kompetencijas ir mažiau priklauso nuo bendros valstybės informacinių sistemų institucionalizavimo procedūros. Kita vertus, ši stiprybė kartu reiškia ir silpnesnę valstybės lygmens tęstinumo garantiją. Po projekto pabaigos scenarijaus tvarumas labiau priklausytų nuo konsorciumo susitarimų, LITNET organizacinio pajėgumo ir sektoriaus vidaus koordinavimo, o ne nuo griežtesnio valstybės informacinės infrastruktūros režimo.

Taigi šis scenarijus laikytinas vienu iš galimų nacionalinio EOSC-LT mazgo organizavimo būdų. Jis yra techniškai pagrįstas, gerai dera su akademinio tinklo infrastruktūra ir išlaiko didesnę sektoriaus autonomiją. Vis dėlto, vertinant pagal institucinį tvarumą, formalų atsakomybių įtvirtinimą, tęstinumo garantijas ir sąveiką su kitomis valstybės duomenų valdymo iniciatyvomis, šis scenarijus laikytinas mažiau optimaliu už pirmąjį.

5.3.3. 3 scenarijus. Konsorciumo savivaldos pagrindu organizuojamo EOSC-LT mazgo, kaip autonominės sektorinės mokslo duomenų erdvės, scenarijus



9 pav. Konsorciumo savivaldos pagrindu organizuojamas EOSC-LT mazgas

Šio scenarijaus esmė – nacionalinė MTD erdvė ir ją įgalinantis EOSC-LT mazgas organizuojami kaip autonominė sektorinė MTD erdvė, grindžiama pačių MSI savivalda, be tiesioginio valstybės institucijų dalyvavimo mazgo valdysenoje. Šiame modelyje pagrindinis mazgo branduolys – EOSC-LT-Core – kartu veikia kaip nacionalinis

centrinis MTD metaduomenų katalogas ir pagrindinis Lietuvos integracijos į EU EOSC taškas, tačiau jo valdymas, eksploatavimas ir plėtra būtų grindžiami tik konsorciumo narių susitarimu ir jų prisiimama atsakomybe.

Scenarijus numato, kad valdytojo funkciją vykdytų pats EOSC-LT konsorciumas, vienijantis į mazgą įsitraukusias MSI. Operacinio ir techninio lygmens atsakomybė būtų pavedama vienam iš konsorciumo narių, turinčiam pakankamą techninį, organizacinį ir projektinį pajėgumą įgyvendinti ir prižiūrėti EOSC-LT-Core komponentą. Toks modelis reikštų, kad mazgo įveiklinimas, techninė priežiūra ir organizacinė plėtra būtų vykdomi visiškai sektoriinės saviorganizacijos pagrindu.

Šio scenarijaus stiprybė – didžiausia akademinio sektoriaus autonomija. Valdymas čia būtų maksimaliai priartintas prie pačios mokslo bendruomenės, o sprendimai dėl infrastruktūros vystymo, prioritetų, paslaugų ir investicijų būtų priimami pačių konsorciumo narių. Tokia schema leistų formuoti išgrynintą sektoriinę MTD erdvę, kurioje institucijos turėtų didžiausią įtaką mazgo plėtros kryptims ir organizacinei logikai. Šis modelis taip pat galėtų būti patrauklus dėl didesnio lankstumo, spartesnio sprendimų priėmimo akademiame rate ir mažesnės priklausomybės nuo valstybės lygmens administracinių procedūrų.

Mokslo duomenų valdytojų tinklas šiame scenarijuje išlieka ne mažiau svarbus negu pirmais dviem variantais ir iš esmės atlieka tą pačią funkciją. Praktinės FAIR principų, metaduomenų kokybės, duomenų aprašymo, atitikčių sudarymo ir sąveikumo palaikymo funkcijos turėtų būti realizuojamos EOSC-LT-Core, teminiuose vienetuose, instituciniuose vienetuose ir bendroje nacionalinėje mokslinių duomenų saugykloje. Kaip ir kituose scenarijuose, centrinis mokslo duomenų valdytojų forumas / taryba veiktų kaip ekspertinis ir metodinis tinklas, jungiantis paskirtus ar funkcinio pagrindu veikiančius valdytojus atskiruose mazgo komponentuose. Esminis skirtumas slypėtų ne šio sluoksnio funkcijoje, bet jo institucinėje koordinavimo logikoje – ji būtų grindžiama vien konsorciumo savivalda ir sektoriiniu susitarimu.

Vis dėlto, lyginant su pirmuoju ir antruoju scenarijais, šis modelis pasižymi silpniausiu formaliu instituciniu įtvirtinimu. Kadangi mazgo valdysenoje tiesiogiai nedalyvautų ministerija ar kita valstybės institucija, nacionalinės MTD erdvės tęstinumas būtų grindžiamas ne valstybės lygmens institucionalizavimu, o pačių konsorciumo narių tarpusavio susitarimais ir jų gebėjimu ilgai išlaikyti bendrą organizacinę valią. Dėl to šis scenarijus būtų stiprus autonomijos ir sektoriinės savivaldos požiūriu, tačiau silpnesnis formalios atsakomybės, ilgalaikio tęstinumo ir horizontalios integracijos su kitomis nacionalinėmis duomenų valdymo iniciatyvomis požiūriu.

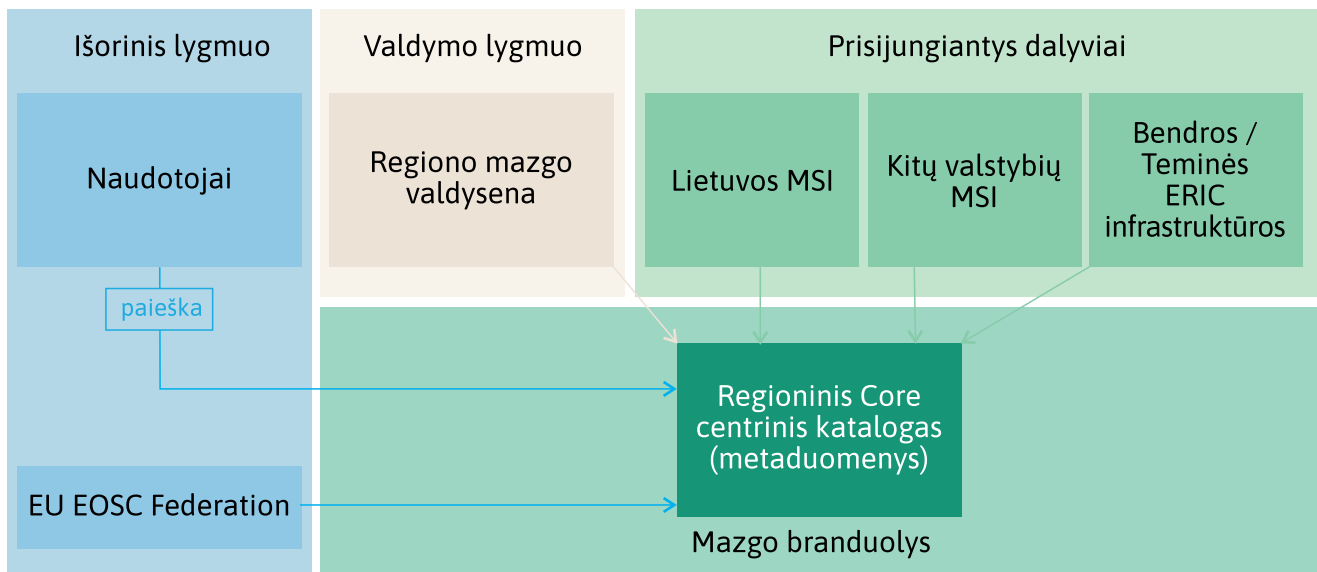
Papildomai pažymėtina, kad šiame scenarijuje ne taip aiškiai apibrėžtas ir ilgalaikio finansavimo modelis po projekto pabaigos. Mazgo veiklos tęstinumas čia dar didesniu mastu priklausytų nuo būsimų konsorciumo narių susitarimų, institucinių prioritetų ir projektinių ar kitų finansavimo sprendimų. Tai reiškia, kad nors šis modelis suteiktų didžiausią veikimo laisvę mokslo sektoriui, kartu jis kurtų ir didžiausią riziką dėl finansinio bei organizacinio tvarumo ilguoju laikotarpiu.

Taigi šis scenarijus laikytinas autonomiškiausiu, tačiau kartu ir instituciniu požiūriu mažiausiai tviriu nacionalinio EOSC-LT mazgo organizavimo būdu. Jis geriausiai atitinka išgrynintos sektoriinės mokslo duomenų erdvės logiką, tačiau, vertinant pagal formalų atsakomybių įtvirtinimą, finansavimo tęstinumo aiškumą, ilgalaikį stabilumą ir integraciją su kitomis valstybės duomenų valdymo iniciatyvomis, yra silpniausias iš nagrinėjamų scenarijų.

5.3.4. Galutinė išvada dėl antros alternatyvos scenarijų

Apibendrinant tris nacionalinio EOSC-LT mazgo organizavimo scenarijus, darytina išvada, kad tinkamiausias yra pirmasis scenarijus – nacionalinio EOSC-LT mazgo kaip valstybės informacinės sistemos scenarijus. Nors antrasis ir trečiasis scenarijai suteikia daugiau sektoriinės autonomijos ir tam tikrais aspektais gali būti lankstesni organizaciniu požiūriu, jie pasižymi silpnesniu formaliu instituciniu įtvirtinimu, mažiau aiškiu ilgalaikio finansavimo modeliu ir didesne tęstinumo rizika po projekto pabaigos. Pirmasis scenarijus geriausiai atitinka nacionalinio tvarumo, aiškios atsakomybės, institucionalizavimo ir integracijos su kitomis valstybės duomenų valdymo iniciatyvomis kriterijus. Būtent šis scenarijus sudaro tvirčiausias prielaidas nacionalinį EOSC-LT mazgą plėtoti ne kaip laikiną projektinę ar sektoriinę iniciatyvą, bet kaip ilgalaikę nacionalinės svarbos skaitmeninę infrastruktūrą, užtikrinančią ir veikiančią nacionalinę MTD erdvę Lietuvoje, ir nuoseklų pasirengimą visavertei integracijai į EOSC federaciją.

5.4. Trečioji alternatyva: regioninis EOSC mazgas



10 pav. Regioninis EOSC mazgas

Šios alternatyvos esmė – Lietuva į EOSC integruotųsi ne per savarankiškai organizuojamą nacionalinį mazgą, o per bendrą kelių valstybių regioninį mazgą. Toks modelis konceptualiai atitinka EOSC federacijos architektūrą, nes EOSC dokumentuose numatyta, kad mazgai gali atstovauti ne tik nacionalinėms ar teminėms, bet ir regioninėms bendruomenėms. Regioninio mazgo logika reiškia, kad kelių valstybių institucijos ir infrastruktūros būtų jungiamos į bendrą koordinavimo, katalogavimo ir sąveikumo sluoksnį, kuris veiktų kaip bendras regioninis įėjimo taškas į EOSC federaciją.

Tokio modelio teorinis privalumas – galimybė iš karto kurti platesnio masto tarpvalstybinę MTD erdvę, dalytis techninėmis ir organizacinėmis sąnaudomis, derinti regioninius standartizavimo sprendimus ir jungti kelių šalių mokslo bendruomenes į bendrą sąveikią aplinką. Tam tikrais atvejais tai galėtų leisti sumažinti dubliavimąsi, sukurti bendrus infrastruktūros pajėgumus ir padidinti regiono kaip visumos matomumą EOSC federacijoje. Tokia alternatyva teoriškai galėtų būti patraukli mažesnėms valstybėms, jei jos turėtų aiškią bendrą politinę valią, sutartą valdymo modelį ir gana panašų brandos lygį mokslo duomenų valdymo srityje. Ši kryptis minima ir platesnėse EOSC diskusijose apie mazgais grįstą federacijos plėtrą.

Vis dėlto, vertinant Lietuvos situaciją, ši alternatyva pasižymi reikšmingais apribojimais. Pirma, ji tik iš dalies sprendžia pagrindinę problemą, dėl kurios apskritai svarstomas nacionalinio mazgo kūrimas, t. y. nacionalinės MTD erdvės stoką Lietuvoje. Regioninis mazgas visų pirma orientuotas į išorinę tarpvalstybinę integraciją, o ne į vidinę nacionalinę konsolidaciją. Todėl tokia alternatyva silpniau prisidėtų prie Lietuvos vidinio MTD sujungimo, bendro nacionalinio katalogavimo ir nuoseklaus sąveikumo modelio formavimo nacionaliniu mastu. Kitaip tariant, regioninis modelis labiau stiprina dalyvavimą platesnėje federacijoje, bet mažiau tiesiogiai kuria Lietuvos vidinę MTD erdvę.

Antra, regioninio mazgo atveju mažėja tiesioginė nacionalinė pridėtinė vertė. Jei centrinis koordinavimo ir katalogavimo taškas būtų bendras kelioms valstybėms, dalis investicijų, organizacinės energijos ir sukuriama vertė būtų nukreipta ne į Lietuvos nacionalinės infrastruktūros stiprinimą, bet į bendrą regioninį centrą. Tokiu atveju Lietuva neišvengiamai taptų labiau priklausoma nuo to, kaip regioniniu lygiu bus nustatomi prioritetai, pa-

skirstomi išteklių ir organizuojamas techninis bei institucinės raidos tempas. Tai reikštų, kad Lietuva nebeturėtų tokio pat lygio tiesioginės kontrolės, kokią turėtų nacionalinio mazgo atveju.

Trečia, šis modelis didintų priklausomybę nuo regioninio centrinio taško. Praktiniu požiūriu tai reikštų, kad Lietuvos prieiga prie EOSC, metaduomenų publikavimas, katalogavimo logika ar net kai kurios sprendimų priėmimo procedūros būtų bent iš dalies priklausomos nuo regioninio mazgo valdymo centro ir su juo susijusių susitarimų. Jei regioninis centras būtų fiziškai ar instituciškai įtvirtintas kitoje valstybėje, Lietuvos galimybės savarankiškai formuoti savo prioritetus, techninius sprendimus ar plėtros tempą būtų ribotesnės. Tai nėra vien teorinė rizika: EOSC viešojoje medžiagoje regioninis bendradarbiavimas pristatomas kaip perspektyvus, tačiau kartu matyti, kad mažesnės šalys regioninio mazgo idėją sieja su papildomais finansavimo ir juridinės struktūros iššūkiais.

Ketvirta, regioninio mazgo alternatyva pasižymi didesne valdysenos ir įgyvendinimo sudėtingumu rizika. Skirtingai nuo nacionalinio mazgo, kuriam pakanka vienos valstybės institucinės valios ir koordinavimo, regioniniam modeliui reikėtų kelių šalių suderintų sprendimų dėl valdymo, atsakomybių, finansavimo, techninės architektūros, standartų ir ilgalaikės eksploatacijos. Tokia alternatyva suponuoja ne tik techninį, bet ir politinį bei teisinį suderinamumą. Būtent šie iššūkiai minimi ir regioninio bendradarbiavimo kontekste Šiaurės Makedonijos atveju, kur regioninio mazgo idėja siejama su neaiškumais dėl finansavimo bei regioninės juridinės struktūros.

Penkta, praktinių precedentų požiūriu ši alternatyva yra silpniau pagrįsta negu nacionalinio mazgo modelis. EOSC *Federation Handbook* ir kiti EOSC dokumentai regioninį mazgo aiškiai pripažįsta kaip galimą kategoriją, tačiau praktikoje gerokai ryškiau matomi nacionalinių ir teminių mazgų pavyzdžiai. Lenkija yra ypač svarbus lyginamasis atvejis, nes ji aiškiai kuria savo nacionalinį mazgo, koordinuojamą remiantis nacionaline struktūra, o ne bendrą regioninį sprendinį. Tai rodo, kad net regiono mastu nacionalinio mazgo kelias šiuo metu laikomas praktiškiau ir labiau institucionalizuotu.

Ši alternatyva taip pat kelia klausimų dėl finansinio ir organizacinio tęstinumo. Jeigu regioninis mazgas būtų grindžiamas kelių valstybių bendradarbiavimu, ilgalaikis jo išlaikymas priklausytų nuo nuolatinio kelių šalių finansinio įsipareigojimo, bendros politinės valios ir gebėjimo išlaikyti regioninį koordinavimo centrą. Tokia schema yra jautresnė išoriniams pokyčiams. Vienos šalies prioritetų pasikeitimas, finansavimo sumažėjimas ar valdysenos nesutarimai galėtų neproporcingai paveikti ir Lietuvos galimybes naudotis regioniniu mazgo. Dėl šios priežasties Lietuvos požiūriu tai būtų viena rizikingiausių alternatyvų tvarumo prasme. Ši rizika atitinka ir viešai aptariamą regioninio bendradarbiavimo problemas EOSC aplinkoje.

Taigi regioninio EOSC mazgo alternatyva yra metodologiškai galima ir su EOSC federacijos architektūra suderinama, tačiau Lietuvos atveju ji vertintina kaip mažiau palanki nei nacionalinio mazgo kūrimo alternatyva. Jos stiprybė – potenciali tarpvalstybinė sinergija ir sąnaudų dalijimasis. Tačiau esminės silpnybės yra mažesnė tiesioginė nacionalinė nauda, didesnė priklausomybė nuo regioninio koordinavimo centro, sudėtingesnė valdysena, neaiškesnis finansinis tęstinumas ir tai, kad ši alternatyva silpniau sprendžia pačią pagrindinę Lietuvos problemą – nacionalinės mokslo duomenų erdvės stygių. Dėl to ši alternatyva laikytina labiau teoriškai įmanoma nei praktiškai optimali.

5.5. Alternatyvų palyginimo išvados

Atlikta analizė rodo, kad Lietuvos įsijungimo į EOSC alternatyvos skiriasi ne tik intervencijos mastu, bet ir jų gebėjimu vienu metu spręsti du strateginius uždavinius: 1) formuoti nacionalinę MTD erdvę Lietuvoje; 2) užtikrinti pasirengimą visavertei integracijai į besiformuojančią EOSC federaciją. Alternatyvų palyginimas leidžia daryti išvadą, kad didžiausią pridėtinę vertę Lietuvai sukurtų nacionalinio mazgo alternatyva, nes ji geriausiai suderina nacionalinį interesą, sąveikumą su EOSC, institucinį tvarumą ir praktinį įgyvendinamumą.

Status quo alternatyva, grindžiama decentralizuotu atskirų institucijų dalyvavimu, leidžia išlaikyti esamą padėtį ir nereikalauja papildomos institucinės intervencijos, tačiau ji neišsprendžia esminės problemos – nacionalinės MTD erdvės stokos. Toks modelis išlaiko fragmentuotą duomenų matomumą, nekuria bendro naciona-

linio katalogavimo sluoksnio ir ribotai prisideda prie nuoseklaus Lietuvos pasirengimo EOSC federacijai. Dėl to *status quo* alternatyva vertintina labiau kaip atskaitos taškas, o ne kaip strategiškai pakankamas sprendimas.

Regioninio mazgo alternatyva yra konceptualiai suderinama su EOSC federacijos architektūra ir teoriškai galėtų suteikti tarpvalstybinės sinergijos bei sąnaudų dalijimosi naudą. Vis dėlto Lietuvos atveju ši alternatyva yra mažiau palanki, nes ji silpniau prisideda prie nacionalinės MTD erdvės formavimo, mažina tiesioginę nacionalinę pridėtinę vertę, didina priklausomybę nuo regioninio koordinavimo centro ir suponuoja sudėtingesnę valdyseną bei neaiškesnį ilgalaikio finansavimo modelį. Be to, praktinių precedentų požiūriu regioninio mazgo modelis šiuo metu yra silpniau pagrįstas negu nacionalinių mazgų kelias. Todėl ši alternatyva laikytina metodologiškai galima, tačiau strategiškai ir praktiškai mažiau tinkama Lietuvai.

Nacionalinio mazgo alternatyva šioje analizėje išsiskiria kaip proporcingiausias ir strategiškai nuosekliausias sprendimas. Ji leidžia sukurti Lietuvoje veikiančią nacionalinę MTD erdvę, užtikrina bendrą metaduomenų katalogavimo ir sąveikumo sluoksnį, stiprina MTD matomumą bei pakartotinį naudojimą ir kartu sudaro struktūriškai suderintas prielaidas integracijai į EOSC federaciją. Šios alternatyvos stiprybė ta, kad jos kuriama vertė išlieka reikšminga net ir tuo atveju, jei visa integracija į EOSC federaciją vyktų palaipsniui. Lietuva bet kuriuo atveju įgytų nacionaliniu mastu funkcionuojantį, pagal EOSC sąveikumo principus organizuotą MTD katalogą ir federuotą prieigos modelį.

Vertinant vidinius nacionalinio mazgo alternatyvos scenarijus, optimaliu laikytinas pirmasis scenarijus – nacionalinio EOSC-LT mazgo kaip valstybės informacinės sistemos scenarijus. Šis scenarijus geriausiai užtikrina formalų institucinį įtvirtinimą, aiškią atsakomybių struktūrą, techninio ir finansinio tęstinumo prielaidas bei galimybę nacionalinę MTD erdvę susieti su kitomis valstybės duomenų valdymo iniciatyvomis. Antrasis ir trečiasis scenarijai suteikia daugiau sektorinės autonomijos, tačiau yra silpnesni ilgalaikio tvarumo, finansavimo aiškumo ir valstybės lygmens institucionalizavimo požiūriu.

Atsižvelgiant į tai, galimybių studijoje rekomenduotina rinktis nacionalinio mazgo alternatyvą, o jai įgyvendinti prioritetu laikyti pirmąjį organizacinį ir valdymo scenarijų, grindžiamą valstybės informacinių išteklių valdymo logika. Būtent ši kryptis geriausiai atitinka Lietuvos poreikį vienu metu stiprinti nacionalinę mokslo duomenų erdvę, užtikrinti ilgalaikį institucinį tvarumą ir pasirengti visavertei integracijai į EOSC federaciją.

20 lentelė. Lietuvos įsijungimo į EOSC alternatyvų daugiakriteris įvertinimas

Kriterijus	Status quo – decentralizuotas modelis	Nacionalinis mazgas	Regioninis mazgas
Pagrindinė logika	Tęsiamas decentralizuotas atskirų institucijų dalyvavimas esamose saugyklose ir tarptautiniuose tinkluose	Kuriamas nacionalinis mazgas su centriniu katalogu ir federuotu institucinių bei teminių vienetų tinklu	Lietuva integruojasi per bendrą kelių valstybių regioninį mazgą
Nacionalinės mokslo duomenų erdvės kūrimas	Nesukuriama	Sukuriama visavertė nacionalinė mokslo duomenų erdvė	Kuriama tik iš dalies, nes centrinis koordinavimo sluoksnis perkeliamas į regioninį lygmenį
Nacionalinė pridėtinė vertė	Maža	Didelė	Vidutinė arba ribota
Integracija su EOSC federacija	Fragmentuota, atskirais MTI prisijungimais	Struktūriškai parengta ir nuosekliai organizuota per nacionalinį centrinį tašką	Priklausoma nuo regioninio mazgo brandos, valdysenos ir susitarimų
Strateginė kontrolė	Išskaidyta tarp MTI	Nacionalinė	Regioninė / pasidalinta
Valdysenos aiškumas	Ribotas	Aukštas	Vidutinis arba žemas
Institucinis įtvirtinimas	Silpnas	Stiprus	Silpnas arba nepakankamai aiškus
Technologinis įgyvendinamumas	Aukštas, nes remiasi esama padėtimi	Aukštas / vidutinis	Vidutinis

Kriterijus	Status quo – decentralizuotas modelis	Nacionalinis mazgas	Regioninis mazgas
Igyvendinimo sudėtingumas	Mažas	Vidutinis	Labai didelis
Koordinavimo sudėtingumas	Mažas	Vidutinis	Labai didelis
Ilgalaikio finansavimo aiškumas	Ribotas ir fragmentuotas	Santykinai aiškiausias	Neaiškus, priklausomas nuo kelių šalių susitarimų
Tęstinumo po projekto prielaidos	Silpnos, nes nėra bendros nacionalinės struktūros	Stiprios	Silpnos / rizikingos
Sąveika su kitomis nacionalinėmis duomenų erdvėmis	Ribota	Didelė	Ribota
Sektoriaus autonomija	Aukšta atskirų MTI lygmeniu, bet be bendro koordinavimo	Subalansuota	Ribota nacionaliniu lygmeniu, nes dalis sprendimų perkeliama į regioninį lygmenį
Priklausomybė nuo išorinių partnerių	Vidutinė	Maža	Didelė
Praktinės patirties pagrįstumas	Aukštas, nes tai esama padėtis	Aukštas, nes EOSC praktikoje dominuoja nacionaliniai mazgai	Ribotas
Tinkamumas Lietuvos poreikiams	Vidutinis	Aukščiausias	Ribotas
Bendras vertinimas	Atskaitinis taškas, bet nepakankamas strateginis sprendimas	Rekomenduojama alternatyva	Metodologiškai galima, bet praktiškai mažiau tinkama

6. Finansinė MSI mokslinių tyrimų duomenų valdymo situacija

6.1. Dabartinis finansavimas

Mokslinių tyrimų duomenų valdymo finansavimas Lietuvoje yra sudėtinė bendros mokslo ir studijų finansavimo sistemos dalis. Jis užtikrinamas instituciniu MTEP finansavimu, LMT administruojamu konkursiniu projektu finansavimu, MTI finansavimu ir tarptautinių programų lėšomis. Šie finansavimo kanalai apima institucinius, projektinius, infrastruktūrinius ir tarptautinio bendradarbiavimo lygmenis.

Pagrindinis finansavimo šaltinis mokslo ir studijų institucijoms yra bazinio ir skatinamojo finansavimo lėšos, skiriamos atsižvelgiant į LMT atliekamo ekspertinio (lyginamojo) vertinimo, vykdomo kas penkerius metus, ir kasmetinio formaliojo vertinimo rezultatus. Šių dviejų vertinimo dedamųjų santykis sudaro atitinkamai 70 ir 30 proc. Toks modelis reiškia, kad dalis mokslinių tyrimų duomenų tvarkymo poreikių Lietuvoje yra dengiama bendru instituciniu MTEP finansavimu, iš kurio palaikoma tyrimų aplinka, žmogiškieji ištekliai, institucinė infrastruktūra ir vidiniai duomenų valdymo pajėgumai.

Kitas svarbus finansavimo šaltinis yra konkursinis finansavimas, kurį administruoja LMT. LMT finansuojamų projektų sistemoje duomenų valdymas yra institucionalizuotas. Projektų vykdytojams teikiamos duomenų valdymo plano formos ir rekomendacijos, o su atviruoju mokslu susiję reikalavimai grindžiami 2024 m. rugpjūčio 30 d. patvirtintu Atvirosios prieigos prie mokslinių tyrimų ir eksperimentinės plėtos rezultatų tvarkos aprašu (Lietuvos mokslo taryba, 2024a). Be to, bent dalyje LMT administruojamų priemonių tinkamų finansuoti išlaidų kategorijose

tiesiogiai numatomos ne tik publikavimo, bet ir MTD tvarkymo bei atvėrimo išlaidos. Todėl konkursinis finansavimas veikia kaip praktinis mechanizmas, per kurį duomenų valdymo veiklos gali būti finansuojamos projekto lygmeniu.

Tam tikrą vaidmenį atlieka ir tikslinės konkursinio finansavimo priemonės, kurios, nors ir nėra skirtos vien MTD, realiai stiprina duomenų tvarkymo, skaitmenizavimo ir skaičiavimo pajėgumus. Šiuo požiūriu reikšminga LMT paskirtinė programa „Informacinės technologijos mokslo ir žinių visuomenės plėtrai“, duomenų valdymo pajėgumus taip pat stiprina kitos tikslinės valstybės finansavimo programos. Pavyzdžiui, pagal ŠMSM programą „Universitetų ekselencijos iniciatyva“ finansuojamas VU projektas „Duomenų centras mašiniam mokymui ir kvantiniams skaičiavimams gamtos ir biomedicinos mokslų srityse“, įgyvendinamas 2023–2027 m. Šiam projektui skirtas 2,563 mln. Eur valstybės biudžeto finansavimas. Tokios iniciatyvos, nors ir nėra tiesiogiai orientuotos į visos nacionalinės MTD sistemos plėtrą, sudaro prielaidas stiprinti institucinius duomenų kaupimo, atvėrimo ir naudojimo pajėgumus.

Atskiras finansavimo blokas yra MTI finansavimas. Lietuvos mokslinių tyrimų infrastruktūrų kelrodis yra politikos dokumentas, nustatantis ilgalaikes nacionalines ir tarptautines MTI politikos kryptis, prioritetus, galimybes ir perspektyvas. 2024 m. Kelrodyje įtrauktos 24 MTI, tarp jų ir tokios, kurios tiesiogiai susijusios su duomenų kaupimu, skaičiavimais, ilgalaikiu saugojimu ar prieiga prie duomenų, pavyzdžiui, „LitGrid-HPC“, QUARTIC, LiDA ir CLARIN-LT. Lietuvos mokslo tarybos 2026–2028 m. strateginis veiklos planas taip pat numato finansavimą dalyvauti tarptautinių mokslinių tyrimų infrastruktūrų veikloje, taip pat valstybės tikslinėms mokslinių tyrimų infrastruktūroms.

Reikšmingas finansavimo šaltinis yra ir tarptautinės programos, ypač ES mokslinių tyrimų ir inovacijų programa „Europos horizontas“. Šioje programoje su MTD ir kitų rezultatų valdymu susijusios išlaidos laikomos tinkamomis finansuoti, jei jos atitinka dotacijos sutarties sąlygas. Tai reiškia, kad projektų vykdytojai gali planuoti tokias išlaidas kaip duomenų tvarkymas, saugojimas, archyvavimas, pakartotinio naudojimo užtikrinimas ir kitos FAIR principų įgyvendinimo veiklos. Be to, „Europos horizontas“ finansuoja ir specifinius projektus, susijusius su integracija į Europos atvirojo mokslo debesį (EOSC), kuriais siekiama kurti federuotą aplinką, sudarančią sąlygas tyrėjams dalytis duomenimis ir jais naudotis tarpvalstybiniu bei tarpdalykiniu mastu.

Šiame kontekste pažymėtinas būsimas kvietimas „Expanding and deepening the EOSC Federation“ pagal priemonę HORIZON-INFRA-2027-01-EOSC-01. Tai bendro finansavimo veiksmas (angl. *co-fund action*), kuriam numatyta iki 40 mln. Eur Europos Sąjungos lėšų, taikant iki 50 proc. tinkamų finansuoti išlaidų finansavimo intensyvumą. Likusi dalis turėtų būti dengiama valstybių narių ar kitų dalyvaujančių institucijų lėšomis. Tokios priemonės svarbios tuo, kad jos finansuoja ne tik pavienius tyrimus, bet ir platesnę federuotų duomenų bei paslaugų aplinkos plėtrą europiniu mastu.

Apibendrinant galima teigti, kad MTD finansavimas Lietuvoje yra mišrus ir daugiakanalis. Jį sudaro institucinis bazinis finansavimas, LMT konkursiniai projektai, MTI finansavimas, teminės nacionalinės programos, tikslinės valstybės investicijos ir tarptautinės programos ir kt. Vis dėlto šie finansavimo srautai daugiausia remia duomenų valdymą kaip platesnių MTEP, infrastruktūros ar skaitmenizacijos veiklų dalį, o ne kaip savarankišką, nuoseklią nacionalinę finansavimo eilutę, apimančią visą MTD gyvavimo ciklą – nuo jų sukūrimo ir aprašymo iki ilgalaikio išsaugojimo, atvėrimo ir pakartotinio naudojimo.

6.2. Būtinasis finansavimas ir jo paskirstymas integruojantis į EOSC

Remiantis atlikta dabartinės situacijos Lietuvoje bei užsienio šalių patirčių analize, apibendrintai galima teigti, jog yra būtinas poreikis atlikti kertinius MTD valdymo situacijos pokyčius, siekiant teigiamo socialinio-ekonominio poveikio. Parengtos Lietuvos įsijungimo į EOSC alternatyvos yra pagrindas, siekiant nustatyti būtinąjį finansavimą ir jo paskirstymą integruojantis į EOSC.

LIETUVOS ĮSIJUNGIMO Į EOSC ALTERNATYVOS

I. Lietuvos įsijungimo į EOSC alternatyva – decentralizuotas *status quo* modelis

Prielaidos *status quo* alternatyvai vertinti

Dabartinai MTD valdymo situacijai Lietuvoje būdingi šie požymiai:

- MTD kūrimas finansuojamas daugiausia valstybės lėšomis (LMT, ESF projektų, taip pat „Horizon Europe“ projektų finansavimu. Mokslinių tyrimų finansavimą iliustruoja 21 lentelė).
- Fragmentuotas MTD valdymas, kuris remiasi dažniausiai (atsitiktiniu – konkurso keliu) gaunamu išoriniu projektiniu finansavimu. Institucijoms toks fragmentuotas finansavimas sukelia sunkumų sukurti MTD politiką ir ją įgyvendinti.
- Žema išorinė atvira prieiga prie MTD bei galimybės pakartotinai juos naudoti. Mokslinių tyrimų rezultatai fragmentuoti, MTD yra institucijų repozitorijuose arba lieka mokslininkų asmeniniuose kompiuteriuose ar kitose laikmenose.
- Dideli sukurtų MTD praradimai. Keičiantis IT sistemų kartoms, kompiuteriams bei kitoms laikmenoms, anksčiau sukurti MTD neperkeliami į naujos kartos laikmenas ir neišvengiamai prarandami.
- Poreikis nuolat atnaujinti tyrimus be galimybės palyginti jų rezultatus su ankstesnių bangų tyrimais dėl prarastų jų rezultatų arba MTD palyginamumo stygiaus.
- Nepasitikėjimas sukurtų MTD kokybe. Eksponentiškai auga mokslinių publikacijų, patentų skaičius, kurie remiasi atliktais moksliniais tyrimais. Tačiau dėl ribotos prieigos (arba jos nebuvimo) prie MTD nėra galimybės tyrimus replikuoti ir patikrinti rezultatų validumą.

MTD valdymo politikos pokyčiai tarptautiniu mastu:

- Reprezentaciniai tarptautiniai mokslo žurnalai skatina tokias publikacijas, kurios atveria MTD arba iš anksto juos registruoja (angl. *pre-register*) AP repozitorijuose (arba publikuoja išimtinai tik tokius mokslinius straipsnius).
- OECD patvirtintos 2021 m. politikos gairės (OECD, 2021), kuriomis skatinami FAIR principai valstybės lėšomis finansuojamų MTD atžvilgiu. Pagrindiniai šios politikos principai – pasitikėjimo stiprinimas, techninių standartų nustatymas, MTD nuosavybės apibrėžimas, paskatų teikimas, tvarios infrastruktūros užtikrinimas, žmogiškųjų išteklių ugdymas ir tarptautinio bendradarbiavimo skatinimas.
- OECD paskelbtas 2025 m. įgyvendinimo priemonių rinkinys¹¹⁰, apimantis šiuos elementus:
 1. *MTD valdymo planavimas*: priemonės duomenims valdyti ir saugoti.
 2. *Teisinės ir etinės gairės*: rekomendacijos dėl intelektinės nuosavybės ir MTD privatumo valdymo.
 3. *Tvari infrastruktūra*: parama MTD saugykloms ir skaitmeninėms MTD dalijimosi priemonėms.
 4. *Paskatos*: MTD dalijimosi pripažinimo skatinimas profesinėje karjeroje.

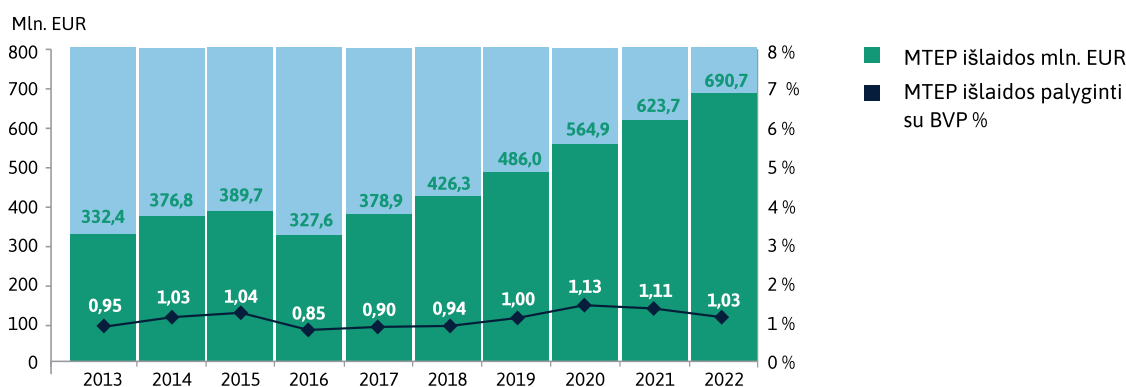
¹¹⁰ https://www.oecd.org/en/publications/access-to-public-research-data-toolkit_a12e8998-en/research-data-repository-development-and-operation-guidelines_d6edc5f8-en.html#:~:text=To%20enhance%20the%20management%20and,privacy%20and%20intellectual%20property%20

Vertinant dabartinę MTD ir jų kūrimo situaciją bei mastelius Lietuvoje svarbūs yra šie aspektai:

- Bendras MSI skaičius.
- MTI skaičius (Kelrodžio duomenimis).
- MSI, vykdančių doktorantūrą, skaičius.
- Bendras mokslininkų (tyrėjų) skaičius.
- Tyrėjų skaičius, metinis vidutinis mokslininkų visos darbo dienos atitikmuo (toliau – MVDDA, angl. *full time equivalent*, FTE).
- Bendras MTEP veiklų finansavimas bei jo dalis nuo BVP.
- MTEP išlaidos aukštojo mokslo sektoriuje.

Dabartinė MTEP situacija Lietuvoje charakterizuojama šiais duomenimis. Oficialios statistikos portalo duomenimis, 2022 m. MTEP išlaidos sudarė 690,7 mln. EUR¹¹¹.

Mokslinių tyrimų ir eksperimentinės plėtros (MTEP) išlaidos



11 pav. MTEP išlaidų Lietuvoje dinamika, 2013-2022 m.

2022 m. daugiausia išlaidų patyrė verslo įmonių sektorius – 334,9 mln. EUR (48,5 proc.) visų MTEP išlaidų, aukštojo mokslo sektorius – 251,7 mln. EUR (36,4 proc.), ir valdžios sektorius – 104,1 mln. EUR (15,1 proc.). Bendrosios Lietuvos išlaidos MTEP sudarė 1,03 proc. BVP (2021 m. – 1,11 proc.). Verslo sektoriaus MTEP išlaidos 2022 m. sudarė 0,5 proc. BVP (2021 m. – 0,55 proc.), aukštojo mokslo ir valdžios sektoriaus – 0,54 proc. (2021 m. – 0,56 proc.).

Verslo įmonių lėšos sudarė didžiausią MTEP išlaidų dalį pagal finansavimo šaltinius – 38,7 proc., valdžios lėšos – 31,2 proc., užsienio lėšos – 26,7 proc., aukštojo mokslo ir ne pelno institucijų lėšos – 3,4 proc.

Didžiausią MTEP išlaidų dalį pagal tikslinę paskirtį sudarė darbui apmokėti skirtos lėšos – 57,6 proc. (397,9 mln. EUR), kitos išlaidos – 30,4 proc. (210,1 mln. EUR), ilgalaikiam materialiajam turtui įsigyti – 12 proc. (82,7 mln. EUR).

2022 m. išlaidos fundamentiniams tyrimams sudarė 21,9 proc. visų išlaidų MTEP, taikomiesiems tyrimams – 45,3 proc., eksperimentinei plėtrai – 32,8 proc. Verslo sektorius didžiausią dalį lėšų – 56,6 proc. – skyrė eksperimentinei plėtrai. Aukštojo mokslo ir valdžios sektoriai didžiausią dalį lėšų skyrė taikomiesiems tyrimams – atitinkamai 52,6 ir 45,5 proc.

¹¹¹ <https://osp.stat.gov.lt/>

Mokslinių tyrimų ir eksperimentinės plėtros (MTEP) išlaidos pagal sektorius

Mln. EUR

	2018	2019	2020	2021	2022
Iš viso	426,3	486,0	564,9	623,7	690,7
Aukštojo mokslo sektorius	153,2	176,8	210,1	218,9	251,7
Valdžios sektorius	94,8	99,1	89,4	98,2	104,1
Verslo įmonių sektorius	178,3	210,12	265,4	306,6	334,9

21 pav. MTEP išlaidų pagal sektorius Lietuvoje dinamika 2018-2022 m.

2022 m. mokslo tiriamąja veikla užsiėmė 26 tūkst. darbuotojų, 9 tūkst. iš jų turėjo mokslo daktaro laipsnį. Aukštojo mokslo ir valdžios institucijose MTEP (pagrindinėje ir ne pagrindinėje darbovietėje) dalyvavo apie 17,9 tūkst. darbuotojų, iš jų 8,3 tūkst. – tyrėjai, turintys mokslo daktaro laipsnį. Verslo sektoriuje MTEP veikloje dalyvavo 8 tūkst. darbuotojų, iš jų 725 – tyrėjai, turintys mokslo daktaro laipsnį. Aukštojo mokslo ir valdžios sektoriuje daugiausia tyrėjų buvo 35–44 metų amžiaus – 3 967 (27,6 proc.), 25–34 metų amžiaus – 3 139 (21,8 proc.), 45–54 metų amžiaus – 3 626 (25,2 proc.), vyresni nei 65 metų tyrėjai sudarė 6,8 proc. Verslo sektoriuje daugiausia tyrėjų buvo 25–34 metų amžiaus – 1 876 (3,9 proc.), 35–44 metų amžiaus – 1 613 (33,4 proc.), 45–54 metų amžiaus – 624 (12,9 proc.), vyresni nei 65 metų tyrėjai sudarė 2,4 proc. visų tyrėjų.

Tyrėjai, turintys mokslo laipsnį

Kriterijus	2018	2019	2020	2021	2022
Iš viso	8 521	8 293	8 521	8 857	9 016
Aukštojo mokslo sektorius	8 561	6 367	6 456	6 731	6 778
Valdžios sektorius	1 466	1 432	1 483	1 480	1 513
Verslo įmonių sektorius	494	492	582	644	725

22 pav. Tyrėjų, turinčių mokslo laipsnius, skaičius Lietuvoje dinamika 2018-2022 m.

Skaičiuojant pagal MVDDA 2023 m. LMT, mokslo palyginamojo vertinimo duomenimis, sudarė 3965,6 MVDDA, t. y. kiek mažiau nei 4000 MVDDA (angl. *full time equivalent*, FTE)¹¹². 2022 m. duomenimis, MTEP dalis vienam MVDDA sudarė apytiksliai 0,172 mn Eur / MVDDA (FTE).

Šiais duomenimis remiantis galima teigti, jog Lietuva yra žemo mokslo intensyvumo valstybė.

2024 m. duomenimis, Lietuvoje veikė platus MSI tinklas, apimantis universitetus ir mokslo tyrimo centrus. Kasmetiniame ir lyginamajame vertinime dalyvavo 28 institucijų. Didelis doktorantūrą vykdančių institucijų skaičius.

¹¹² <https://lmt.lrv.lt/lt/veiklos-sritys/mokslo-kokybe/mokslo-meno-vertinimas/universitetai-ir-moksliniu-tyrimu-institutai/kasmetinis-mokslo-meno-veiklos-vertinimas/2024-m/vertinimo-rezultatai2024/universitetu-ir-moksliniu-tyrimu-institutu-gautu-tarptautiniu->

Toliau pateikiamas pagrindinių Lietuvos MSI, kurios vykdo doktorantūrą, t. y. atlieka mokslinius tyrimus bei rengia mokslininkus, sąrašas.

- VU – vykdo doktorantūrą daugelyje mokslo sričių (humanitarinių, socialinių, fizinių, biomedicinos, technologijos).
- VILNIUS TECH – stiprus technologijų ir inžinerijos doktorantūros centras.
- KTU – vykdo įvairių sričių, ypač technologijos ir fizinių mokslų, doktorantūrą.
- VDU – vykdo socialinių, humanitarinių, menų doktorantūros studijas, dažnai kartu su kitais universitetais.
- MRU – specializuojasi socialinių mokslų (teisės, vadybos, psichologijos) doktorantūroje.
- LSMU – biomedicinos mokslų doktorantūra.
- KU – vykdo jūrinės inžinerijos, humanitarinių ir socialinių mokslų doktorantūrą.
- LSU (Lietuvos sporto universitetas).
- ISM Vadybos ir ekonomikos universitetas.

Mokslinių tyrimų institutai, turintys doktorantūros teisę (dažnai kartu su universitetais):

- Lietuvių kalbos institutas.
- Lietuvos istorijos institutas.
- LAMMC.
- Fizinių ir technologijos mokslų centras (FTMC).
- Lietuvos socialinių mokslų centras.
- Gamtos tyrimų centras.
- Lietuvos energetikos institutas.
- Lietuvos agrarinių ir miškų mokslų centras.

Mokslo veiklos, tyrimų mastus bei mokslinių tyrimų efektyvumą rodo „WoS Incites“ apibendrinti Lietuvos tyrėjų publikavimo bei publikacijų cituojamumo (2025 m.) duomenys¹¹³.

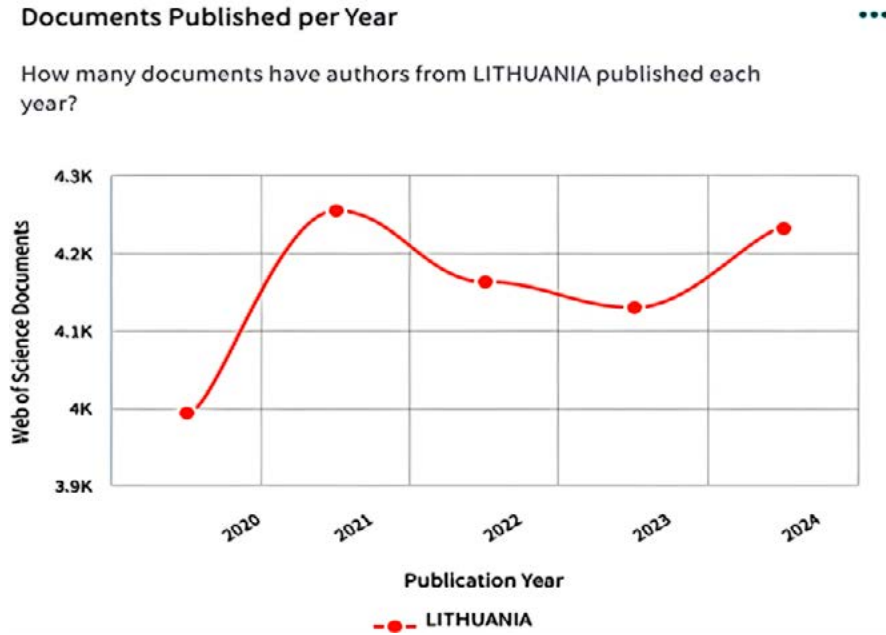
Overview	
Documents Published	Times Cited
20,779	284,514
% Documents in Top 1%	% Documents in Top 10%
1.56	10.73
% Documents Cited	Average CNCI
86.33	1.16

12 pav. Lietuvos tyrėjų publikacijų, įtrauktų į WoS, skaičius bei jų cituojamumas 2025 m.

¹¹³ WoS Incites https://incites.clarivate.com/#/tabreports/dbtbr_111?selectedReportId=dbr_291

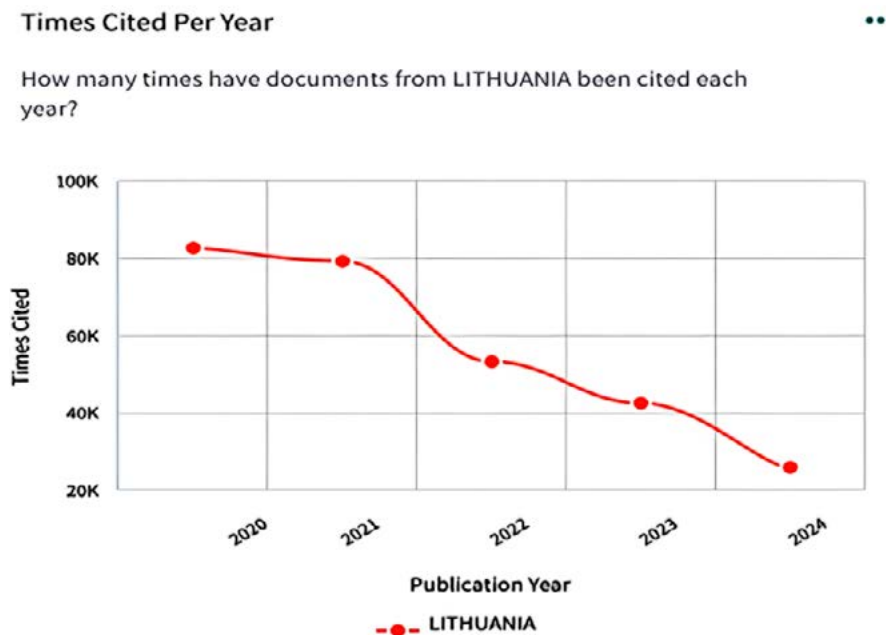
Mokslinių straipsnių publikacijų (tik žurnaluose) WoS leidiniuose 2020–2024 m. su Lietuvos MSI prieskyromis pateiktos 13 paveikslėlyje.

13 pav. Lietuvos tyrėjų publikacijų, įtrauktų į WoS, dinamika 2020–2024 m.



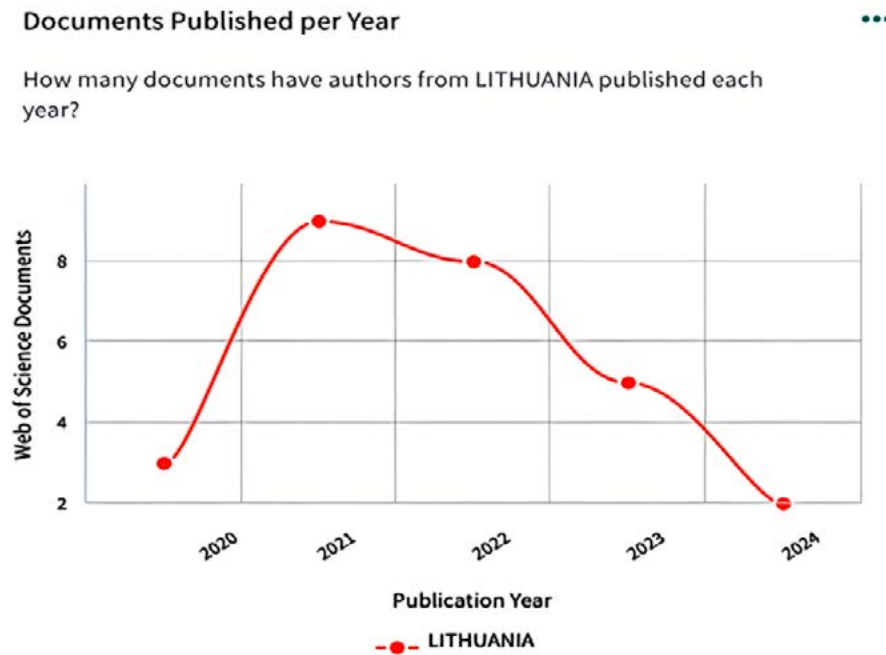
Mokslinių straipsnių (tik žurnaluose) WoS leidiniuose 2020–2024 m. su Lietuvos MSI prieskyromis cituojamumas pateiktas 14 paveikslėlyje.

14 pav. Lietuvos tyrėjų publikacijų, įtrauktų į WoS, cituojamumo dinamika 2020–2024 m.

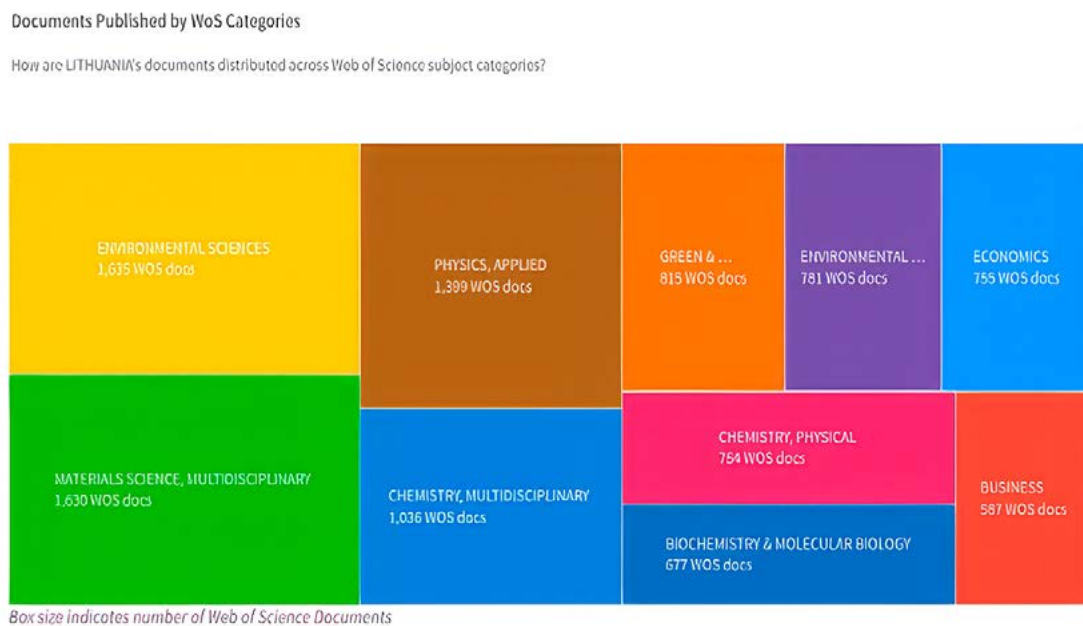


Duomenų pagrindu WoS publikuoti straipsniai (angl. *Data papers*) 2020–2024 m. su Lietuvos MSI prieskyromis pateikti 15 paveikslėlyje.

15 pav. Lietuvos tyrėjų straipsnių (angl. *Data papers*), įtrauktų į WoS, dinamika 2020–2024 m.



Mokslinių straipsnių (tik žurnaluose) WoS leidiniuose 2020–2024 m. su Lietuvos MSI prieskyromis pasiskirstymas pagal mokslo sritis pateikti 16 paveikslėlyje.



16 pav. Straipsnių Lietuvos MSI prieskyromis, įtrauktų į WoS, pasiskirstymas pagal mokslo sritis, 2020–2024 m.

Šie Lietuvos tyrėjų publikacijų WoS duomenys demonstruoja pastarojo meto MTD kūrimo mastus bei jų pasiskirstymą pagal mokslo sritis Lietuvoje.

Tačiau nepaisant gana išplėtos mokslo tyrimų institucinės struktūros ir tinklo, Lietuva pasižymi fragmentuota MTD kūrimo ir valdymo struktūra. Tokia struktūra, nuolat augant mokslinių tyrimų, sukuriančių MTD skaičiui, nesuformuoja prielaidų kokybiniam MTD valdymo šuoliui ir sukelia praradimus (nuostolius).

Taigi *status quo* alternatyva pasižymi:

- lėtu AP prie MTD centrų formavimusi;
- fragmentuota MTD bei metaduomenų apdorojimo standartizavimo, sistemų suderinamumo ir dalijimosi MTD praktika;
- papildomu tikslinio finansavimo esamoms infrastruktūroms, jų duomenų saugykloms, poreikiu, užtikrinant Lietuvoje sukurtam MTD srautą į EOSC;
- metodinės ir organizacinės paramos stygiumi nacionalinėms MTI jų integracijos procese.

Patiriama praradimų (nuostolių) dėl *status quo* išlaikymo.

Remiantis užsienio tyrimų (European Commission, 2019) duomenimis, išlaikant *status quo*, institucijos patiria nuostolių, kai nėra efektyvios MTD valdymo sistemos:

1. *Neigiamas poveikis mokslinių tyrimų vykdymui gali būti vertinamas remiantis šiais kriterijais:*

a) Tyrėjų laikas, prarastas ieškant duomenų.

Tyrėjai reikšmingą savo laiko dalį praleidžia ieškodami MTD. Tai kompleksinis rodiklis, susidedantis iš kelių elementų. Prarasto laiko sąnaudos dėl neefektyvios duomenų paieškos gali būti apskaičiuojamos derinant tyrėjų skaičių (FTE)¹¹⁴, jų vidutinius atlyginimus¹¹⁵ ir laiką, skirtą MTD paieškai. Ekspertų vertinimu, 31,52 % laiko gali būti sutaupyta, taikant standartizuotus metaduomenų saugojimo sprendimus.

b) Laikas, prarastas dėl perteklinio duomenų saugojimo.

Sąnaudos yra susijusios su atsarginių kopijų sudarymu, saugojimu ir jo trukme. ES duomenimis, vidutiniškai MTD saugomi 4,63 laikmenose / repozitorijuose 10 metų. Tinkami MTD valdymo sprendimai galėtų saugomų MTD kiekius sumažinti 25,67 %.

c) Licencijų sąnaudos

Organizacijų MTD finansavimo išlaidos, skirtos duomenų licencijoms įsigyti, kai nėra duomenų alternatyvios atvirosios priegios. Tai turi poveikį mokslinių tyrimų vystymo perspektyvai.

2. *Neigiamas poveikis tolesnių mokslinių tyrimų plėtrai gali būti vertinamas remiantis šiais kriterijais:*

a) MTD atkūrimo ar pakartotinio kūrimo sąnaudos

Kai MTD nėra randami ar prieinami, tyrėjams tenka juos sukurti iš naujo. Toks pastangų dubliavimas lemia neefektyvumą ir papildomas sąnaudas. Bendros prarasto laiko sąnaudos formuojasi dėl MTD atkūrimo poreikio.

b) Dvigubo mokslinių tyrimų finansavimo sąnaudos

ES „Horizon Europe“ duomenys rodo, kad stebima mokslinių tyrimų tematikų galimai dvigubo finansavimo tendencija. Tikėtina, kad tokia tendencija pasireiškia ir finansuojant Lietuvos mokslinius projektus. Vis dėlto, nesant duomenų, sudėtinga šią tendenciją įvertinti skaičiais.

¹¹⁴ <https://lmt.lrv.lt/lt/veiklos-sritis/mokslo-kokybe/mokslo-meno-vertinimas/palyginamasis-ekspertinis-mokslo-veiklos-vertinimas/>

¹¹⁵ <https://www.ftmc.lt/vidutinis-darbo-uzmokestis/>

c) Tarpdisciplininių tyrimų vykdymo galimybių praradimas

Tarpdisciplininių tyrimų plėtra priklauso nuo geresnės MTD valdymo rezultatų kokybės, glaudesnio tyrėjų bendradarbiavimo ir pakartotinio MTD naudojimo galimybių.

d) Nepakankamai išnaudojamas MTI potencialas

Taip nėra skatinamas MTI sukurtų MTD platesnis naudojimas, taip pat ir pakartotinis naudojimas (angl. *reusability*).

3. *Poveikis inovacijoms šalyje vertinamas kaip potencialus poveikis ekonomikos augimui*

Formuojasi nepakankamas potencialo išnaudojimas. Potencialus poveikis ekonomikos augimui (BVP augimui, %) priklauso nuo mokslinių tyrimų sričių bei krypčių, poveikio inovacijų kūrimui, startuolių atsiradimui. Vis dėlto, nesant duomenų, sudėtinga šią tendenciją įvertinti skaičiais.

Reputacijos mokslinių tyrimų ir jų duomenų valdymo srityje praradimo nuostoliai, tiek Lietuvoje, tiek užsienyje. Lietuvoje išlieka neigiamos reputacijos nuostatos verslo sektoriuje. Nepakankamai išnaudojamos tarptautinio mokslinio bendradarbiavimo galimybės.

II. Lietuvos įsijungimo į EOSC alternatyva – nacionalinio EOSC mazgo sukūrimo alternatyva

Prielaidos nacionalinio EOSC mazgo sukūrimo alternatyvai įvertinti

Šioje alternatyvoje išskiriami trys organizacinio valdymo scenarijai, kuriems bendra esminė mazgo architektūra, tačiau kurie skiriasi valdymo, koordinavimo ir institucinės atsakomybės pasiskirstymu. Visi trys scenarijai pasižymi bendromis investicijų finansavimo prielaidomis.

Numatomos investicijų finansavimo kryptys:

1. EOSC-LT-Core – centrinis nacionalinio mazgo sukūrimas, kurio branduolį sudaro nacionalinis centrinis MTD metaduomenų katalogas, atliekantis metaduomenų agregavimo, katalogavimo, sąveikumo koordinavimo ir centralizuotų paslaugų teikimo funkcijas.

2. Administruojančios struktūros valstybės lygmeniu sukūrimas.

3. Investicijos į federacinio tinklo sukūrimą – integracija su instituciniais ir teminiais vienetais (institucijų valdomas ar centralizuotai palaikomas mokslinių duomenų saugyklos, metaduomenų atitikčių sudarymo sprendimai, duomenų aprašymo, sąveikumo ir konsultavimo veiklos).

4. Investicijos į duomenų valdytojų sistemos bei duomenų valdytojų tinklo sukūrimą.

Finansinių skaičiavimų prielaidos:

- Projekto trukmė – treji metai.
- Pradinė projekto investicija – 9 mln. Eur.
- MTI infrastruktūrų skaičius – 24 (Lietuvos mokslo taryba, 2024b).
- Duomenų valdytojų skaičiaus poreikis MTI infrastruktūrai, vidutinis – 2 MVDDA (FTE).
- MSI ir MTI institucijų skaičius – 28¹¹⁶.
- Duomenų valdytojų skaičiaus poreikis MSI ir MTI institucijai, vidutinis – 4 MVDDA (FTE).¹¹⁷.

¹¹⁶ Turinčios doktorantūros teisę, žr. LMT, ŠSMM įsakymai

¹¹⁷ <https://opensciencefair.eu/lightning-talks/5-data-stewards-per-100-researchers-the-development-of-a-postgraduate-certificate-data-steward-at-the-university-of-vienna>

Centrinio nacionalinio mazgo sukūrimas

- Pradinė investicija ir palaikymas per dvejus metus – 2,6 mn Eur.

Administruojančios struktūros valstybės lygmeniu sukūrimas

- Numatomas įsteigti etatų) skaičius – 4 MVDDA (FTE).
- Projektinis etatų finansavimas – 2 metai.
- Metinės vieno darbuotojo darbo užmokesčio sąnaudos, atsižvelgiant į infliaciją – 45814 Eur/m.¹¹⁸

Duomenų valdytojų sistemos ir duomenų valdytojų tinklo sukūrimas

- Duomenų valdytojų sistemos bei mokymų programos parengimas pirmais projekto metais
- Duomenų valdytojų tinklo sukūrimas (mokymai, konsultavimas per dvejus projekto metus) – vieno duomenų valdytojo apmokymo kaina – 3000 Eur (15 ECTS)¹¹⁹.

MSI ir MTI institucinių duomenų valdymo struktūrų sukūrimas einant į EOSC

- Darbuotojų skaičius per metus – 40 MVDDA (FTE).
- Metinės vieno darbuotojo darbo užmokesčio sąnaudos, atsižvelgiant į infliaciją – 51 900 Eur/m.
- Patalpų išlaikymo išlaidos – 5 % nuo darbo užmokesčio, suma, Eur.
- Pridėtinės išlaidos – 30 % nuo darbo užmokesčio, suma, Eur.

Nacionalinio mazgo ir federacinės sistemos išlaikymo sąnaudos po projekto pabaigos (neatsižvelgiant į infliaciją)

- Centrinis nacionalinio mazgo palaikymas – 100,0 tūkst. Eur.
- Administruojančios struktūros valstybės lygmeniu palaikymas:
 - Numatomas etatų skaičius – 4 MVDDA (FTE).
 - Metinės vieno darbuotojo darbo užmokesčio sąnaudos, atsižvelgiant į infliaciją – 45 814 Eur/m.
- MSI ir MTI institucinių duomenų valdymo struktūrų palaikymas:
 - Duomenų valdytojų skaičius per metus – 160 MVDDA (FTE).
 - Metinės vieno duomenų valdytojų darbo užmokesčio sąnaudos, atsižvelgiant į infliaciją – 51 900 Eur/m.
 - Patalpų išlaikymo išlaidos – 5 % nuo darbo užmokesčio, suma, Eur.
 - Pridėtinės išlaidos, 30% nuo darbo užmokesčio, suma, Eur.

¹¹⁸ <https://www.ftmc.lt/vidutinis-darbo-uzmokestis/>

¹¹⁹ <https://rdm.univie.ac.at/data-stewards-at-the-university/become-a-data-steward/>

23 lentelė. Investicijos projekto laikotarpiu

Investicijos projekto laikotarpiu

Investicijų / išlaidų pobūdis	1 metai	2 metai	3 metai	Iš viso per 3 metus
<i>I. Nodo (centrinio katalogo pagrindu) sukūrimas ir išlaikymas</i>				
Nodo (centrinio katalogo pagrindu) sukūrimas (paslaugos pirkimas), Eur	2350000			
Metinė Nodo (centrinio katalogo pagrindu) priežiūra, palaikymas po sukūrimo (paslaugos pirkimas), Eur		150000	100000	
Bendra suma, Eur	2350000	150000	100000	2600000

<i>II. Administravimas, struktūros palaikymas</i>				
Darbo užmokestis, suma, Eur		183256	183256	
Darbuotojų skaičius (4 FTE)		4	4	
Vidutinis darbo užmokestis, Eur		45814	45814	
Patalpų išlaikymo išlaidos, 5% nuo darbo užmokesčio, suma, Eur		9162,8	9162,8	
Pridėtinės išlaidos, 30% nuo darbo užmokesčio, suma, Eur		54976,8	54976,8	
Bendra suma, Eur		247395,6	247395,6	494791,2

<i>III. MTI institucinių duomenų valdymo struktūrų sukūrimas einant į EOSC</i>				
Darbo užmokestis, suma, Eur	0	2076000	2076000	
Darbuotojų skaičius kasmet (40 FTE)		40	40	
Vidutinis darbo užmokestis, Eur		51900	51900	
Patalpų išlaikymo išlaidos, 5% nuo darbo užmokesčio, suma, Eur		103800	103800	
Pridėtinės išlaidos, 30% nuo darbo užmokesčio, suma, Eur		622800	622800	
Bendra suma, Eur	0	2802600	2802600	5605200
Iš viso, Eur				8699991,2

300000

PROJEKTINIS FINANSAVIMAS				8999991,2
---------------------------------	--	--	--	------------------

24 lentelė. Investicijos į duomenų valdytojų sistemos ir mokslo duomenų valdytojų tinklo sukūrimą

Investicijų / išlaidų pobūdis	1 metai	2 metai	3 metai	Iš viso per 3 metus	Viso parengta DS	Parengta DS institucijai
Duomenų stiuartų sistemos ir programos parengimas (paslaugos pirkimas), Eur	600000					
Duomenų stiuartų programos įgyvendinimas (mokymai) (paslaugos pirkimas), Eur		120000	120000			
Programos kaina vienam dalyviui, 15 ECTS, Eur		3000	3000			
MSI ir MTI institucijų skaičius - 28		14	14	28	28	
Duomenų stiuartų skaičius institucijai, vidutinis - 4 FTE		2	2	4	4	
MTI infrastruktūrų skaičius - 24		12	12	24	24	
Duomenų stiuartų skaičius MTI infrastruktūrai, vidutinis - 2 FTE		1	1	2	2	
Bendra suma	60000	120000	120000	300000	160	3,8

Investicijų paskirstymas tarp mokslo sričių per projekto laikotarpį

Finansinių skaičiavimų prielaidos:

- Projekto trukmė – treji metai.
- Pradinė projekto investicija – 5,61 mln. Eur.
- Bendras tyrėjų skaičius, remiantis 2022 m. duomenimis, 6778 tyrėjai.
- Bendras tyrėjų skaičius MVDDA (FTE), remiantis 2023 m. LMT lyginamojo vertinimo duomenimis – 3965,6 MVDDA (FTE).
- Investicijoms paskirstyti tarp mokslo sričių naudojami kriterijai:
 - 1) bendro tyrėjų skaičiaus dalis mokslo srityje % (2023 m. LMT palyginamojo vertinimo duomenimis);
 - 2) vidutinė MTEP projektinio finansavimo dalis mokslo srityje % (2020–2023 m. duomenimis);
 - 3) apytikslė bendro doktorantų skaičiaus mokslo srityje dalis, % (2022–2024 m. priėmimo duomenimis)ʹ.
- Kriterijų svoriai – atitinkamai 40 %, 40 % ir 20 % (ekspertinis vertinimas).
- Skaičiuoti taikomas svertinio indekso modelis, kurio formulė yra ši:

$$S_i = w_p \cdot p_i + w_f \cdot f_i + w_D \cdot d_i$$

MTEP projektų finansavimo dalis:

$$p_i = \frac{P_i}{\sum_{j=1}^k P_j}$$

tyrėjų skaičiaus dalis:

$$f_i = \frac{F_i}{\sum_{j=1}^k F_j}$$

doktorantų skaičiaus dalis:

$$d_i = \frac{D_i}{\sum_{j=1}^k D_j}$$

čia w_p – MTEP projektų dalies rodiklio svoris; w_f – tyrėjų dalies rodiklio svoris;

w_D – doktorantų dalies rodiklio svoris;

S – bendra finansavimo suma (eurais);

A – investicijų dalis, tenkanti mokslo sričiai:

$$A_i = \frac{S_i}{\sum_{j=1}^k S_j}$$

Duomenys skaičiavimams pateikiami 25, 26 ir 27 lentelėse, remiantis apytiksliais pastarųjų metų kriterijų rezultatų vertinimais. Bendro tyrėjų skaičiaus pasiskirstymas pagal mokslo sritis skaičiuojamas pagal tyrėjų, dirbančių atitinkamoje mokslo srityje, MVDDA ir remiasi pastarojo LMT lyginamojo vertinimo rezultatais. Tikslių duomenų apie MTEP projektinio finansavimo bei doktorantų pasiskirstymą pagal mokslo sritis Lietuvoje nėra.

Tyrėjų mokslo sritis	MVDDA	Dalis, proc.
GTM	2021,4	51 %
HSM	1336,1	34 %
Medicinos ir sveikatos mokslai	358,29	9 %
Žemės ūkio mokslai	249,82	6 %
	3965,6	

25 pav. Bendro tyrėjų skaičiaus (MVDDA) pasiskirstymas pagal mokslo sritis, %

Mokslo sritis	LMT konkursinis finansavimas	Horizon Europe	ESF fondai
GTM	55–60 %	70–75 %	60–65 %
HSM, menai	20–25 %	10–12 %	10–15 %
Medicinos ir sveikatos mokslai	15–18 %	10–15 %	10–12 %
Žemės ūkio ir aplinkos mokslai	7–10 %	5–8 %	10–12 %

26 pav. MTEP projektinio finansavimo pasiskirstymas pagal mokslo sritis, %

Doktorantūros mokslų sritis	Dalis, proc.
GTM	40–45 %
HSM, menai	25–30 %
Medicinos ir sveikatos mokslai	15–18 %
Žemės ūkio mokslai	5–7 %

27 pav. Bendro doktorantų skaičiaus pasiskirstymas pagal mokslo sritis, %

Projekto laikotarpiu numatomos skirti 5,61 mn Eur sumos paskirstymas tarp mokslo sričių investicijoms įgyvendinti pateikiamas 28 lentelėje remiantis pirmiau aptartais trimis kriterijais.

Mokslo sritis	Tyrėjų dalis, MVDDA, %	Doktorantų dalis, %	MTEP projektų dalis, %	Normalizuotas svoris	Paskirstymas, mn Eur
GTM	50	45	55	50	2,80
HSM, menai	35	30	20	27	1,52
Medicinos ir sveikatos mokslai	9	18	154	16	0,90
Žemės ūkio ir aplinkos mokslai	6	7	10	7	0,38
	100	100	100	100	5,61

28 pav. Investicijų paskirstymas tarp mokslo sričių, mln. Eur

Taigi GTM mokslo sričiai rekomenduojama skirti 2,8 mln. Eur nuo visoms mokslo sritims finansavimui skirtos 5,61 mln. Eur sumos. HSM mokslo sričiai rekomenduojama skirti 1,52 mln. Eur, medicinos ir sveikatos mokslų sričiai – 0,9 mln. Eur, žemės ūkio ir aplinkos mokslų sričiai – 0,38 mln. Eur nuo finansavimui skirtos 5,61 mln. Eur sumos.

Sukurtos infrastruktūros išlaikymo po projekto pabaigos potencialios metinės išlaidos

Finansinių skaičiavimų prielaidos:

- Nacionalinio mazgo ir federacinės sistemos išlaikymo sąnaudos po projekto pabaigos (neatsižvelgiant į infliaciją)
 - Centrinis nacionalinio mazgo palaikymas – 100,0 tūkst. Eur.
 - Administruojančios struktūros valstybės lygmeniu palaikymas:
 - Numatomas darbuotojų etatų skaičius – 4 (FTE).
 - Metinės darbo užmokesčio sąnaudos, atsižvelgiant į infliaciją – 45814 Eur/m.
 - Bendra suma per metus – 247,4 tūkst. Eur.
 - MSI ir MTI institucijų duomenų valdymo struktūrų palaikymas:
 - DS skaičius per metus – nuo 40 iki 160 MVDDA (FTE) (priklausomai nuo duomenų kiekių augimo).
 - Metinės darbo užmokesčio sąnaudos, atsižvelgiant į infliaciją – 51 900 Eur/m.
 - Patalpų išlaikymo išlaidos, 5 % nuo darbo užmokesčio, suma, Eur.
 - Pridėtinės išlaidos, 30 % nuo darbo užmokesčio, suma, Eur.

Apskaičiuotos sukurtos infrastruktūros išlaikymo po projekto pabaigos potencialios metinės išlaidos:

- kai duomenų valdytojų skaičius yra 160 MVDDA (FTE), **bendra suma per metus** – 11 210,4 tūkst. Eur/metus.

III. Lietuvos įsijungimo į EOSC alternatyva – regioninio EOSC mazgo sukūrimo alternatyva

Prielaidos regioninio EOSC mazgo sukūrimo alternatyvai vertinti

Lietuva į EOSC integruotųsi ne per savarankiškai organizuojamą nacionalinį mazgą, o per bendrą kelių valstybių regioninį mazgą. Toks modelis suteikia galimybę integruotis į platesnio masto tarpvalstybinę MTD erdvę, dalytis techninėmis ir organizacinėmis sąnaudomis, derinti regioninius standartizavimo sprendimus ir jungti kelių šalių mokslo bendruomenes į bendrą sąveikią aplinką.

Modelis tam tikrais atvejais galėtų leisti sumažinti dubliavimąsi, sukurti bendrus infrastruktūrinius pajėgumus ir padidinti regiono kaip visumos matomumą EOSC federacijoje.

Numatomos investicijų finansavimo kryptys:

1. Investicijos į institucijų valdomas ar centralizuotai palaikomas MTD saugyklas, metaduomenų atitikčių sudarymo sprendimai, MTD aprašymo, sąveikumo ir konsultavimo veiklos, teisiniai, reguliaciniai sprendimai institucijų lygmeniu.
2. Administruojančios struktūros valstybės lygmeniu sukūrimas.
3. Investicijos į duomenų valdytojų sistemos bei duomenų valdytojų tinklo sukūrimą ir nuolatinį palaikymą.

Šio finansavimo modelio atveju pagrindinė investicijų našta būtų perkeliama po galimo projekto finansavimo pabaigos. Pagrindinė našta tektų MSI savarankiškai finansuoti įsijungimą į EOSC. Alternatyvos įgyvendinimas truktų ilgiau nei trejus metus. Dėl šios alternatyvos įgyvendinimo sėkmės atsiranda rizikų.

Finansavimo modelį būtų galima svarstyti, jei pradinei investicijai trejų metų laikotarpiu finansavimas nebūtų skiriamas.

7. Literatūra

- Aksenova, A., Johnny, A., Adams, T., Gribbon, P., Jacobs, M., & Hofmann-Apitius, M. (2024). Current state of data stewardship tools in life science. *Frontiers in Big Data*, 7, Article 1428568. <https://doi.org/10.3389/fdata.2024.1428568>
- Ammar, A., Bonaretti, S., Winckers, L., Quik, J., Bakker, M., Maier, D., Lynch, I., van Rijn, J., & Willighagen, E. (2020). A semi-automated workflow for FAIR maturity indicators in the Life Sciences. *Nanomaterials*, 10(10), Article 2068. <https://doi.org/10.3390/nano10102068>
- Antol, M., Marek, J., Capandova, M., Juracek, J., & Matyska, L. (2024). EOSC CZ: Towards the development of Czech national ecosystem for FAIR research data. *arXiv preprint arXiv:2402.13343*. <https://doi.org/10.48550/arXiv.2402.13343>
- Bailo, D., Paciello, R., Sbarra, M., Rabissoni, R., Vinciarelli, V., & Cocco, M. (2020). Perspectives on the implementation of FAIR principles in solid earth research infrastructures. *Frontiers in Earth Science*, 8. <https://doi.org/10.3389/feart.2020.00003>
- Bite, K., Daugavietis, J., Kampars, J., Kreicbergs, J., Kučma, I., Ločmele, E., Vecpuise, E., Ostrovska, D., Veisa, K., & Želve, M. (2020). *Pētījuma par atvērto zinātni un rīcībpolitikas ceļa kartes izstrādi*. Ministry of Education and Science Republic of Latvia. https://www.izm.gov.lv/sites/izm/files/petijums-atverta_zinatne_21_2.pdf.
- Cabrera, N. B., Bongartz, E. C., Dörrenbächer, N., Goebel, J., Kaluza, H., & Siegers, P. (2020). White Paper on implementing the FAIR principles for data in the Social, Behavioural, and Economic Sciences. *RatsWD Working Paper 274*. German Data Forum (RatsWD). <https://doi.org/10.17620/02671.60>
- Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J.D., Anderson, J., & Hudson, M. (2020). The CARE principles for indigenous data governance. *Data Science Journal*, 19(1), Article 43. <https://doi.org/10.5334/dsj-2020-043/>
- Celebi, R., Moreira, J. R., Hassan, A. A., Ayyar, S., Ridder, L., Kuhn, T., & Dumontier, M. (2020). Towards FAIR protocols and workflows: The OpenPREDICT use case. *PeerJ Computer Science*, 6, Article e281. <https://doi.org/10.7717/peerj-cs.281>
- Chue Hong, N. P., Katz, D. S., Barker, M., Lamprecht, A.-L., Martinez, C., Psomopoulos, F. E., Harrow, J., Castro, L. J., Gruenpeter, M., Martinez, P. A., Honeyman, T., Struck, A., Lee, A., Loewe, A., van Werkhoven, B., Jones, C., Garijo, D., Plomp, E., Genova, F., ... RDA FAIR4RS WG. (2022). *FAIR principles for research software (FAIR4RS principles)*. Research Data Alliance. <https://zenodo.org/records/6623556>.
- Clare, H. (2026, Jan. 29). *Thematic track 3: Competences and training rapporteur*. [Oral presentation]. EOSC Winter School. Nice, France.
- Clarke, D. J. B., Wang, L., Jones, A., Wojciechowicz, M. L., Torre, D., Jagodnik, K. M., Jenkins, S. L., McQuilton, P., Flamholz, Z., Silverstein, M. C., Schilder, B. M., Robasky, K., Castillo, C., Idaszak, R., Ahalt, S. C., Williams, J., Schurer, S., Cooper, D. J., Azevedo, R. de M., ... Ma'ayan, A. (2019). FAIRshake: Toolkit to evaluate the FAIRness of research digital resources. *Cell Systems*, 9(5), 417–421. <https://doi.org/10.1016/j.cels.2019.09.011>
- CoreTrustSeal Board. (2016). *Core trustworthy data repositories requirements (v01.00)*. https://www.coretrustseal.org/wp-content/uploads/2017/01/Core_Trustworthy_Data_Repositories_Requirements_01_00.pdf
- Corleto, A., Di Giorgio, S., Paolini, G., Berberi, L., Candela, L., Costantini, A., Gaido, L., Green, D., JANIK, J., Tuminauskas, R., Whyte, A., Lazzeri, E., Prandoni, C., & Evangelinou, B. (2025). *D7.3 Report on CCs and user support networks and recommendations for networks evolution*. [Project deliverable]. Skills4EOSC - Skills for the European Open Science commons. <https://doi.org/10.5281/zenodo.15262091>
- Council of the EU. (2025). *Council conclusions - Towards the EU strategy on artificial intelligence in science – ES strategijos dėl dirbtinio intelekto naudojimo moksle kūrimas – Tarybos išvados, patvirtintos 2025 m. gegužės 23 d.* Brussels, 23 May 2025 (OR. en) 9292/25.
- Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and conducting mixed methods research* (3rd ed.). Sage.
- da Silva Santos, L. O. B. D. S., Sales, T. P., Fonseca, C. M., & Guizzardi, G. (2023a). Towards a conceptual model for the FAIR Digital Object Framework. In N. Aussenac-Gilles, T. Hahmann, A. Galton, & M. M. Hedblom (Eds.), *Formal ontology in information systems, FOIS 2023* (Vol. 377, pp. 227–241). Ios Press. <https://doi.org/10.3233/FAIA231131>
- da Silva Santos, L. O. B., Burger, K., Kaliyaperumal, R., & Wilkinson, M.D. (2023b). FAIR data point: a FAIR-oriented approach for metadata publication. *Data Intelligence*, 5, 163–183. https://doi.org/10.1162/dint_a_00160
- DataCite & Make Data Count. (2025). *Data citation corpus data file (v4.1)* [Data set]. DataCite. <https://doi.org/10.5281/zenodo.16901115>.
- Demchenko, Y., Belloum, A., & Wiktorski, T. (2017). *Edison Data Science Framework: Part 1. Data Science Competence Framework (Cf-Ds) Release 2*. <https://doi.org/10.5281/ZENODO.1044345>
- Demchenko, Y., & Stoy, L. (2021). Research data management and data stewardship competences in university curriculum. *IEEE Global Engineering Education Conference, EDUCON, 2021-April*, 1717–1726. <https://doi.org/10.1109/EDUCON46332.2021.9453956>
- Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on Open Data and the re-use of public sector information (recast) = Europos Parlamento ir Tarybos direktyva (ES) 2019/1024, 2019 m. birželio 20 d. dėl atvirųjų duomenų ir viešojo sektoriaus informacijos pakartotinio naudojimo (nauja redakcija). <http://data.europa.eu/eli/dir/2019/1024/oj>

- Doldirina, C., Eisenstadt, A., Onsrud, H., & Uhler, P. (2018). *Legal approaches for open access to research data*. <https://doi.org/10.31228/osf.io/n7gfa>
- Dyson, R. G. (2004). Strategic development and SWOT analysis at the University of Warwick. *European Journal of Operational Research*, 152(3), 631–640. [https://doi.org/10.1016/S0377-2217\(03\)00062-6](https://doi.org/10.1016/S0377-2217(03)00062-6)
- EOSC Association. (2025). *EOSC Federation Handbook (v1.0)*. Zenodo. <https://doi.org/10.5281/zenodo.14999577>.
- EOSC Association. (2026). *EOSC Federation Handbook (v2.0)*. Zenodo. <https://doi.org/10.5281/zenodo.18454649>.
- European Commission. (2016a). *Guidelines on FAIR data management in Horizon 2020. Version 3.0*. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
- European Commission. (2016b). *Open innovation, open science, open to the world: A vision for Europe*. Publications Office of the European Union. <https://op.europa.eu/en/publication-detail/-/publication/3213b335-1c6c-11e6-ba9a-01aa75ed71a1/language-en>
- European Commission. (2017). *ANNEX to the Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. European Interoperability Framework - Implementation strategy*. = Komisijos komunikatas Europos Parlamentui, Tarybai, Europos Ekonomikos ir socialinių reikalų komitetui ir Regionų komitetui. Europos sąveikumo sistema. Įgyvendinimo strategija. Brussels, 23.3.2017 COM(2017) 134 final ANNEX 2.
- European Commission. (2018). *Commission Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information*. = Komisijos rekomendacija (ES) 2018/790 2018 m. balandžio 25 d. dėl prieigos prie mokslinės informacijos ir tos informacijos išsaugojimo. <http://data.europa.eu/eli/reco/2018/790/oj>.
- European Commission. (2019). *Cost of not having FAIR research data. Cost-benefit analysis for FAIR research data [Report]*. Publications Office of the European Union. <https://doi.org/10.2777/02999>
- European Commission. (2020). *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European strategy for data*. = Komisijos komunikatas Europos Parlamentui, Tarybai, Europos Ekonomikos ir socialinių reikalų komitetui ir Regionų komitetui. Europos duomenų strategija. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066>.
- European Commission. (2021a). *EOSC Interoperability Framework. Report from the EOSC Executive Board working groups FAIR and Architecture*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2777/620649>
- European Commission. (2021b). *Digital skills for FAIR and Open Science. Report from the EOSC Executive Board Skills and Training working group*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2777/59065>.
- European Commission. (2024). *Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC) (v1.3)*. 10.5281/zenodo.17551218.
- European Commission. (2025a). *AGA – Annotated Grant Agreement: EU funding programmes 2021–2027. Version 2.0*. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/aga_en.pdf.
- European Commission. (2025b). *Living guidelines on the responsible use of generative AI in research*. Publications Office of the European Union. https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en
- European Commission. (2025c). *Communication from the Commission to the European Parliament and the Council. A European strategy for artificial intelligence in science. Paving the way for the Resource for AI Science in Europe (RAISE)* = Komisijos komunikatas Europos Parlamentui ir Tarybai Europos dirbtinio intelekto naudojimo moksle strategija. Klojami Europos DI mokslo išteklių centro (RAISE) pamatai. Brussels, 8.10.2025, COM(2025) 724 final.
- European Commission. (2025d). *Proposal for a Regulation of the European Parliament and of the Council amending Regulations (EU) 2016/679, (EU) 2018/1724, (EU) 2018/1725, (EU) 2023/2854 and Directives 2002/58/EC, (EU) 2022/2555 and (EU) 2022/2557 as regards the simplification of the digital legislative framework, and repealing Regulations (EU) 2018/1807, (EU) 2019/1150, (EU) 2022/868, and Directive (EU) 2019/1024 (Digital Omnibus)* = Pasiūlymas Europos Parlamento ir Tarybos Reglamentas, kuriuo siekiant supaprastinti skaitmeninio sektoriaus teisės aktų sistemą iš dalies keičiami reglamentai (ES) 2016/679, (ES) 2018/1724, (ES) 2018/1725, (ES) 2023/2854 ir direktyvos 2002/58/EB, (ES) 2022/2555 ir (ES) 2022/2557 ir panaikinami reglamentai (ES) 2018/1807, (ES) 2019/1150, (ES) 2022/868 ir Direktyva (ES) 2019/1024, (Skaitmeninio sektoriaus bendrasis rinkinys). Brussels, 19.11.2025, COM(2025) 837 final, 2025/0360(COD).
- European Commission. (2025e). *Communication from the Commission to the European Parliament and the Council. Data Union Strategy: Unlocking data for AI* = Komisijos komunikatas Europos Parlamentui ir Tarybai. Duomenų sąjungos strategija. Prieigos prie DI reikalingų duomenų atvėrimas. Brussels, 19.11.2025 COM(2025) 835 final.
- European Commission. (2025f). *ANNEX to the Proposal for a Council recommendation on the European Research Area Policy Agenda 2025–2027* = Priedas prie pasiūlymo dėl Tarybos rekomendacijos dėl 2025–2027 m. Europos mokslinių tyrimų erdvės politikos darbotvarkės. Brussels, 3 March 2025 (OR. en) 6693/25 ADD 1
- European Commission Joint Research Centre. (2018). *DigComp into action, get inspired make it happen: A user guide to the European Digital Competence framework*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2760/112945>.
- European University Association. (2022). *The EUA Open Science agenda 2025*. European University Association asbl. <https://www.eua.eu/publications/positions/the-eua-open-science-agenda-2025.html>

- Frey, J., & Hellmann, S. (2021). FAIR linked data — towards a linked data backbone for users and machines. *Companion Proceedings of the Web Conference 2021 (WWW '21)*, 431–435. <https://doi.org/10.1145/3442442.3451364>
- Gaia-X European Association for Data and Cloud. (2024). *An introduction to the Gaia-X Trust Framework (V4)*. https://gaia-x.eu/wp-content/uploads/2024/05/An-Introduction-to-the-Gaia-X-Trust-Framework_2024-V4.pdf
- Gaignard, A., Rosnet, T., De Lamotte, F., Lefort, V., & Devignes, M.-D. (2023). FAIR-Checker: Supporting digital resource findability and reuse with Knowledge Graphs and Semantic Web standards. *Journal of Biomedical Semantics*, 14(1), Article 7. <https://doi.org/10.1186/s13326-023-00289-5>
- GEO Data Sharing Working Group (GEO DSWG). 2014. *White paper: Mechanisms to share data as part of GEOSS Data-CORE*. https://www.earthobservations.org/documents/dswg/Annex%20V1%20-%20%20Mechanisms%20to%20share%20data%20as%20part%20of%20GEOSS%20Data_CORE.pdf
- Goble, C., Cohen-Boulakia, S., Soiland-Reyes, S., Garijo, D., Gil, Y., Crusoe, M. R., Peters, K., & Schober, D. (2020). FAIR computational workflows. *Data Intelligence*, 2(1-2), 108–121. https://doi.org/10.1162/dint_a_00033
- Gruber, A., Schranzhofer, H., Knopper, S., Stryeck, S., & Hasani-Mavriqi, I. (2021). Kompetenzen von Data Stewards an österreichischen Universitäten. *Mitteilungen Der Vereinigung Österreichischer Bibliothekarinnen Und Bibliothekare*, 74(1), 12–32. <https://doi.org/10.31263/voebm.v74i1.6255>
- Hansen, K. K., White, A., Horton, L., Koteska, B., Rauste, P., & van Leersum, N. (2025). *Data steward: Minimum viable skills profile*. Zenodo. <https://doi.org/10.5281/zenodo.16760455>
- Huber, R. (2025). *FAIRsFAIR data object assessment metrics (08)*. Zenodo. <https://doi.org/10.5281/zenodo.15045911>
- Hugo, W., Thorpe, D., & Katsanakis, M. (2025). *EOSC beyond D13.3 digital rights management state of the art*. <https://doi.org/10.5281/zenodo.17225379>
- Jetten, M., Grootveld, M., Mordant, A., Jansen, M., Bloemers, M., Miedema, M., & Van Gelder, C. W. G. (2021). *Professionalising data stewardship in the Netherlands. Competences, training and education. Dutch roadmap towards national implementation of FAIR data stewardship (1.1)*. Zenodo. <https://doi.org/10.5281/zenodo.4623713>
- Jones, S., & Grootveld, M. (2017). *How FAIR are your data?* <https://doi.org/10.5281/zenodo.1065991>
- Kaur, P., Kaushik, A., & Kapoor, A. (2023). Skills and responsibilities of data wrangler. In M. Niranjanamurthy, K. Sheoran, G. Dhand, & P. Kaur (Eds.), *Data wrangling: Concepts, applications and tools* (pp. 19–51). <https://doi.org/10.1002/9781119879862.ch2>
- Kleemola, M., & Wong, M. W. (2025). *The battle for trust: Advancing the social science data ecosystem*. Zenodo: CESSDA. <https://doi.org/10.5281/zenodo.15125538>
- Krisnawijaya, N. N. K., Tekinerdogan, B., Catal, C., van der Tol, R., & Herdiyeni, Y. (2025). Implementing FAIR principles in data management systems: A multi-case study in precision farming. *Computers and Electronics in Agriculture*, 230, Article 109855. <https://doi.org/10.1016/j.compag.2024.109855>
- Lammy, R. (2019). How publishers can work with Crossref on data citation. *Science Editing* 6(2), 166–170. <https://doi.org/10.6087/kcse.165>
- Lamprecht, A.-L., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Martin Del Pico, E., Dominguez Del Angel, V., van de Sandt, S., Ison, J., Martinez, P. A., McQuilton, P., Valencia, A., Harrow, J., Psomopoulos, F., Gelpi, J. L., Chue Hong, N., Goble, C., & Capella-Gutierrez, S. (2019). Towards FAIR principles for research software. *Data Science*, 3(1), 37–59. <https://doi.org/10.3233/DS-190026>
- Lehväslaiho, H., Parland-von Essen, J., Behnke, C., Laine, H., Riungu-Kalliosaari, L., Le Franc, Y., & Staiger, C. (2019). *D2.1 Report on FAIR requirements for persistence and interoperability (v1.0 Draft)*. FAIRsFAIR. <https://doi.org/10.5281/zenodo.3557381>
- Lietuvos mokslo taryba. (2016). *Atvirosios prieigos prie mokslo publikacijų ir duomenų gairės, 2016 m. vasario 29 d. Nr.VIII-2*. <https://www.e-tar.lt/portal/lt/legalAct/dceeeb10e05711e59cc8b27b54efaf6e>
- Lietuvos mokslo taryba. (2024a). *Atvirosios prieigos prie mokslinių tyrimų ir eksperimentinės plėtros rezultatų tvarkos aprašas, patv. 2024 m. rugpjūčio 30 d. įsakymu Nr. V-385*. <https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/a1a39ae1696211efad96fef38de4d94a>
- Lietuvos mokslo taryba. (2024b). *Lietuvos mokslinių tyrimų infrastruktūrų kelrodis*. Vilnius: LĮ „Kriventa“. <https://lmt.lrv.lt/public/canonical/1735535380/3894/LMT%20infrastrukturu%20kelrodis%202024.pdf>
- Lietuvos Respublikos Seimas. (2000). *Lietuvos Respublikos teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatymas, Valstybės žinios, 2000-02-02, Nr. 10-236*. <https://www.e-tar.lt/portal/lt/legalAct/TAR.FA13E28615F6/asr>
- Lietuvos Respublikos Seimas. (2009). *Lietuvos Respublikos mokslo ir studijų įstatymas, Nr. XI-242*. <https://www.e-tar.lt/portal/lt/legalAct/TAR.C595FF45F869/OGxFOEfsdd>
- Lietuvos Respublikos Vyriausybė. (2020). *2021–2030 metų nacionalinis pažangos planas*. <https://lr.lt/uploads/main/documents/files/NPP%20planas.pdf>
- Lietuvos Respublikos Vyriausybė. (2021). *Nutarimas. Dėl Aštuonioliktosios Lietuvos Respublikos Vyriausybės programos nuostatų įgyvendinimo plano patvirtinimo*. <https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/bef7d43286fe11eb998483d0ae31615c?jfwid=n4udyawj6>

- Lin, D., Crabtree, J., Dillo, I., Downs, R. R., Edmunds, R., Giaretta, D., De Giusti, M., L'Hours, H., Hugo, W., Jenkyns, R., Khodiyar, V., Martone, M. E., Mokrane, M., Navale, V., Petters, J., Sierman, B., Sokolova, D. V., Stockhause, M. & Westbrook, J. (2020). The TRUST Principles for digital repositories. *Sci Data*, 7, Article 144. <https://doi.org/10.1038/s41597-020-0486-7>
- L'Hours, H., Kleemola, M., Parkes, O., Recker, J., Duvaud, S., van Horik, R., Alaterä, T. J., Liberante, F., Conzett, P., Kaartinen, H., Bäckman, S., & Esteves, E. (2025). *FIDELIS TTRAMatrix v01.00 Guide (v01.00)*. Zenodo. <https://doi.org/10.5281/zenodo.17144141>
- McDowall, R. D. (2017). Understanding data governance, Part II. *Spectroscopy*, 32(4), 12–18.
- McMurry, J. A., Juty, N., Blomberg, N., Burdett, T., Conlin, T., Conte, N., Courtot, M., Deck, J., Dumontier, M., Fellows, D. K., Gonzalez-Beltran, A., Gormanns, P., Grethe, J., Hastings, J., Hériché, J.-K., Hermjakob, H., Ison, J. C., Jimenez, R. C., Jupp, S., ... Parkinson, H. (2017). Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data. *PLOS Biology*, 15(6), 1–18. <https://doi.org/10.1371/journal.pbio.2001414>
- Memorandum of Understanding on Preparing Operational Integration within the Envisaged EOSC Federation (n.d.). https://eosc.eu/wp-content/uploads/2025/10/20251103_MoU_Operational-Preparation_EOSC-Federation.pdf
- Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L. O. B., & Wilkinson, M. D. (2017). Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud. *Information Services and Use*, 37(1), 49–56. <https://doi.org/10.3233/ISU-170824>
- Mons, B., Schultes, E., Liu, F., & Jacobsen, A. (2020). The FAIR principles: First generation implementation choices and challenges. *Data Intelligence*, 2(1–2), 1–9. https://doi.org/10.1162/dint_e_00023
- Murray, J. D., & vanRyper, W. (1996). *Encyclopedia of graphics file formats* (2nd. ed.). O'Reilly & Associates.
- National Open Research Forum. (2022, November 18). *National action plan for open research*. Digital Repository of Ireland. <https://doi.org/10.7486/DRI.ff36jz222>
- Neumann, J., & Brase, J. (2014). DataCite and DOI names for research data. *Journal of Computer-Aided Molecular Design*, 28(10), 1035–1041. <https://doi.org/10.1007/s10822-014-9776-5>
- Noble, H., & Heale, R. (2019). Triangulation in research, with examples. *Evidence-Based Nursing*, 22(3), 67–68. <https://doi.org/10.1136/ebnurs-2019-103145>
- O'Brien, M., Duerr, R., Taitingfong, R., Martinez, A., Vera, L., Jennings, L. L., Downs, R. R., Antognoli, E., ten Brink, T., Halmaj, N. B., Carroll, S. R., David-Chavez, D., Hudson, M., & Buttigieg, P. L. (2024). Earth science data repositories: Implementing the CARE principles. *Data Science Journal*, 23(37), 1–29. <https://doi.org/10.5334/dsj-2024-037>
- OECD. (2020). Building digital workforce capacity and skills for data-intensive science. *OECD Science, Technology and Industry Policy Papers*, 90. OECD Publishing. <https://doi.org/10.1787/e08aa3bb-en>
- OECD. (2021). *Recommendation of the Council concerning access to research data from public funding*. OECD Legal Instruments. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347>
- Open Data Institute. (2020). *Data Skills Framework*. <https://theodi.org/insights/tools/data-skills-framework/>
- Palmer, F. (2014). Analytics stewardship and the community of practice. *IBM Data Management Magazine*, 7. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84906075501&partnerID=40&md5=8bc95188f00a1dc7ff4fd5c3483b91f4>
- Paulus, B., & Biskup, T. (2023). Towards more reproducible and FAIRer research data: Documenting provenance during data acquisition using the Infile format. *Digital Discovery*, 2(1), 234–244. <https://doi.org/10.1039/D2DD00131D>
- Pergl, R., Hooft, R., Suchánek, M., Knaisl, V., & Slifka, J. (2019). “Data Stewardship Wizard”: A tool bringing together researchers, data stewards, and data experts around data management planning. *Data Science Journal*, 18, Article 59. <https://doi.org/10.5334/dsj-2019-059>
- Peters, D., & Schindler, S. (2024). FAIR for digital twins. *CEAS Space Journal*, 16(3), 367–374. <https://doi.org/10.1007/s12567-023-00506-y>
- RDA FAIR Data Maturity Model Working Group. (2020). *FAIR Data Maturity Model. Specification and guidelines*. (1.0). <https://zenodo.org/records/3909563>
- Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) = Europos Parlamento ir Tarybos direktyva (ES) 2022/868 2022 m. gegužės 30 d. dėl Europos duomenų valdymo, kuriuo iš dalies keičiamas Reglamentas (ES) 2018/1724 (Duomenų valdymo aktas). <http://data.europa.eu/eli/reg/2022/868/oj>
- Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) = Europos Parlamento ir Tarybos reglamentas (ES) 2022/2065 2022 m. spalio 19 d. dėl bendrosios skaitmeninių paslaugų rinkos, kuriuo iš dalies keičiama Direktyva 2000/31/EB (Skaitmeninių paslaugų aktas) <http://data.europa.eu/eli/reg/2022/2065/oj>
- Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act) (Text with EEA relevance) = Europos Parlamento ir Tarybos direktyva (ES) 2023/2854 2023 m. gruodžio 13 d. dėl suderintų sąžiningos prieigos prie duomenų ir jų naudojimo taisyklių, kuriuo iš dalies keičiamas Reglamentas (ES) 2017/2394 ir Direktyva (ES) 2020/1828 (Duomenų aktas). <http://data.europa.eu/eli/reg/2023/2854/oj>

- Rinaldi, L., Bonacorsi, D., Bozzi, C., Pra, S. D., Dell'Agnello, L., Gaido, L., Marchegiani, F., & Piergentili, I. (2025). So FAIR, so good: The INFN strategy for data stewardship. *EPJ Web of Conferences*, 337. <https://doi.org/10.1051/epjconf/202533701140>
- Scardaci, D. O., Sciacca, E., Hériché, J.-K., Van De Sanden, M., Wierenga, K., Manghi, P., Tamburri, D., Mazon, J. N., López García, Á., Hugo, W., Pansanel, J., & Krøl Andersen, L. (2023). *A landscape overview of the EOSC Interoperability Framework – Capabilities and gaps*. Zenodo. <https://doi.org/10.5281/zenodo.8399710>
- Scholtens, S., Jetten, M., Böhmer, J., Staiger, C., Slouwerhof, I., Geest, M. V. D., & Gelder, C. W. G. (2019). *Final report: Towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise (Version 1)*. Zenodo. <https://doi.org/10.5281/ZENODO.3471707>
- Schultes, E., Magagna, B., Hettne, K. M., Pergl, R., Suchanek, M., & Kuhn, T. (2020). Reusable FAIR implementation profiles as accelerators of FAIR convergence. In G. Grossmann, & S. Ram (Eds.), *Advances in Conceptual Modeling, ER 2020* (Vol. 12584, pp. 138–147). Springer. https://doi.org/10.1007/978-3-030-65847-2_13
- Singh, N. K., Wang, S., Maineri, A., Hofstra, T., Bruyneel, M., van de Sandt, S., Siebes, R., van Ossenbruggen, J., & Kuhn, T. (2025). Aligning data management plans with community standards using FAIR implementation profiles. In M. Sfakakis, E. Garoufallou, M. Damigos, A. Salaba, & C. Papatheodorou (Eds.), *Metadata and semantic research, MTSR 2024* (Vol. 2331, pp. 159–171). Springer. https://doi.org/10.1007/978-3-031-81974-2_14
- Stueber, M., & Frey, G. (2022). FAIRness for models and simulations. *ATP Magazine*, 10, 58–65.
- Sylvia, M. L., & Terhaar, M. F. (2023). Data governance and stewardship. In M. L. Sylvia, & M. F. Terhaar (Eds.), *Clinical analytics and data management for the DNP* (3rd ed.) (pp. 363–371). <https://doi.org/10.1891/9780826163240.0018>
- Tautkevičienė, G., Gribauskienė, A., Kepalienė, F., Kiršienė, J., Kupčinskas, L., Medzvieckaitė, G., Petrauskaitė, R., & Umbrasaitė, J. (2022). *Atviroslausios prieigos prie mokslo publikacijų ir duomenų gairių taikymo praktikos ir pasiektos pažangos įvertinimas: studija*. Lietuvos mokslo taryba. <https://lmt.lrv.lt/media/viesa/saugykla/2023/10/gT5VZIFkVE.pdf>
- Thompson, C. A. (2018). Cultivating data expertise: Learning strategies & conditions for research institutions. In L. Freund (Ed.), *Proceedings of the Association for Information Science and Technology* (pp. 514–523). Wiley. <https://doi.org/10.1002/pr2.2018.14505501056>
- van Reisen, M., Stokmans, M., Basajja, M., Otieno, O., Kirkpatrick, C., & Mons, B. (2020). Towards the tipping point for FAIR implementation. *Data Intelligence*, 2(1–2), 264–275. https://doi.org/10.1162/dint_a_00049
- Wendelborn, C., Anger, M., & Schickhardt, C. (2023). What is data stewardship? Towards a comprehensive understanding. *Journal of Biomedical Informatics*, 140, Article 104337. <https://doi.org/10.1016/j.jbi.2023.104337>
- Wildgaard, L., Vlachos, E., Nondal, L., Larsen, A. V., & Svendsen, M. (2020). *National coordination of data steward education in Denmark: Final report to the National Forum for research data management (DM Forum)* (Version 1). <https://doi.org/10.5281/zenodo.3609516>
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Santos, L. B. da S., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., ... Mons, B. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3, Article 160018. <https://doi.org/10.1038/sdata.2016.18>
- Wilkinson, M. D., Sansone, S.-A., Schultes, E., Doorn, P., Bonino da Silva Santos, L. O., & Dumontier, M. (2018). A design framework and exemplar metrics for FAIRness. *Scientific Data*, 5(1), Article 180118. <https://doi.org/10.1038/sdata.2018.118>
- Wilkinson, M. D., Dumontier, M., Sansone, S.-A., Bonino da Silva Santos, L. O., Prieto, M., Batista, D., McQuilton, P., Kuhn, T., Rocca-Serra, P., Crosas, M., & Schultes, E. (2019). Evaluating FAIR maturity through a scalable, automated, community-governed framework. *Scientific Data*, 6(1), Article 174. <https://doi.org/10.1038/s41597-019-0184-5>
- Wolf, M., Logan, J., Mehta, K., Jacobson, D., Cashman, M., Walker, A. M., Eisenhauer, G., Widener, P., & Cliff, A. (2021). Reusability first: Toward FAIR workflows. *2021 IEEE International Conference on Cluster Computing (CLUSTER 2021)*, 444–455. <https://doi.org/10.1109/Cluster48925.2021.00053>

8. Priedai

1 PRIEDAS

Lietuvos mokslo tarybos kreipimosi laiškas į institucijas, kviečiant užpildyti institucijoms skirtą klausimyną



LIETUVOS MOKSLO TARYBA

Biudžetinė įstaiga, Gedimino pr. 3, 01103 Vilnius, tel. (+370) 670 32435, el. p. info@lmt.lt,
elektroninio pristatymo dėžutės adresas 188716281.

Duomenys kaupiami ir saugomi Juridinių asmenų registre, kodas 188716281

Pagal adresatų sąrašą

2025-04-

Nr.

DĖL KLAUSIMYNO PILDYMO

DĖL KLAUSIMYNO PILDYMO

Lietuvos mokslo taryba (toliau – Taryba), įgyvendindama Lietuvos Respublikos švietimo, mokslo ir sporto ministerijos mokslo plėtros programos pažangos priemonę Nr. 12-001-01-01-01 „Gerinti mokslo ir studijų aplinką“, veiklą „Įsijungti į Europos atvirojo mokslo debesį“, projekto kodas 10-073-P-0001 (toliau – projektas), rengia galimybių studiją „Įsijungti į Europos atvirojo mokslo debesį – galimybių studijos dėl būtinų pokyčių ir investicijų krypčių parengimas“ (toliau – studija), daugiau apie projektą: <https://lmt.lrv.lt/lt/lmt-projektai/es-projektai/>.

Projektu siekiama spręsti mokslinių tyrimų duomenų infrastruktūros fragmentacijos, nevienodų metaduomenų standartų ir duomenų valdymo kompetencijų stokos problemas. Prisijungimas

prie Europos atvirojo mokslo debesies (toliau – EOSC) sudarys sąlygas naudotis bendra Europos skaitmenine infrastruktūra, didins Lietuvos tyrimų duomenų matomumą ir plės tarptautinio bendradarbiavimo galimybes.

Studijoje bus analizuojama Lietuvos mokslinių tyrimų duomenų infrastruktūra, mokslo ir studijų institucijų (toliau – institucijos) kompetencijos bei duomenų valdymo praktikos. Remiantis analizės rezultatais bus parengtos išvados ir rekomendacijos, dėl būtinų pokyčių ir investicijų krypčių, reikalingų Lietuvai integruojantis į EOSC.

Prašome jūsų institucijos iki 2025 m. gegužės 16 d. (imtinai) užpildyti pirminį elektroninį klausimyną (nuoroda: <https://forms.office.com/e/ewDm5DxQ8j>). Suprantame, kad didesnėse institucijose gali būti taikomi skirtingi duomenų valdymo ir tvarkymo principai, tačiau šiame etape siekiame tik bendro institucijų vaizdo. Konkreti institucijų praktika bus nagrinėjama vėlesniuose etapuose – išsamiose apklausose ir teminėse diskusijose, į kurias institucijas kviesime atskiru pranešimu.

Studijos rezultatai taps pagrindu investicinėms kryptims ir nacionalinėms finansavimo priemonėms, skirtoms duomenų valdymo kompetencijoms stiprinti, sąveikumui didinti ir infrastruktūrai modernizuoti. Jūsų institucijos dalyvavimas padės užtikrinti, kad būsimas finansavimas atlieptų realius institucijų poreikius ir būtų suplanuotas racionaliai.

Kilus klausimų dėl klausimyno pildymo prašome kreiptis į projekto informacijos analitikę Indrę Nazarenko, tel. +370 676 16 779, el. p. indre.nazarenko@lmt.lt.

Dėkojame už bendradarbiavimą.

Pirmininko pavaduotoja

Vaiva

Priudokienė

2 PRIEDAS

Institucijų apklausa dėl mokslinių tyrimų duomenų valdymo praktikų

Lietuvos mokslo taryba vykdo projektą „Įsijungti į Europos atvirojo mokslo debesį – galimybių studijos dėl būtinų pokyčių ir investicijų krypčių parengimas“, kurio tikslas – įvertinti Lietuvos mokslo ir studijų institucijų pasirengimą integruotis į Europos atvirojo mokslo debesį (EOSC), nustatyti būtinus pokyčius bei investicines kryptis.

Šis klausimynas skirtas surinkti informaciją apie Lietuvos mokslo ir studijų institucijų taikomas mokslinių tyrimų duomenų valdymo praktikas, institucinius pajėgumus bei poreikius. Gauti rezultatai bus naudojami tiek apibendrintai, tiek analizuojant atskirų institucijų duomenis. Apklausos rezultatai bus viešai skelbiami kaip dalis mokslinių tyrimų duomenų valdymo situacijos Lietuvoje analizės.

Dalyvavimas apklausoje yra **savanoriškas**, tačiau labai svarbus siekiant parengti išsamų, faktais grįstą vertinimą ir rekomendacijas.

Šioje apklausoje renkama informacija apie institucijų praktiką, o ne apie konkrečius asmenis. Jei pateikiami kontaktiniai duomenys (pvz., atsakingo asmens vardas ar el. paštas), jie bus naudojami tik komunikacijai, susijusiai su šia galimybių studija ir tolesniu bendradarbiavimu. Visi duomenys bus tvarkomi laikantis Bendrojo duomenų apsaugos reglamento reikalavimų.

Pastaba. *Mokslinių tyrimų duomenys* šioje apklausoje apibrėžiami, kaip skaitmeniniai ir (ar) fiziniai objektai, kurie yra renkami ar generuojami vykdant mokslinius tyrimus ir yra naudojami mokslinėms įžvalgoms gauti ir mokslinių tyrimų išvadoms bei rezultatams patvirtinti. Į šį apibrėžimą įeina ne tik laboratorinių bandymų ar aplinkos stebėjimų duomenys, bet ir įvairių apklausų ar interviu rezultatai, tekstynai, garsynai, lentelių ar kitokia forma organizuotos duomenų bazės, vaizdo ir garso įrašai, skaitmeniniai modeliai, programinis kodas ir kt.

Mokslinių tyrimų duomenų valdymas šioje apklausoje suprantamas, kaip bet kokie veiksmai atliekami su mokslinių tyrimų duomenimis po jų sukūrimo, pvz., duomenų saugojimas, metaduomenų (kontekstinės informacijos apie duomenis) kaupimas, duomenų atvėrimas ir pan. Duomenys institucijoje gali būti valdomi tiek centralizuotai, tiek individualiai, t. y. atskirų jos narių, naudojant institucijos suteiktas arba asmenines priemones.

Kilus klausimų prašome kreiptis el. paštu indre.nazarenko@lmt.lt arba telefonu +370 676 16 779

1. Įrašykite institucijos pavadinimą

2. Ar jūsų institucijoje kuriami ir valdomi (archyvuojami, sisteminami bei išsaugomi pakartotiam panaudojimui) mokslinių tyrimų duomenys?

- a) Taip
- b) Ne

3. Jei 2 klausime pasirinkote atsakymą „Taip“, nurodykite kokia apimtimi kuriami ir valdomi (archyvuojami, sisteminami ir išsaugomi pakartotiniam panaudojimui) mokslinių tyrimų duomenys?

- a) Individualių mokslininkų ar mokslininkų grupių apimtimi
- b) Visos institucijos ar atskirų jos padalinių apimtimi
- c) Bendrai su kitomis institucijomis, per bendras mokslinių tyrimų infrastruktūras ar kt.
- d) Kita (atviras)

4. Ar Jūsų institucija yra reglamentavusi mokslinių tyrimų duomenų valdymą, ar teikiate metodinę pagalbą?

- a) Taip, teikiame pagalbą mokslininkams, pvz., mokymus apie metaduomenis, duomenų valdymo planus ir kt.
- b) Taip, turime vidinius teisės aktus, skirtus mokslinių tyrimų duomenų valdymo politikai
- c) Ne
- d) Kita (atviras)

5. Ar jūsų institucija būtų suinteresuota prisidėti prie Lietuvos mokslinių tyrimų duomenų ekosistemos kūrimo ir integracijos į EOSC? Nuoroda į projektą: <https://lmt.lrv.lt/lt/lmt-projektai/es-projektai/>

- a) Taip
- b) Ne
- c) Kita (atviras)

6. Prašytume nurodyti atsakingą kontaktinį asmenį(-is), į kurį(-iuos) galėtume kreiptis, dėl papildomos informacijos suteikimo. *Pateikite asmens vardą, pavardę, el. p. bei kontaktinį telefono numerį.*

3 PRIEDAS**Lietuvos mokslo tarybos kreipimosi laiškas į institucijas dėl antrojo klausimyno užpildymo****LIETUVOS MOKSLO TARYBA**

Biudžetinė įstaiga, Gedimino pr. 3, 01103 Vilnius, tel. (+370) 670 32435, el. p. info@lmt.lt,
elektroninio pristatymo dėžutės adresas 188716281.
Duomenys kaupiami ir saugomi Juridinių asmenų registre, kodas 188716281

Pagal adresatų sąrašą

2025-07-08

Nr.

DĖL ANTROJO KLAUSIMYNO PILDYMO

Lietuvos mokslo taryba (toliau – Taryba), įgyvendindama Lietuvos Respublikos švietimo, mokslo ir sporto ministerijos mokslo plėtros programos pažangos priemonę Nr. 12-001-01-01-01 „Gerinti mokslo ir studijų aplinką“, veiklą „Įsijungti į Europos atvirojo mokslo debesį“, projekto kodas 10-073-P-0001 (toliau – projektas), rengia galimybių studiją „Įsijungti į Europos atvirojo mokslo debesį – galimybių studijos dėl būtinų pokyčių ir investicijų kryptių parengimas“ (toliau – studija), daugiau apie projektą: <https://lmt.lrv.lt/lt/lmt-projektai/es-projektai/>.

Projektu siekiama spręsti mokslinių tyrimų duomenų infrastruktūros fragmentacijos, nevienodų metaduomenų standartų ir duomenų valdymo kompetencijų stokos problemas. Prisijungimas prie Europos atvirojo mokslo debesies (toliau – EOSC) sudarys sąlygas naudotis bendra Europos skaitmenine infrastruktūra, didins Lietuvos tyrimų duomenų matomumą ir plės tarptautinio bendradarbiavimo galimybes.

Studijoje bus analizuojama Lietuvos mokslinių tyrimų duomenų infrastruktūra, mokslo ir studijų institucijų (toliau – institucijos) kompetencijos bei duomenų valdymo praktikos. Remiantis analizės rezultatais bus parengtos išvados ir rekomendacijos, dėl būtinų pokyčių ir investicijų kryptių, reikalingų Lietuvai integruojantis į EOSC.

Nuoširdžiai dėkojame, kad sutikote dalyvauti tęstinėje institucijų apklausoje ir taip prisidėti prie galimybių studijos.

Prašome jūsų institucijos iki 2025 m. rugsėjo 15 d. (imtinai) užpildyti pirminį elektroninį klausimyną (nuoroda: <https://forms.office.com/e/igWWjp8DWh>).

Suprantame, kad klausimynas yra sudėtingas ir reikalauja laiko, koordinacijos bei vidinių resursų. Daugeliu atvejų tam gali prireikti institucinio bendradarbiavimo, papildomos informacijos paieškos ar net inventorizacijos. Tačiau net ir apytiksliai preliminarūs atsakymai mums yra ypač svarbūs – jie leis išryškinti tiek turimus pajėgumus, tiek problemas, su kuriomis susiduria skirtingų sričių tyrėjai ir institucijos. Ši informacija sudarys pagrindą pasiūlyti įrodymais grįstas rekomendacijas dėl ateities sprendimų, politikos krypčių ir finansavimo modelių, atitinkančių realius poreikius.

Manome, kad šis klausimynas bus naudingas ir Jūsų institucijai, nes padės įsivertinti esamus duomenų valdymo procesus, identifikuoti galimas spragas ar stiprybes ir pasiruošti tolimesniems žingsniams. Visoms institucijoms, pateikusioms atsakymus, bus pristatyti apibendrinti galimybių studijos rezultatai, kurie leis suprasti bendrą situaciją Lietuvoje ir palyginti savo patirtis su kitomis tyrimų organizacijomis.

Kilus klausimų dėl klausimyno pildymo prašome kreiptis į projekto informacijos analitikę Indrę Nazarenko, tel. +370 676 16 779, el. p. indre.nazarenko@lmt.lt.

Dėkojame už bendradarbiavimą.

Pirmininko pavaduotoja,
pavadojanti pirmininką

Vaiva Priudokienė

4 PRIEDAS

Antroji institucijų apklausa dėl mokslinių tyrimų duomenų valdymo praktikų

Lietuvos mokslo taryba vykdo projektą „Įsijungti į Europos atvirojo mokslo debesį – galimybių studijos dėl būtinų pokyčių ir investicijų krypčių parengimas“, kurio tikslas – įvertinti Lietuvos mokslo ir studijų institucijų pasirengimą integruotis į Europos atvirojo mokslo debesį (EOSC), nustatyti būtinus pokyčius bei investicines kryptis.

Šis klausimynas skirtas informacijai apie Lietuvos mokslo ir studijų institucijų taikomas mokslinių tyrimų duomenų valdymo praktikas, institucinius pajėgumus bei poreikius surinkti. Gauti rezultatai bus naudojami tiek apibendrintai, tiek analizuojant atskirų institucijų duomenis. Apklausos rezultatai bus viešai skelbiami kaip dalis mokslinių tyrimų duomenų valdymo situacijos Lietuvoje analizės.

Dalyvavimas apklausoje yra **savanoriškas**, tačiau labai svarbus siekiant parengti išsamų, faktais grįstą vertinimą ir rekomendacijas.

Šioje apklausoje renkama informacija apie institucijų praktiką, o ne apie konkrečius asmenis. Jei pateikiami kontaktiniai duomenys (pvz., atsakingo asmens vardas ar el. paštas), jie bus naudojami tik komunikacijai, susijusiai su šia galimybių studija ir tolesniu bendradarbiavimu. Visi duomenys bus tvarkomi laikantis Bendrojo duomenų apsaugos reglamento reikalavimų.

Pastaba. *Mokslinių tyrimų duomenys* šioje apklausoje apibrėžiami kaip skaitmeniniai ir (ar) fiziniai objektai, kurie yra renkami ar generuojami vykdant mokslinius tyrimus ir yra naudojami mokslinėms išvaidoms gauti ir mokslinių tyrimų išvadoms bei rezultatams patvirtinti. Į šį apibrėžimą įeina ne tik laboratorinių bandymų ar aplinkos stebėjimų duomenys, bet ir įvairių apklausų ar interviu rezultatai, tekstynai, garsynai, lentelių ar kitokia forma organizuotos duomenų bazės, vaizdo ir garso įrašai, skaitmeniniai modeliai, programinis kodas ir kt.

Mokslinių tyrimų duomenų valdymas šioje apklausoje suprantamas, kaip bet kokie veiksmai atliekami su mokslinių tyrimų duomenimis po jų sukūrimo, pvz., duomenų saugojimas, metaduomenų (kontekstinės informacijos apie duomenis) kaupimas, duomenų atvėrimas ir pan. Duomenys institucijoje gali būti valdomi tiek centralizuotai, tiek individualiai, t. y. atskirų jos narių, naudojant institucijos suteiktas arba asmenines priemones.

Duomenų rinkiniu apklausoje laikome tvarkingą skaitmeninių duomenų rinkinį, apibrėžtą pagal turinį, grupavimą, paskirtį ir duomenų tarpusavio ryšius.

Kilus klausimų prašome kreiptis el. paštu indre.nazarenko@lmt.lt arba telefonu +370 676 16 779

1. Įrašykite institucijos pavadinimą.

Mokslinių tyrimų duomenims naudojama infrastruktūra

Kaip jūsų institucija kaupia ir tvarko mokslinių tyrimų duomenis?

2. Kokią turimą aparatinę įrangą naudoja jūsų institucija mokslinių tyrimų duomenims (toliau – MTD) kaupti, tvarkyti ir atverti (serveriai, saugyklos, archyvai, talpyklos, aukšto našumo skaičiavimo įranga, didelio pralaidumo tinklai ir kt.)?

3. Kokia turima programine įranga jūsų institucija naudojasi kompiuterių tinkle duomenims tvarkyti ir mainams (Web puslapiai su paieškos galimybe, REST (angl. *Representational State Transfer*) sąsajos, duomenų įkėlimo ir parsisiuntimo žiniatinklio sąsajos ir kt.)?

(Vardinkite tik tuos išteklius, kuriuos valdo jūsų institucija.)

4. Kurie iš jūsų 2 klausime paminėtų technologinių išteklių yra nurodyti jūsų institucijos norminiuose / teisiniuose dokumentuose?

5. Ar jūsų institucijos valdomos duomenų talpyklos / saugyklos yra sertifikuotos ar planuojamos sertifikuoti kuriuo nors iš toliau nurodytų sertifikatų?

Standartai, sertifikatai	Taip	Ne, bet pradėtas sertifikavimo procesas	Ne, bet planuojama sertifikuoti	Ne, neplanuojama sertifikuoti
„CoreTrustSeal“				
„nestor Seal“				
ISO 16363				
ISO 27001				
Kita (įrašykite)				

6. Jei 5 klausime pasirinkote „Kita“ standartai, sertifikatai, įrašykite.

7. Kokias kitų Lietuvos, užsienio ir tarptautinių organizacijų valdomas talpyklas jūsų institucija, jos mokslininkai naudoja MTD saugoti („Zenodo“, PDB, NCBI, LiDA, „Clarín LT“ ir pan.)? Ar jų naudojimas koordinuojamas institucijos lygmeniu?

8. Kokiose tarptautinėse mokslinių tyrimų infrastruktūrose (toliau – MTI) jūsų institucija dalyvauja?

9. Ar jūsų institucija toje (-ose) mokslinių tyrimų infrastruktūroje (-ose) generuoja MTD?

Taip

Ne

Kita

10. Ar jūsų institucija toje (-ose) mokslinių tyrimų infrastruktūroje (-ose) teikia paslaugas?

Taip

Ne

Kita

11. Ar jūsų institucija toje (-ose) mokslinių tyrimų infrastruktūroje (-ose) naudojasi bendrais ištekliais ir paslaugomis?

Taip

Ne

Kita

12. Ar mokslinių tyrimų infrastruktūra (-os), kuriose dalyvaujate, įtrauktos į Lietuvos mokslinių tyrimų infrastruktūrų kelrodį 2024 ar į kokį kitą tarptautinį MTI sąvadą?

Lietuvos mokslinių tyrimų infrastruktūrų kelrodis 2024

<https://lmt.lrv.lt/public/canonical/1735535380/3894/LMT%20infrastrukturu%20kelrodis%202024.pdf>

Taip, yra Lietuvos MTI kelrodyje

Ne, nėra jokiam tarptautiniame sąvade

Kita

Mokslinių tyrimų duomenų politika, rengimo, valdymo, tvarkymo ir archyvavimo sąnaudos

Prašome žemiau pateiktus klausimus pakomentuoti, jei nežinoma ar nėra, taip ir nurodykite.

13. Ar Jūsų institucija turi norminius dokumentus, skirtus MTD valdymo politikai? Jei taip, pateikite nuorodą žemiau arba atsiųskite juos el. paštu indre.nazarenko@lmt.lt

14. Ar juose nustatytas reikalavimas rengti duomenų valdymo planus? (Pakomentuokite.)

15. Ar juose nustatyta archyvavimo tvarka? (Pakomentuokite.)

16. Ar yra numatyta duomenų sunaikinimo tvarka? Ar tvarkoje yra numatyta išsaugoti metaduomenis? (Pakomentuokite.)

17. Kiek laiko MTD saugomi iki sunaikinimo? (Pakomentuokite.)

18. Ar sunaikinimo procedūros dokumentuojamos (pvz., rengiami naikinimo sertifikatai (taikoma BDAR reglamentuojantiems duomenims) ir kt.)?

19. Ar Jūsų institucija turi intelektinės duomenų nuosavybės valdymą institucijoje apibrėžiantį (-ius) dokumentą (-us)? Jei taip, pateikite nuorodą žemiau arba atsiųskite juos el. paštu indre.nazarenko@lmt.lt

20. Ar Jūsų institucija vykdo stebėseną, kaip yra laikomasi MTD valdymą reglamentuojančių teisės aktų? Kokie kiti MTD valdymo kontrolės mechanizmai taikomi (MTD inventorizacija ar kt.)?

21. Ar Jūsų institucija seka / registruoja joje dirbančių tyrėjų institucijos talpykloje ir kitose talpyklose pavišintus MTD? Jeigu taip, paaiškinkite, kaip vyksta MTD pavišintų rinkinių stebėseną Jūsų institucijoje.

22. Su kokiais iššūkiais tyrėjai susiduria valdant, tvarkant ir atveriant MTD? (Pakomentuokite.)

23. Ar yra žinoma apytikslė vidutinė 2024 m. metinė kaina (EUR) už 1 TB ilgalaikio saugojimo (įskaitant atsargines kopijas)? (Jei nežinoma, taip ir nurodykite.)

24. Kaip (kokiais finansavimo šaltiniais) dengiami duomenų saugojimo kaštai? (Pakomentuokite.)

Kaupiamų ir saugojamų duomenų parametrai (tipai, apimtys, pobūdis).

Nurodykite, kokių tipų duomenis turi jūsų institucija ir apytikslę bendrą jų apimtį. Jei tikslaus skaičiaus neturite, pasirinkite artimiausią intervalą arba pažymėkite „Nežinoma“.

25. Kiek apytikriai duomenų rinkinių (angl. *datasets*) jūsų institucija yra iki šiol sukaupusi?

	iki 100 vnt.	101-1000 vnt.	1001-10 000 vnt.	>10 000 vnt.	Tokių duomenų nėra
Gamtos moksluose					
Technologijos moksluose					
Medicinos ir sveikatos moksluose					
Žemės ūkio moksluose					

	iki 100 vnt.	101–1000 vnt.	1001–10 000 vnt.	>10 000 vnt.	Tokių duomenų nėra
Socialiniuose moksluose					
Humanitariniuose moksluose					
Scenos ir ekrano menuose					
Vaizduojamuosiuose menuose					

26. Kiek iš 25 klausime pateiktų rinkinių yra tarpdisciplininiai?

27. Kiek apytikriai duomenų rinkinių jūsų institucija sukuria per metus?

	iki 100 vnt.	101–1000 vnt.	1001–10 000 vnt.	>10 000 vnt.	Tokių duomenų nėra
Gamtos moksluose					
Technologijos moksluose					
Medicinos ir sveikatos moksluose					
Žemės ūkio moksluose					
Socialiniuose moksluose					
Humanitariniuose moksluose					
Scenos ir ekrano menuose					
Vaizduojamuosiuose menuose					

28. Kiek iš 27 klausime pateiktų rinkinių yra tarpdisciplininiai?

29. Kokios bendros iki šiol sukauptų duomenų rinkinių apimtys?

	iki 100 vnt.	101–1000 vnt.	1001–10 000 vnt.	>10 000 vnt.	Tokių duomenų nėra
Gamtos moksluose					
Technologijos moksluose					
Medicinos ir sveikatos moksluose					
Žemės ūkio moksluose					
Socialiniuose moksluose					
Humanitariniuose moksluose					
Scenos ir ekrano menuose					
Vaizduojamuosiuose menuose					

30. Kokių tipų MTD generuoja Jūsų institucija (statistinius, tekstinius, vaizdo nuotraukų, geoerdvinius, garso, vaizdo, programinės įrangos, interaktyvių išteklių, 3D formato, kitokius)?

	iki 100 vnt.	101–1000 vnt.	1001–10 000 vnt.	>10 000 vnt.	Tokių duomenų nėra
Gamtos moksluose					
Technologijos moksluose					
Medicinos ir sveikatos moksluose					

	iki 100 vnt.	101–1000 vnt.	1001–10 000 vnt.	>10 000 vnt.	Tokių duomenų nėra
Žemės ūkio moksluose					
Socialiniuose moksluose					
Humanitariniuose moksluose					
Scenos ir ekrano menuose					
Vaizduojamuosiuose menuose					

31. Jei 28 klausime pažymėjote „Kita“, prašome nurodyti žemiau.

32. Kur šie duomenys saugomi?

33. Kokia apytiksliai dalis jūsų turimų duomenų rinkinių yra atvirai surandami pagal metaduomenis (aprašus)?

	iki 100 vnt.	101–1000 vnt.	1001–10 000 vnt.	>10 000 vnt.	Tokių duomenų nėra
Gamtos moksluose					
Technologijos moksluose					
Medicinos ir sveikatos moksluose					
Žemės ūkio moksluose					
Socialiniuose moksluose					
Humanitariniuose moksluose					
Scenos ir ekrano menuose					
Vaizduojamuosiuose menuose					

34. Kokia apytiksliai dalis jūsų institucijos parengtų turimų duomenų rinkinių metaduomenų (aprašų) yra atvirai prieinami?

	iki 100 vnt.	101–1000 vnt.	1001–10 000 vnt.	>10 000 vnt.	Tokių duomenų nėra
Gamtos moksluose					
Technologijos moksluose					
Medicinos ir sveikatos moksluose					
Žemės ūkio moksluose					
Socialiniuose moksluose					
Humanitariniuose moksluose					
Scenos ir ekrano menuose					
Vaizduojamuosiuose menuose					

Žmogiškieji išteklių ir kompetencijos, MTD duomenų atvėrimo skatinimas.

35. Kiek ir kokių darbuotojų (nuolatinių ar laikinų) Jūsų institucijoje atlieka funkcijas, susijusias su MTD valdymu ir tvarkymu (pvz., duomenų vadybininkai, prižiūrėtojai, inžinieriai)? Kokia apytikslė jų etato dalis skirta šioms funkcijoms?

36. Kokius su MTD valdymo ir tvarkymo tema susijusius mokymus jūsų institucija siūlo joje dirbantiems tyrėjams?

37. Ar darbo su MTD valdymu pagrindai yra įtraukti į mokymo programas? Jei taip, tai į kokio lygmens (bakaluro, magistro, daktaro) ir kokios temos?

38. Ar ir kokia individuali pagalba / konsultacijos teikiamos MTD valdymo ir tvarkymo srityje? Ar jos teikiamos nuolat ar pagal poreikį ir kas jas teikia?

39. Kokią su MTD valdymu ir tvarkymu susijusią praktinę pagalbą (pvz., duomenų saugojimas, dokumentavimas, metaduomenų parengimas ir kt.) jūsų institucija siūlo joje dirbantiems tyrėjams? Kas teikia šią pagalbą?

40. Ar MTD tvarkymas, dalijimasis ir duomenų atvėrimas yra skatinami jūsų institucijoje? Jei taip, tai kokiomis priemonėmis tai daroma? Kokios paskatos (finansinės, karjeros vertinimo, pripažinimo) taikomos?

41. Ar institucijoje organizuojami informaciniai / sklaidos renginiai apie MTD valdymą? Jei taip, ar jie vyksta reguliariai, ar galėtumėte pateikti pavyzdžių?

42. Pakomentuokite, kokie yra kiti veiksniai, darantys įtaką MTD praktikoms ir valdymui instituciniu lygmeniu (pvz., dalyvavimas „Horizonto“ programose, nacionalinių ar ES gairių laikymasis, narystė tarptautinėse infrastruktūrose, institucijų vidaus politika ar motyvavimo modeliai).

5 PRIEDAS

Lietuvos mokslo tarybos kreipimosi laiškas ir klausimynas, skirtas mokslininkams

Laba diena,

Lietuvos mokslo taryba (toliau – Taryba), įgyvendindama Lietuvos Respublikos švietimo, mokslo ir sporto ministerijos mokslo plėtros programos pažangos priemonę Nr. 12-001-01-01-01 „Gerinti mokslo ir studijų aplinką“, veiklą „Įsijungti į Europos atvirojo mokslo debesį“, projekto kodas 10-073-P-0001 (toliau – projektas), rengia galimybių studiją „Įsijungti į Europos atvirojo mokslo debesį – galimybių studijos dėl būtinų pokyčių ir investicijų kryptių parengimas“ (toliau – studija), daugiau apie projektą: <https://lmt.lrv.lt/lt/lmt-projektai/es-projektai/>.

Projektu siekiama spręsti mokslinių tyrimų duomenų infrastruktūros fragmentacijos, nevienodų meta-duomenų standartų ir duomenų valdymo kompetencijų stokos problemas. Tikimasi, kad ši galimybių studija, padės politikos formuotojams priimti tinkamus sprendimus dėl investicijų į mokslinių tyrimų duomenų ekosistemą bei galimybių jungtis į Europos atvirojo mokslo debesį (toliau – EOSC).

Studijoje ketinama išanalizuoti šalies mokslinių tyrimų duomenų infrastruktūrą, mokslo ir studijų institucijų (toliau – institucijos) kompetencijas, duomenų valdymo praktikas ir kt. Remiantis analizės rezultatais bus parengtos įrodymais grįstos išvados ir rekomendacijos dėl būtinų pokyčių ir investicijų kryptių, reikalingų Lietuvai.

Viena iš rengiamos studijos dalių yra empirinė, kuri orientuota į patirčių ir praktikų įvertinimą dviem lygmenimis: institucijų (tam skirtas atskiras klausimynas) ir individualių mokslininkų lygmeniu. Siekiame identifikuoti ir analizuoti įvairiapuses veiklos praktikas bei sukauptą patirtį. Todėl būtume dėkingi už Jūsų skirtą laiką užpildant klausimyną ir pasidalijant patirtimi. Surinkta ir apibendrinta informacija taps pagrindu rekomendacijoms dėl ateities sprendimų, politikos formavimo ir finansavimo modelių, atliepančių realius poreikius.

Kreipiamės į Jus, nes esate teikę paraišką viename iš Tarybos skelbtų mokslininkų grupių projektų konkursų. Norime patikinti, kad apklausos atsakymai yra visiškai anoniminiai – jie nebus siejami nei su konkrečia paraiška, nei su joje nurodyta duomenų valdymo praktika, nei su Jumis kaip asmeniu.

Jei esate teikę kelias paraiškas, klausimyne kviečiame pasirinkti tą variantą, kuris geriausiai atspindi Jūsų įprastą duomenų valdymo praktiką. Surinkti duomenys bus naudojami tik apibendrintai analizei, užtikrinant konfidencialumą ir laikantis BDAR bei nacionalinių teisės aktų reikalavimų.

Daugiau informacijos apie asmens duomenų apsaugą: <https://lmt.lrv.lt/lt/asmens-duomenu-apsauga>

Jūsų atsakymas mums labai svarbus, kviečiam iki 2025 m. spalio 29 d. (imtinai) užpildyti elektroninį klausimyną (nuoroda: <https://forms.office.com/e/6MkBvA0ShK>).

Iš anksto dėkojame už Jūsų bendradarbiavimą ir indėlį į šį tyrimą.

Kilus klausimų prašome kreiptis.

Klausimynas mokslininkams apie jų patirtį, dirbant su mokslinių tyrimų duomenimis

1. Koks Jūsų santykis su mokslinių tyrimų duomenimis (toliau – MTD)? (pažymėkite vieną arba kelis variantus):
 - a. Juos kuriate (generuojate, renkate, kaupiate, sujungiate kitų surinktus duomenis)
 - b. Juos tvarkote ir saugote (archyvuojate)
 - c. Juos validuojate (duomenų validavimas – tai procesas, kuriuo tikrinama, ar surinkti ar įvesti duomenys yra teisingi, išsamūs, nuoseklūs ir atitinka nustatytus kokybės kriterijus / taisykles. Jis padeda užtikrinti, kad duomenys būtų patikimi ir tinkami naudoti analizėje, sprendimų priėmimui ar DI sistemų veikimui)
 - d. Naudojatės kitų parengtais nemokamais duomenimis
 - e. Naudojatės kitų parengtais mokamais duomenimis
 - f. Netinka nė vienas variantas
 - g. Kita (įrašyti)

2. Kokią aparatinę įrangą naudojate MTD apdorojimui ir kaupimui?
 - a. Serverius
 - d. Aukšto našumo skaičiavimo įrangą
 - e. Didelio pralaidumo tinklus
 - f. Asmeninį kompiuterį
 - g. Kita

3. Kokia programine įranga naudojate duomenų mainams kompiuterių tinkle?
 - a. Žiniatinklio puslapiams su paieškos galimybe
 - b. REST (angl. *Representational State Transfer*) sąsajomis, FTP ir kt.
 - c. Failų persiuntimo paslaugomis
 - d. Debesijos tiekėjų siūlomos bendradarbiavimo paslaugomis
 - e. Archyvų ar talpyklų siūlomą įrangą
 - f. Kita
 - g. Nesinaudoju

4. Su kokio tipo MTD įprastai dirbate?
 - a. Eksperimentiniai duomenys (duomenys, gauti atliekant kontroliuojamus bandymus ar laboratorinius tyrimus)
 - b. Stebėjimų duomenys (duomenys, surinkti stebint reiškinius ar procesus be aktyvaus įsikišimo)
 - c. Apklausų duomenys (duomenys, gauti iš respondentų, tyrimų, interviu ar anketų)
 - d. Skaičiavimų / modeliavimo duomenys (apdoroti ar sujungti duomenys, gauti iš kitų šaltinių)
 - e. Tekstiniai ir dokumentiniai duomenys (tekstai, rankraščiai, transkripcijos, dokumentai, ataskaitos, archyvai)
 - f. Vaizdiniai duomenys (paveikslėliai, nuotraukos, rentgeno ar mikroskopiniai vaizdai, žemėlapiai)

- g. Erdviniai duomenys (duomenys, turintys vietos (koordinatės) ir laiko matmenį)
- h. Garso ir vaizdo duomenys (garso ar vaizdo įrašai iš tyrimų, interviu, stebėjimų)
- j. Išvestiniai (antriniai) duomenys (apdoroti ar sujungti duomenys, gauti iš kitų šaltinių)
- k. Programinės įrangos ir kodo duomenys (tyrimuose naudojamas ar sukurtas kodas, algoritmai, skriptai, modeliai)
- l. Dirbtinio intelekto modeliai – tai programiniu būdu (mokymosi procesu) gauti modeliai, sukurti remiantis duomenimis
- m. Prognozavimo modeliai (angl. *predictive model*) – mašininio mokymo būdu sukurti modeliai, skirti prognozuoti ateities reikšmes ar tendencijas pagal istorinius duomenis
- n. Interaktyvūs ir 3D išteklių (skaitmeniniai modeliai, žaidimų aplinkos, interaktyvūs žemėlapiai)
- o. Skaitmenizuoti archyviniai šaltiniai (istoriniai ar kultūros paveldo duomenys, perkelti į skaitmeninę formą)

5. Kokiu būdu dalinatės MTD?

- a. Įkeliate į atviras duomenų saugyklas (pvz., „Zenodo“, LiDA, „Clarín LT“ ir kt.)
- b. Skelbiate kaip priedą prie publikacijų / straipsnių
- c. Dalinatės tik pagal bendradarbiavimo sutartis
- d. Dalinatės tik savo institucijos viduje
- e. Dalinatės tik gavę užklausą (asmeniškai)
- f. Nesidalinu (*pereiti į 6 klausimą*)
- g. Kita (*įrašykite*)

6. Jei nesidalinate MTD, nurodykite dažniausias priežastis?

- a. Duomenų konfidencialumas
- b. Reikalavimai dėl autorių teisių
- c. Etiniai apribojimai
- d. Techninės kliūtys (pvz., saugyklų nebuvimas, duomenų kiekis per didelis) duomenys nesutvarkyti/ neparengti dalijimuisi
- e. Asmens duomenų apsaugos
- f. Trečiųjų šalių intelektinės nuosavybės
- g. Profesinių, komercinių ar valstybės ir tarnybos paslapčių
- h. Nacionalinio saugumo ir gynybos, teisėsaugos ir viešojo saugumo
- i. Paskatų ar pripažinimo stygius už duomenų atvėrimą
- j. Baimė, kad duomenimis pasinaudos kiti tyrėjai (pvz., pavogs idėjas ar publikuos rezultatus anksčiau)
- k. Kita

7. Kur, kokiose Lietuvos ar užsienio talpyklose saugote savo MTD?

- a. COD – Atvira kristalografinė duomenų bazė (*Crystallography Open Database*)
- b. „Clarín LT – Kalbos išteklių ir technologijų infrastruktūra Lietuvoje (*Common Language Resources and Technology Infrastructure*)

- c. LiDA – Lietuvos duomenų archyvas humanitariniams ir socialiniams mokslams
- d. MIDAS – Nacionalinis atvirosios prieigos mokslinių tyrimų duomenų archyvas (Lietuva)
- e. „Dryad“ – Tarptautinė atvirų mokslinių tyrimų duomenų saugykla
- f. „Figshare“ – Atvirų mokslinių tyrimų rezultatų dalinimosi platforma
- g. „Harvard Dataverse“ – Harvardo universiteto mokslinių tyrimų duomenų saugykla
- h. NCBI – JAV nacionalinis biotechnologijos informacijos centras (pvz., genetinių duomenų talpyklos)
- i. OSF – Atvirojo mokslo infrastruktūra („Open Science Framework“)
- j. PDB – Baltymų duomenų bankas („Protein Data Bank“)
- k. „Zenodo“ – Europos Sąjungos atvirų mokslinių tyrimų saugykla („Zenodo“)
- l. Kita (įrašykite)

8. Ar Jums aiškūs ir suprantami MTD FAIR reikalavimai?

- a. Taip
- b. Ne
- c. Iš dalies

9. Jei 8 klausime pasirinkote „Taip“ arba „Iš dalies“, kaip FAIR principus taikote savo MTD?

- a. Duomenys yra randami (angl. *Findable*)
- b. Duomenys yra pasiekiami (angl. *Accessible*)
- c. Duomenys yra suderinami su kitais duomenimis (angl. *Interoperable*)
- d. Duomenys yra pakartotinai panaudojami (angl. *Re-usable*)
- e. Netaikau

10. Ar Jūsų institucijoje sudarytos tinkamos MTD tvarkymui reikalingos sąlygos?

- a. Taip, sąlygos tinkamos (pereiti prie 12 klausimo)
- b. Iš dalies tinkamos
- c. Netinkamos
- d. Kita

11. Jei Jūsų manymu sąlygos MTD tvarkymui tinkamos tik iš dalies arba netinkamos, nurodykite pagrindines priežastis

- a. Trūksta žmogiškųjų išteklių (specialistų)
- b. Trūksta infrastruktūros (serverių, saugyklų, tinklų)
- c. Trūksta programinės įrangos
- d. Trūksta mokymų / kvalifikacijos
- e. Institucija neskiria pakankamo finansavimo MTD tvarkyti
- f. Institucija negauna (arba neieško) išorinio finansavimo (pvz., ES projektų ir kt.)
- g. Kita

12. Ar dalyvaujate tarptautinių mokslinių tyrimų infrastruktūrų ar užsienio mokslo institucijų veikloje, susijusioje su MTD?

- a. Taip, per bendrus projektus
- b. Taip, bet tik epizodiškai (pvz., dalyvauju renginiuose)
- c. Ne
- d. Kita (įrašykite)

13. Jei 12 klausime pasirinkote „Taip“, nurodykite kokioje veikloje dalyvaujate.

- a. Duomenų mainuose (pvz., dalijimasis duomenų rinkiniais, prisijungimas prie tarptautinių duomenų bazių)
- b. Bendros infrastruktūros kūrime (pvz., prisidėjimas prie konsorciūmų, platformų, standartų kūrimo)
- c. Duomenų analizėje (pvz., tarptautinės tyrėjų grupės, bendra analizė su užsienio partneriais)
- d. Kita (įrašykite)

14. Ar rengiate duomenų valdymo planus? (galimi keli variantai)

- a. Taip, jei reikalauja finansuojanti organizacija
- b. Taip, jei reikalauja institucija, kurioje dirbate
- c. Taip, asmenine iniciatyva
- d. Ne
- e. Kita

15. Su kokiais sunkumais dažniausiai susiduriate kaupdami, archyvuodami ir (ar) atverdami MTD?

- a. Trūksta finansavimo (pvz., lėšų infrastruktūrai, licencijoms, specialistams)
- b. Trūksta kompetencijų (pvz., žinių apie standartus, teisinius aspektus, metaduomenų rengimą)
- c. Trūksta laiko (didelis darbo krūvis, duomenų tvarkymas nėra prioritetas)
- d. Trūksta infrastruktūros (pvz., saugyklų, serverių, tinkamų sistemų)
- e. Neaiškūs reikalavimai (pvz., skirtingos finansuotojų / institucijų taisyklės, trūksta gairių)
- f. Kita

16. Kokių kompetencijų susijusių su MTD norėtumėte įgyti papildomai?

- a. Duomenų valdymas ir valdysena
- b. Duomenų kuravimas
- c. Duomenų ir informacijos saugumas
- d. Kibernetinis saugumas
- e. Etiniai klausimai
- f. Dirbtinio intelekto naudojimosi kompetencijos
- g. Duomenų valdymo ir naudojimo teisiniai aspektai
- h. Kita

17. Ar esate savo institucijos skatinami kaupti, archyvuoti ir dalintis MTD?

- a. Taip
- b. Ne
- c. Iš dalies

18. Jei taip ar iš dalies, koku būdu?

- a. Finansinės paskatos
- b. Karjeros vertinimo atžvilgiu
- c. Institucinis ar profesinis pripažinimas
- d. Kita

19. Kokie veiksniai ar motyvai Jums svarbiausi saugant, tvarkant ir atveriant duomenis? (Pažymėkite visus tinkamus variantus.)

- a. Asmeninės paskatos (vidinis pasitenkinimas, noras prisidėti prie mokslo pažangos)
- b. Finansinės paskatos (papildomas finansavimas, premijos)
- c. Karjeros vertinimas (įskaičiuojama vertinant akademinis ar profesinius pasiekimus)
- d. Institucijos pripažinimas (įvertinimas institucijos viduje)
- e. Platesnis profesinis pripažinimas (pvz., matomumas tarptautinėje bendruomenėje, cituojamumas)
- f. Bendradarbiavimo galimybės (galimybė užmegzti naujas partnerystes, dalyvauti projektuose)
- g. Dalyvavimas projektuose (pvz., galimybės prisijungti prie tarptautinių ar nacionalinių iniciatyvų, tokių kaip OPUS, EOSC ir kt.)
- h. Kita

20. Ar naudojate duomenų vadybininkų (angl. *data steward*) paslaugomis?

- a. Taip
- b. Ne
- c. Ne, bet naudočiausi atsiradus galimybei
- d. Kita (paaiškinkite)

21. Kiek apytikriai duomenų rinkinių Jūs sukuriate per metus? Duomenų rinkinys (*dataset*) paprastai reiškia savarankišką struktūrizuotą duomenų visumą, o ne pavienį įrašą. Pvz., viena apklausa = vienas rinkinys, vienas eksperimentas = vienas rinkinys, vienas lauko tyrimo duomenų failas = vienas rinkinys.

	iki 10 vnt.	11–50 vnt.	51–100 vnt.	>100 vnt.	Tokių duomenų nėra
Eksperimentiniai duomenys (duomenys, gauti atliekant kontroliuojamus bandymus ar laboratorinius tyrimus)					
Stebėjimų duomenys (duomenys, surinkti stebint reiškinius ar procesus be aktyvaus įsikišimo)					

	iki 10 vnt.	11–50 vnt.	51–100 vnt.	>100 vnt.	Tokių duomenų nėra
Apklausų duomenys (duomenys, gauti iš respondentų, tyrimų, interviu ar anketų)					
Skaičiavimų / modeliavimo duomenys (apdoroti ar sujungti duomenys, gauti iš kitų šaltinių)					
Tekstiniai ir dokumentiniai duomenys (tekstai, rankraščiai, transkripcijos, dokumentai, ataskaitos, archyvai)					
Vaizdiniai duomenys (paveikslėliai, nuotraukos, rentgeno ar mikroskopiniai vaizdai, žemėlapiai)					
Erdviniai duomenys (duomenys, turintys vietas (koordinatės) ir laiko matmenį)					
Garso ir vaizdo duomenys (garso ar vaizdo įrašai iš tyrimų, interviu, stebėjimų)					
Išvestiniai (antriniai) duomenys (apdoroti ar sujungti duomenys, gauti iš kitų šaltinių)					
Programinės įrangos ir kodo duomenys (tyrimuose naudojami ar sukurtas kodas, algoritmai, skriptai, modeliai)					
Interaktyvūs ir 3D ištekliai (skaitmeniniai modeliai, žaidimų aplinkos, interaktyvūs žemėlapiai)					
Skaitmenizuoti archyviniai šaltiniai (istoriniai ar kultūros paveldo duomenys, perkelti į skaitmeninę formą)					

22. Kokia yra vidutinė vieno Jūsų tyrimuose sukuriamo duomenų rinkinio apimtis?

MTD tipas	<50 GB	51–100 GB	101 GB – 10 TB	>10 TB	Tokių duomenų nėra
Ekspertiniai duomenys (duomenys, gauti atliekant kontroliuojamus bandymus ar laboratorinius tyrimus)					
Stebėjimų duomenys (duomenys, surinkti stebint reiškinius ar procesus be aktyvaus įsikišimo)					
Apklausų duomenys (duomenys, gauti iš respondentų, tyrimų, interviu ar anketų)					
Skaičiavimų / modeliavimo duomenys (apdoroti ar sujungti duomenys, gauti iš kitų šaltinių)					
Tekstiniai ir dokumentiniai duomenys (tekstai, rankraščiai, transkripcijos, dokumentai, ataskaitos, archyvai)					
Vaizdiniai duomenys (paveikslėliai, nuotraukos, rentgeno ar mikroskopiniai vaizdai, žemėlapiai)					

MTD tipas	<50 GB	51–100 GB	101 GB – 10 TB	>10 TB	Tokių duomenų nėra
Erdviniai duomenys (duomenys, turintys vietos (koordinatės) ir laiko matmenį)					
Garso ir vaizdo duomenys (garso ar vaizdo įrašai iš tyrimų, interviu, stebėjimų)					
Išvestiniai (antriniai) duomenys (apdoroti ar sujungti duomenys, gauti iš kitų šaltinių)					
Programinės įrangos ir kodo duomenys (tyrimuose naudojamasis ar sukurtas kodas, algoritmai, skriptai, modeliai)					
Interaktyvūs ir 3D išteklių (skaitmeniniai modeliai, žaidimų aplinkos, interaktyvūs žemėlapiai)					
Skaitmenizuoti archyviniai šaltiniai (istoriniai ar kultūros paveldo duomenys, perkelti į skaitmeninę formą)					
Kita					

23. Kas Jūsų manymu svarbiausia saugant ir valdant MTD? (*neprivalomas*)

24. Prašome nurodyti savo mokslinės veiklos patirtį

- Pradedantysis tyrėjas (asmuo turintis magistro kvalifikacinį laipsnį ar jam lygiavertę aukštojo mokslo kvalifikaciją; jis vykdo mokslinę (meno) veiklą vadovaujant pripažintam arba pirmaujančiam tyrėjui)
- Patvirtintas tyrėjas (mokslininkas (meno daktaras), kurio mokslinė (meno) veikla nėra visiškai savarankiška)
- Pripažintas tyrėjas (mokslininkas (meno daktaras), pasiekęs mokslinės (meno) veiklos savarankiškumo lygį)
- Pirmaujantysis tyrėjas (savarankiškas mokslininkas (meno daktaras), pirmaujantis savo tyrimų ar mokslo (meno) srityje)

25. Prašome nurodyti savo darbo patirtį metais

26. Nurodykite savo pagrindinę mokslo (meno) sritį:

- Gamtos mokslai
- Technologijos mokslai
- Medicinos ir sveikatos mokslai
- Žemės ūkio mokslai
- Socialiniai mokslai
- Humanitariniai mokslai
- Scenos ir ekrano menai
- Vaizduojamieji menai
- Tarpdisciplininė

27. Nurodykite savo pagrindinę mokslo kryptį(-is) .

6 PRIEDAS

Institucijų ir mokslininkų apklausų kryžminė analizė

1 lentelė. MTD reglamentavimas

Ar Jūsų institucija turi norminius dokumentus, skirtus MTD valdymo politikai	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Ne	9	40,9	22,2	40,0	62,5
Taip	7	31,8	55,6*	20,0	12,5
Iš dalies	6	27,3	22,2	40,0	25,0
Atvejų skaičius	22	100,0	9	5	8

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05, * p < 0,1.

2 lentelė. MTD archyvavimo tvarkų reglamentavimas

Ar juose nustatyta archyvavimo tvarka	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Ne	12	54,5	33,3	80,0	62,5
Taip	6	27,3	33,3	20,0	25,0
Iš dalies	4	18,2	33,3	0,0	12,5

3 lentelė. MTD sunaikinimo tvarkų reglamentavimas

Ar juose numatyta duomenų sunaikinimo tvarka	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Ne	15	68,2	55,6	80,0	75,0
Taip	2	9,1	22,2	0,0	0,0
Iš dalies	5	22,7	22,2	20,0	25,0

4 lentelė. MTD intelektinės nuosavybės reglamentavimas

Ar MSI apsibrėžusi MTD intelektinės nuosavybės valdymo tvarką	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Taip	18	81,8	100	40	87,5
Ne	1	4,5	0	20	0
Planuoja apsibrėžti	1	4,5	0	20	0
Atsakymo į šį klausimą nepateikė	2	9,1	0	20	12,5

5 lentelė. Reikalavimo rengti DVP paplitimas MTD reglamentuojančiuose dokumentuose

Ar MSI reikalauja rengti DVP	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Taip	7	31,8	44,4	20,0	25
Ne	10	45,5	22,2	60,0	62,5
Iš dalies	3	13,6	22,2	20,0	0,0
Atsakymo į šį klausimą nepateikė	2	9,1	11,1	0	12,5

6 lentelė. MTD tvarkų laikymosi stebėseną Lietuvos MSI

Ar MSI vykdo MTD tvarkų laikymosi stebėseną	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Taip	2	9,1	11,1	20,0	0,0
Ne	13	59,1	66,7	20,0	75,0
Iš dalies	5	22,7	11,1	40,0	25,0
Atsakymo į šį klausimą nepateikė	2	9,1	11,1	20,0	0,0

7 lentelė. MTD saugyklų sertifikavimo paplitimas tarp Lietuvos MSI

Ar institucijos valdomos duomenų talpyklos/saugyklos yra sertifikuotos	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Ne	18	52,9	66,7	100,0	87,5
Planuojama/pradėta	4	11,8	33,3	0,0	12,5
Atvejų skaičius	22	100,0	9	5	8

8 lentelė. MTD saugojimo vietos

MTD saugojimo vieta	Iš viso pažymėjo	Iš viso pažymėjo (%)	Universitetas (%)	Kolegija (%)	Institutas (%)
Institucinė talpykla	15	44,1	87,5	50,0	75,0
Mokslininkų kompiuteriai	10	29,4	37,5	75,0	50,0
Debesija	7	20,6	50,0	50,0	12,5
Specializuoti archyvai	5	14,7	50,0	0,0	12,5
Kita	2	5,9	0,0	0,0	25,0
Atvejų skaičius		20	8	4	8

9 lentelė. Sukauptas MTD rinkinių skaičius

Kiek apytikriai duomenų rinkinių institucija yra iki šiol sukaupusi	Kiekis	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Gamtos moksluose	Iki 100 vnt.	6	27,3	33,3	40,0	12,5
	101-1000 vnt.	2	9,1	11,1	0,0	12,5
	1001-10000 vnt.	3	13,6	22,2	0,0	12,5
	>10000 vnt.	1	4,5	0,0	0,0	12,5
	Tokių duomenų nėra	10	45,5	33,3	60,0	50,0
Technologijos moksluose	Iki 100 vnt.	6	27,3	33,3	60,0	0,0
	101-1000 vnt.	0	0,0	0,0	0,0	0,0
	1001-10000 vnt.	1	4,5	0,0	0,0	12,5
	>10000 vnt.	1	4,5	11,1	0,0	0,0
	Tokių duomenų nėra	14	63,6	55,6	40,0	87,5

Kiek apytikriai duomenų rinkinių institucija yra iki šiol sukaupusi	Kiekis	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Medicinos ir sveikatos moksluose	Iki 100 vnt.	0	0,0	0,0	0,0	0,0
	101-1000 vnt.	4	18,2	22,2	0,0	25,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	1	4,5	0,0	0,0	12,5
	Tokių duomenų nėra	17	77,3	77,8	100,0	62,5
Žemės ūkio moksluose	Iki 100 vnt.	3	13,6	11,1	40,0	0,0
	101-1000 vnt.	1	4,5	11,1	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	1	4,5	0,0	0,0	12,5
	Tokių duomenų nėra	17	77,3	77,8	60,0	87,5
Socialiniuose moksluose	Iki 100 vnt.	8	36,4	44,4	80,0	0,0
	101-1000 vnt.	2	9,1	22,2	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	12	54,5	33,3	20,0	100,0
Humanitariniuose moksluose	Iki 100 vnt.	8	36,4	44,4	60,0	12,5
	101-1000 vnt.	1	4,5	11,1	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	13	59,1	44,4	40,0	87,5
Scenos ir ekrano menuose	Iki 100 vnt.	1	4,5	0,0	20,0	0,0
	101-1000 vnt.	0	0,0	0,0	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	21	95,5	100,0	80,0	100,0
Vaizduojamuosiuose menuose	Iki 100 vnt.	2	9,1	11,1	20,0	0,0
	101-1000 vnt.	1	4,5	11,1	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	19	86,4	77,8	80,0	100,0
Atvejų skaičius		22	--	9	5	8

10 lentelė. Per metus sukaupiamas MTD rinkinių skaičius

Kiek apytikriai duomenų rinkinių Jūsų institucija sukuria per metus	Kiekis	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Gamtos moksluose	Iki 100 vnt.	8	36,4	33,3	40,0	37,5
	101-1000 vnt.	2	9,1	22,2	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	1	4,5	0,0	0,0	12,5
	Tokių duomenų nėra	11	50,0	44,4	60,0	50,0

Kiek apytikriai duomenų rinkinių Jūsų institucija sukuria per metus	Kiekis	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Technologijos moksluose	Iki 100 vnt.	6	27,3	22,2	60,0	12,5
	101-1000 vnt.	1	4,5	11,1	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	15	68,2	66,7	40,0	87,5
Medicinos ir sveikatos moksluose	Iki 100 vnt.	0	0,0	0,0	0,0	0,0
	101-1000 vnt.	0	0,0	0,0	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	1	4,5	0,0	0,0	12,5
	Tokių duomenų nėra	21	95,5	100,0	100,0	87,5
Žemės ūkio moksluose	Iki 100 vnt.	4	18,2	22,2	40,0	0,0
	101-1000 vnt.	1	4,5	0,0	0,0	12,5
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	17	77,3	77,8	60,0	87,5
Socialiniuose moksluose	Iki 100 vnt.	9	40,9	55,6	80,0	0,0
	101-1000 vnt.	0	0,0	0,0	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	13	59,1	44,4	20,0	100,0
Humanitariniuose moksluose	Iki 100 vnt.	9	40,9	55,6	60,0	12,5
	101-1000 vnt.	0	0,0	0,0	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	13	59,1	44,4	40,0	87,5
Scenos ir ekrano menuose	Iki 100 vnt.	0	0,0	0,0	0,0	0,0
	101-1000 vnt.	0	0,0	0,0	0,0	0,0
	1001-10000 vnt.	0	0,0	0,0	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	22	100,0	100,0	100,0	100,0
Vaizduojamuosiuose menuose	Iki 100 vnt.	2	9,1	11,1	20,0	0,0
	101-1000 vnt.	0	0,0	0,0	0,0	0,0
	1001-10000 vnt.	1	4,5	11,1	0,0	0,0
	>10000 vnt.	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	19	86,4	77,8	80,0	100,0
Atvejų skaičius		22	--	9	5	8

11 lentelė. Sukauptų MTD rinkinių apimtys

Kokios bendros iki šiol sukauptų duomenų rinkinių apimtys	Kiekis	N	%	Universitetas (%)	Kolegija (%)	Institutas (%)
Gamtos moksluose	Iki 100 GB	5	22,7	44,4	20,0	0,0
	Nuo 100 GB iki 10 TB	2	9,1	11,1	0,0	12,5
	> 10 TB	2	9,1	11,1	0,0	12,5
	Tokių duomenų nėra	13	59,1	33,3	80,0	75,0
Technologijos moksluose	Iki 100 GB	3	13,6	22,2	20,0	0,0
	Nuo 100 GB iki 10 TB	2	9,1	11,1	20,0	0,0
	> 10 TB	2	9,1	11,1	0,0	12,5
	Tokių duomenų nėra	15	68,2	55,6	60,0	87,5
Medicinos ir sveikatos moksluose	Iki 100 GB	7	31,8	44,4	40,0	12,5
	Nuo 100 GB iki 10 TB	2	9,1	0,0	0,0	0,0
	> 10 TB	0	0,0	0,0	20,0	12,5
	Tokių duomenų nėra	13	59,1	55,6	40,0	75,0
Žemės ūkio moksluose	Iki 100 GB	3	13,6	22,2	20,0	0,0
	Nuo 100 GB iki 10 TB	1	4,5	0,0	0,0	12,5
	> 10 TB	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	18	81,8	77,8	80,0	87,5
Socialiniuose moksluose	Iki 100 GB	7	31,8	55,6	40,0	0,0
	Nuo 100 GB iki 10 TB	2	9,1	0,0	20,0	12,5
	> 10 TB	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	13	59,1	44,4	40,0	87,5
Humanitariniuose moksluose	Iki 100 GB	5	22,7	44,4	20,0	0,0
	Nuo 100 GB iki 10 TB	1	4,5	0,0	0,0	12,5
	> 10 TB	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	16	72,7	55,6	80,0	87,5
Scenos ir ekrano menuose	Iki 100 GB	1	4,5	0,0	20,0	0,0
	Nuo 100 GB iki 10 TB	0	0,0	0,0	0,0	0,0
	> 10 TB	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	21	95,5	100,0	80,0	100,0
Vaizduojamuosiuose menuose	Iki 100 GB	2	9,1	11,1	20,0	0,0
	Nuo 100 GB iki 10 TB	1	4,5	11,1	0,0	0,0
	> 10 TB	0	0,0	0,0	0,0	0,0
	Tokių duomenų nėra	19	86,4	77,8	80,0	100,0
Atvejų skaičius		22	--	9	5	8

12 lentelė. MTD surandamumas pagal metaduomenis pagal mokslo sritis

Mokslo sritis	Didžioji dalis arba visi surandami	Apie pusę surandama	Surandama tik maža dalis arba nė vieni
Gamtos mokslai	30,8	7,7	61,5
Technologijų mokslai	27,3	0	72,7
Medicinos ir sveikatos mokslai	50	0	50
Žemės ūkio mokslai	0	0	100
Socialiniai mokslai	46,2	7,7	46,2
Humanitariniai mokslai	25	8,3	66,7
Scenos ir ekrano menai	0	0	100
Vaizduojamieji menai	0	0	100

13 lentelė. MTD metaduomenų atviras prieinamumas pagal mokslo sritis

Mokslo sritis	Didžioji dalis arba visi surandami	Apie pusę surandama	Surandama tik maža dalis arba nė vieni
Gamtos mokslai	23,1	0	76,9
Technologijų mokslai	27,3	0	72,7
Medicinos ir sveikatos mokslai	33,3	0	66,7
Žemės ūkio mokslai	0	0	100
Socialiniai mokslai	38,5	7,7	53,8
Humanitariniai mokslai	33,3	0	66,7
Scenos ir ekrano menai	0	0	100
Vaizduojamieji menai	0	0	100

7 PRIEDAS

Apdoroti mokslininkų apklausos duomenys

1 lentelė. Santykis su MTD būdingas skirtingoms mokslo sritims.

Santykis su MTD	Iš viso pažymėjo	Iš viso pažymėjo (%)	GMT (%)	HSM (%)
Kuria	213	91,8	92,7	90,2
Tvarko ir saugo	120	51,7	53,3	48,8
Validuoja	94	40,5	40,7	40,2
Naudojasi nemokamais	99	42,7	39,3	48,8
Naudojasi mokamais	23	9,9	9,3	11,0
Nei vienas netinka	5	2,2	2,7	1,2
Kita	2	0,9	0,7	1,2
Atvejų skaičius			150	82

2 lentelė. Duomenų tipai ir jų naudojimas skirtingose mokslo sričių grupėse.

Duomenų tipas	Iš viso pažymėjo	Iš viso pažymėjo (%)	GMT (%)	HSM (%)
Eksperimentiniai duomenys	142	61,2	84,0 ^{***}	19,5
Stebėjimų duomenys	69	29,7	35,3 ^{**}	19,5
Apklausų duomenys	95	40,9	20,0	79,3 ^{***}
Skaičiavimų / modeliavimo duomenys	105	45,3	54,0 ^{***}	29,3
Tekstiniai ir dokumentiniai duomenys	126	54,3	43,3	74,4 ^{***}
Vaizdiniai duomenys	102	44,0	52,7 ^{***}	28,0
Erdviniai duomenys	30	12,9	13,3	12,2
Audio ir video duomenys	40	17,2	7,3	35,4 ^{***}
Išvestiniai (antriniai) duomenys	58	25,0	22,0	30,5 [*]
Programinės įrangos ir kodo duomenys	44	19,0	28,0 ^{***}	2,4
Dirbtinio intelekto modeliai	27	11,6	15,3 ^{**}	4,9
Prognozavimo modeliai	27	11,6	14,7 [*]	6,1
Interaktyvūs ir 3D ištekčiai	7	3,0	4,7 ^{**}	0,0
Skaitmenizuoti archyviniai šaltiniai	18	7,8	3,3	15,9 ^{**}
Nedirba	1	0,4	0,7	0,0
Atvejų skaičius			150	82

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal χ^2 testą):

^{***} p < 0,001, ^{**} p < 0,05, ^{*} p < 0,1.

3 lentelė. Sukuriamų duomenų rinkinių kiekis bendrai.

Duomenų tipas	Iki 10 vnt.		11-50 vnt.		51-100 vnt.		> 100 vnt.		Tokių duomenų nėra		Iš viso	
	N	%	N	%	N	%	N	%	N	%	N	%
Eksperimentiniai duomenys	73	31,5	36	15,5	16	6,9	24	10,3	83	35,8	232	100
Stebėjimų duomenys	56	24,1	18	7,8	9	3,9	7	3	142	61,2	232	100
Skaičiavimų/modeliavimo duomenys	87	37,5	30	12,9	7	3	14	6	94	40,5	232	100
Tekstiniai ir dokumentiniai duomenys	85	36,6	49	21,1	16	6,9	16	6,9	66	28,4	232	100
Vaizdiniai duomenys	44	19	40	17,2	21	9,1	25	10,8	102	44	232	100
Erdviniai duomenys	24	10,3	5	2,2	5	2,2	9	3,9	189	81,5	232	100
Audio ir video duomenys	47	20,3	10	4,3	5	2,2	4	1,7	166	71,6	232	100
Išvestiniai (antriniai) duomenys	69	29,7	20	8,6	12	5,2	9	3,9	122	52,6	232	100
Programinės įrangos ir kodo duomenys	50	21,6	16	6,9	3	1,3	4	1,7	159	68,5	232	100
Interaktyvūs ir 3D ištekčiai	16	6,9	3	1,3	0	0	4	1,7	209	90,1	232	100
Skaitmenizuoti archyviniai šaltiniai	17	7,3	4	1,7	5	2,2	5	2,2	201	86,6	232	100

4 lentelė. Per metus sukuriamų duomenų rinkinių kiekis mokslo sričių grupėse.

Duomenų tipas	Apimtis per metus	GMT (%)	HSM (%)
Eksperimentiniai duomenys	Iki 10 vnt.	36,0**	23,2
	11-50 vnt.	22,7***	2,4
	51-100 vnt.	10,7**	0,0
	> 100 vnt.	14,7**	2,4
	Tokių duomenų nėra	16,0	72,0***
Stebėjimų duomenys	Iki 10 vnt.	23,3	25,6
	11-50 vnt.	8,7	6,1
	51-100 vnt.	6,0**	0,0
	> 100 vnt.	4,0	1,2
	Tokių duomenų nėra	58,0	67,1
Skaičiavimų / modeliavimo duomenys	Iki 10 vnt.	40,0	32,9
	11-50 vnt.	19,3***	1,2
	51-100 vnt.	4,0	1,2
	> 100 vnt.	8,7**	1,2
	Tokių duomenų nėra	28,0	63,4***
Tekstiniai ir dokumentiniai duomenys	Iki 10 vnt.	30,7	47,6**
	11-50 vnt.	24,0	15,9
	51-100 vnt.	6,7	7,3
	> 100 vnt.	6,7	7,3
	Tokių duomenų nėra	32,0	22,0
Vaizdiniai duomenys	Iki 10 vnt.	18,0	20,7
	11-50 vnt.	24,7***	3,7
	51-100 vnt.	10,7	6,1
	> 100 vnt.	14,0**	4,9
	Tokių duomenų nėra	32,7	64,6***

Duomenų tipas	Apimtis per metus	GMT (%)	HSM (%)
Erdviniai duomenys	Iki 10 vnt.	10,7	9,8
	11-50 vnt.	3,3	0,0
	51-100 vnt.	2,7	1,2
	> 100 vnt.	4,0	3,7
	Tokių duomenų nėra	79,3	85,4
Audio ir video duomenys	Iki 10 vnt.	14,7	30,5 ^{***}
	11-50 vnt.	4,7	3,7
	51-100 vnt.	0,7	4,9
	> 100 vnt.	0,7	3,7
	Tokių duomenų nėra	79,3 ^{**}	57,3
Išvestiniai (antriniai) duomenys	Iki 10 vnt.	25,3	37,8
	11-50 vnt.	9,3	7,3
	51-100 vnt.	4,0	7,3
	> 100 vnt.	5,3	1,2
	Tokių duomenų nėra	56,0	46,3
Programinės įrangos ir kodo duomenys	Iki 10 vnt.	26,0 ^{**}	13,4
	11-50 vnt.	10,0 ^{**}	1,2
	51-100 vnt.	0,7	2,4
	> 100 vnt.	2,7	0,0
	Tokių duomenų nėra	60,7	82,9 ^{***}
Interaktyvūs ir 3D ištekčiai	Iki 10 vnt.	9,3 ^{**}	2,4
	11-50 vnt.	2,0	0,0
	51-100 vnt.	0,0	0,0
	> 100 vnt.	2,0	1,2
	Tokių duomenų nėra	86,7	96,3 ^{**}
Skaitmenizuoti archyviniai šaltiniai	Iki 10 vnt.	4,7	12,2 ^{**}
	11-50 vnt.	1,3	2,4
	51-100 vnt.	1,3	3,7
	> 100 vnt.	0,0	6,1 ^{**}
	Tokių duomenų nėra	92,7 ^{***}	75,6
Atvejų skaičius		150	82

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): ^{***} p < 0,001, ^{**} p < 0,05.

5 lentelė. Sukuriamų duomenų rinkinių apimtys.

Duomenų tipas	Iki 50 GB		51 - 100 GB		101 GB - 10 TB		> 10 TB		Tokių duomenų nėra		Iš viso	
	N	%	N	%	N	%	N	%	N	%	N	%
Eksperimentiniai duomenys	111	47,8	22	9,5	13	5,6	2	0,9	84	36,2	232	100
Stebėjimų duomenys	69	29,7	14	6	4	1,7	1	0,4	144	62,1	232	100
Apklausų duomenys	80	34,5	9	3,9	2	0,9	1	0,4	140	60,3	232	100
Skaičiavimų/modeliavimo duomenys	89	38,4	28	12,1	10	4,3	1	0,4	104	44,8	232	100
Tekstiniai ir dokumentiniai duomenys	127	54,7	17	7,3	3	1,3	1	0,4	84	36,2	232	100
Vaizdiniai duomenys	76	32,8	30	12,9	16	6,9	3	1,3	107	46,1	232	100
Erdviniai duomenys	33	14,2	7	3	5	2,2	0	0	187	80,6	232	100

Duomenų tipas	Iki 50 GB		51 - 100 GB		101 GB - 10 TB		> 10 TB		Tokių duomenų nėra		Iš viso	
	N	%	N	%	N	%	N	%	N	%	N	%
Audio ir video duomenys	35	15,1	12	5,2	9	3,9	2	0,9	174	75	232	100
Išvestiniai (antriniai) duomenys	69	29,7	15	6,5	4	1,7	1	0,4	143	61,6	232	100
Programinės įrangos ir kodo duomenys	59	25,4	6	2,6	5	2,2	1	0,4	161	69,4	232	100
Interaktyvūs ir 3D išteklių	18	7,8	4	1,7	3	1,3	0	0	207	89,2	232	100
Skaitmenizuoti archyviniai šaltiniai	21	9,1	5	2,2	6	2,6	1	0,4	199	85,8	232	100
Kita	28	12,1	2	0,9	4	1,7	1	0,4	197	84,9	232	100

6 lentelė. Sukuriamų duomenų rinkinių apimtys skirtingose mokslo sričių grupėse.

Duomenų tipas	Duomenų rinkinio apimtis	GMT (%)	HSM (%)
Eksperimentiniai duomenys	Iki 50 GB	60,7 ^{***}	24,4
	51 - 100 GB	13,3 ^{**}	2,4
	101 GB - 10 TB	8,7 ^{**}	0,0
	> 10 TB	0,7	1,2
	Tokių duomenų nėra	16,7	72,0 ^{***}
Stebėjimų duomenys	Iki 50 GB	32,7	24,4
	51 - 100 GB	7,3	3,7
	101 GB - 10 TB	2,7	0,0
	> 10 TB	0,7	0,0
	Tokių duomenų nėra	56,7	72,0 ^{**}
Apklausų duomenys	Iki 50 GB	22,0	57,3 ^{**}
	51 - 100 GB	2,7	6,1
	101 GB - 10 TB	0,7	1,2
	> 10 TB	0,7	0,0
	Tokių duomenų nėra	74,0 ^{***}	35,4
Skaičiavimų/modeliavimo duomenys	Iki 50 GB	46,7 ^{**}	23,2
	51 - 100 GB	14,7	7,3
	101 GB - 10 TB	6,0	1,2
	> 10 TB	0,7	0,0
	Tokių duomenų nėra	32,0	68,3 ^{***}
Tekstiniai ir dokumentiniai duomenys	Iki 50 GB	56,0	52,4
	51 - 100 GB	4,7	12,2 ^{**}
	101 GB - 10 TB	0,0	3,7 ^{**}
	> 10 TB	0,7	0,0
	Tokių duomenų nėra	38,7	31,7
Vaizdiniai duomenys	Iki 50 GB	36,7	25,6
	51 - 100 GB	18,0 ^{**}	3,7
	101 GB - 10 TB	8,7	3,7
	> 10 TB	1,3	1,2
	Tokių duomenų nėra	35,3	65,9 ^{***}
Erdviniai duomenys	Iki 50 GB	16,7	9,8
	51 - 100 GB	3,3	2,4
	101 GB - 10 TB	2,0	2,4
	> 10 TB	0,0	0,0
	Tokių duomenų nėra	78,0	85,4

Duomenų tipas	Duomenų rinkinio apimtis	GMT (%)	HSM (%)
Audio ir video duomenys	Iki 50 GB	13,3	18,3
	51 - 100 GB	2,7	9,8**
	101 GB - 10 TB	4,0	3,7
	> 10 TB	0,0	2,4
	Tokių duomenų nėra	80,0**	65,9
Išvestiniai (antriniai) duomenys	Iki 50 GB	26,7	35,4
	51 - 100 GB	7,3	4,9
	101 GB - 10 TB	2,0	1,2
	> 10 TB	0,0	1,2
	Tokių duomenų nėra	64,0	57,3
Programinės įrangos ir kodo duomenys	Iki 50 GB	31,3**	14,6
	51 - 100 GB	4,0	0,0
	101 GB - 10 TB	3,3	0,0
	> 10 TB	0,0	1,2
	Tokių duomenų nėra	61,3	84,1***
Interaktyvūs ir 3D ištekliai	Iki 50 GB	10,7**	2,4
	51 - 100 GB	2,7	0,0
	101 GB - 10 TB	2,0	0,0
	> 10 TB	0,0	0,0
	Tokių duomenų nėra	84,7	97,6**
Skaitmenizuoti archyviniai šaltiniai	Iki 50 GB	6,0	14,6**
	51 - 100 GB	2,0	2,4
	101 GB - 10 TB	1,3	4,9
	> 10 TB	0,0	1,2
	Tokių duomenų nėra	90,7**	76,8
Kiti duomenys	Iki 50 GB	15,3**	6,1
	51 - 100 GB	1,3	0,0
	101 GB - 10 TB	2,7	0,0
	> 10 TB	0,0	1,2
	Tokių duomenų nėra	80,7	92,7**
Atvejų skaičius		150	82

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

7 lentelė. Duomenų dalijimosi praktikos.

Dalijimasis MTD	Iš viso pažymėjo	Iš viso pažymėjo (%)	GMT (%)	HSM (%)
Per atviras duomenų saugyklas	63	27,2	28,7	24,4
Kaip priedas prie publikacijų	128	55,2	62,7**	41,5
Pagal bendradarbiavimo sutartis	55	23,7	29,3**	13,4
Institucijos viduje	86	37,1	44,7**	23,2
Tik gavę (asmeninę) užklausą	119	51,3	48,7	56,1
Nesidalina	36	15,5	13,3	19,5
Kita	6	2,6	3,3	1,2
Atvejų skaičius			150	82

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal χ^2 testą): *** p < 0,001, ** p < 0,05, * p < 0,1.

8 lentelė. Dalijimosi duomenimis praktikos pagal tyrėjo karjeros pakopas.

Dalijimasis MTD	Iš viso pažymėjo	Iš viso pažymėjo (%)	Pradedantysis / patvirtintas tyrėjas (%)	Pripažintas tyrėjas (%)	Pirmaujantysis tyrėjas (%)
Per atviras duomenų saugyklas	63	27,2	19,0	24,6	33,7
Kaip priedas prie publikacijų	128	55,2	61,9	50,8	60,2
Pagal bendradarbiavimo sutartis	55	23,7	28,6	17,5	32,5 ^{***}
Institucijos viduje	86	37,1	61,9 ^{**}	34,1	36,1
Tik gavę (asmeninę) užklausą	119	51,3	42,9	54,0	49,4
Nesidalina	36	15,5	14,3	19,8 ^{**}	8,4
Kita	6	2,6	4,8	1,6	3,6
		Atvejų skaičius	21	126	83

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

9 lentelė. Duomenų valdymo planų rengimas ir dalijimosi duomenimis praktikos.

Dalijimasis MTD	Ar rengia duomenų valdymo planus							
	Jeigu reikalauja finansuojanti organizacija		Jeigu reikalauja darbo institucija		Asmeninė iniciatyva		Nerengia	
	Ne	Taip	Ne	Taip	Ne	Taip	Ne	Taip
Per atviras duomenų saugyklas	7,9	42,0	25,9	31,9	27,5	25,0	36,1	8,1
Kaip priedas prie publikacijų	52,5	57,3	55,1	55,3	53,5	65,6	57,0	51,4
Pagal bendradarbiavimo sutartis	18,8	27,5	19,5	40,4	23,0	28,1	25,3	20,3
Institucijos viduje	37,6	36,6	33,0	53,2	33,5	59,4	39,9	31,1
Tik gavę (asmeninę) užklausą	45,5	55,7	53,0	44,7	49,5	62,5	56,3	40,5
Nesidalina	20,8	11,5	17,3	8,5	16,5	9,4	11,4	24,3
Kita	3,0	2,3	2,7	2,1	3,0	0,0	1,9	4,1
Atvejų skaičius	101	131	185	47	200	32	158	74

10 lentelė. Nesidalijimo duomenimis priežastys būdingos skirtingoms mokslo sričių grupėms.

Nesidalijimo MTD priežastys	Iš viso pažymėjo	Iš viso pažymėjo (%)	GMT (%)	HSM (%)
Duomenys konfidencialūs	71	30,6	30,7	30,5
Autorinės teisės	43	18,5	18,7	18,3
Etiniai apribojimai	29	12,5	8,7	19,5 ^{***}
Techninės kliūtys	14	6,0	6,0	6,1
Duomenys nesutvarkyti	60	25,9	28,0	22,0
Asmens duomenų apsauga	36	15,5	11,3	23,2 ^{**}
Trečiųjų šalių intelektinė nuosavybė	26	11,2	9,3	14,6

Nesidalijimo MTD priežastys	Iš viso pažymėjo	Iš viso pažymėjo (%)	GMT (%)	HSM (%)
Komercinės ir pan. paslaptys	20	8,6	9,3	7,3
Nacionalinis saugumas, viešasis saugumas ir pan.	5	2,2	2,0	2,4
Nėra paskatų ar pripažinimo sistemos	44	19,0	16,7	23,2
Baimė, kad pasinaudos kiti	56	24,1	21,3	29,3
Kita	6	2,6	2,0	3,7
Atvejų skaičius			150	82

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal χ^2 testą): *** p < 0,001, ** p < 0,05, * p < 0,1.

11 lentelė. Kokie pagrindiniai iššūkiai kuruojant MTD pagal skirtingas mokslo sričių grupes.

Sunkumai kuruojant MTD	Iš viso pažymėjo	Iš viso pažymėjo (%)	GMT (%)	HSM (%)
Trūksta finansavimo	78	33,6	35,3	30,5
Trūksta kompetencijų	104	44,8	40,7	52,4*
Trūksta laiko	160	69,0	68,0	70,7
Trūksta infrastruktūros	80	34,5	40,7**	23,2
Neaiškūs reikalavimai	102	44,0	42,0	47,6
Kita	4	1,7	2,7	0,0
Jokių	7	3,0	2,7	3,7
Atvejų skaičius			150	82

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal χ^2 testą): *** p < 0,001, ** p < 0,05, * p < 0,1.

12 lentelė. Kokie pagrindiniai iššūkiai kuruojant MTD pagal tyrėjo karjeros pakopą.

Sunkumai kuruojant MTD	Iš viso pažymėjo	Iš viso pažymėjo (%)	Pradedantysis / patvirtintas tyrėjas (%)	Pripažintas tyrėjas (%)	Pirmaujantysis tyrėjas (%)
Trūksta finansavimo	78	33,6	33,3	27,0	43,4**
Trūksta kompetencijų	104	44,8	61,9	42,1	44,6
Trūksta laiko	160	69,0	66,7	63,5	78,3**
Trūksta infrastruktūros	80	34,5	33,3	35,7	33,7
Neaiškūs reikalavimai	102	44,0	52,4	39,7	47,0
Kita	4	1,7	4,8	1,6	1,2
Jokių	7	3,0	0,0	4,0	2,4
Atvejų skaičius			21	126	83

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

13 lentelė. Institucinės sąlygos ir tyrėjų patiriami sunkumai.

Sunkumai kuruojant MTD	Ar institucijoje sudarytos tinkamos MTD tvarkymui reikalingos sąlygos			
	Netinkamos (%)	Iš dalies (%)	Tinkamos (%)	Kita/nežinau (%)
Trūksta finansavimo	30,8	44,1**	28,9	16,7
Trūksta kompetencijų	57,7	54,8**	30,1	43,3
Trūksta laiko	65,4	71,0	71,1	60,0
Trūksta infrastruktūros	76,9***	48,4***	12,0	16,7
Neaiškūs reikalavimai	57,7	46,2	34,9	50,0
Kita	7,7	0,0	0,0	6,7
Jokių	3,8	1,1	2,4	10,0
Atvejų skaičius	26	93	83	30

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

14 lentelė. FAIR principų supratimo lygmuo.

Ar aiškūs ir suprantami MTD FAIR principų reikalavimai	N	%	GMT (%)	HSM (%)
Ne	79	34,1	31,3	39,0
Iš dalies	79	34,1	33,3	35,4
Taip	74	31,9	35,3	25,6
Atvejų skaičius			150	82

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

15 lentelė. FAIR principų supratimo lygmuo pagal skirtingą tyrėjų karjeros pakopą.

Ar aiškūs ir suprantami MTD FAIR principų reikalavimai	N	%	Pradedantysis / patvirtintas tyrėjas (%)	Pripažintas tyrėjas (%)	Pirmaujantysis tyrėjas (%)
Ne	79	34,1	23,8	34,9	34,9
Iš dalies	79	34,1	42,9	39,7	22,9
Taip	74	31,9	33,3	25,4	42,2
Atvejų skaičius			21	126	83

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

16 lentelė. FAIR principų taikymas mokslinėje veikloje.

Kaip FAIR principus taiko savo MTD	Ar aiškūs ir suprantami MTD FAIR principų reikalavimai	
	Iš dalies (%)	Taip (%)
Surandami	44,3	70,3***
Prieinami	38,0	74,3***
Sąveikūs	39,2	66,2***

Kaip FAIR principus taiko savo MTD	Ar aiškūs ir suprantami MTD FAIR principų reikalavimai	
	Iš dalies (%)	Taip (%)
Pakartotinai panaudojami	26,6	54,1 ^{***}
Netaiko	36,7 ^{***}	13,5
Atvejų skaičius	79	74

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): ^{***} p < 0,001, ^{**} p < 0,05.

17 lentelė. Duomenų dalijimosi praktikos ir FAIR principų aiškumo lygmuo.

Dalijimasis MTD	Ar aiškūs ir suprantami MTD FAIR principų reikalavimai		
	Ne (%)	Iš dalies (%)	Taip (%)
Per atviras duomenų saugyklas	7,6	26,6	48,6 ^{***}
Kaip priedas prie publikacijų	48,1	54,4	63,5
Pagal bendradarbiavimo sutartis	15,2	21,5	35,1 [*]
Institucijos viduje	30,4	38	43,2
Tik gavę (asmeninę) užklausą	48,1	59,5	45,9
Nesidalina	25,3 ^{**}	11,4	9,5
Kita	3,8	2,5	1,4
Atvejų skaičius	79	79	74

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): ^{***} p < 0,001, ^{**} p < 0,05.

18 lentelė. Atskirų FAIR principų įgyvendinimas.

Įgyvendintų FAIR principų skaičius	N taikančiųjų	% taikančiųjų
Visi 4 principai	38	24,8
FAIR principų netaiko	37	24,1
3 iš 4 principų	28	18,3
2 iš 4 principų	27	17,6
1 iš 4 principų	23	15,0

19 lentelė. Institucinės sąlygos ir mokslininkų dalijimosi duomenimis praktikos.

Dalijimasis MTD	Ar institucijoje sudarytos tinkamos MTD tvarkymui reikalingos sąlygos			
	Netinkamos (%)	Iš dalies (%)	Tinkamos (%)	Kita/nežinau (%)
Per atviras duomenų saugyklas	19,2	25,8	34,9	16,7
Kaip priedas prie publikacijų	53,8	61,3	51,8	46,7
Pagal bendradarbiavimo sutartis	15,4	21,5	32,5 ^{**}	13,3
Institucijos viduje	19,2	44,1	39,8	23,3
Tik gavę (asmeninę) užklausą	46,2	58,1	49,4	40,0
Nesidalina	19,2	8,6	13,3	40,0 ^{***}

Dalijimasis MTD	Ar institucijoje sudarytos tinkamos MTD tvarkymui reikalingos sąlygos			
	Netinkamos (%)	Iš dalies (%)	Tinkamos (%)	Kita/nežinau (%)
Kita	7,7	2,2	1,2	3,3
Atvejų skaičius	26	93	83	30

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

20 lentelė. Institucinės sąlygos ir mokslininkų FAIR principų naudojimas MTD tvarkymui.

Kaip FAIR principus taiko savo MTD	Ar institucijoje sudarytos tinkamos MTD tvarkymui reikalingos sąlygos			
	Netinkamos (%)	Iš dalies (%)	Tinkamos (%)	Kita/nežinau (%)
Surandami	19,2	52,7***	34,9	13,3
Prieinami	30,8	40,9	41,0	16,7
Sąveikūs	23,1	36,6	39,8	23,3
Pakartotinai panaudojami	26,9	31,2	26,5	10,0
Netaiko	19,2	12,9	20,5	16,7
Atvejų skaičius	26	93	83	30

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

21 lentelė. DVP rengimas ir institucinės paskatos.

Ar rengia duomenų valdymo planus	Ar institucija skatina kaupiti, archyvuoti ir dalintis MTD			
	Ne (%)	Iš dalies (%)	Taip (%)	Kita/nežinau (%)
Jei reikalauja finansuojanti organizacija	59,6	33,3	63,2	6,2
Jei reikalauja darbo institucija	14,7	0,0	40,4***	6,2
Asmeninė iniciatyva	11,5	0,0	21,1	12,5
Ne	30,8	66,7	19,3	81,2***
Atvejų skaičius	156	3	57	16

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): *** p < 0,001, ** p < 0,05.

22 lentelė. Institucinės paskatos ir mokslininkų dalijimosi MTD praktikos.

Dalijimasis MTD	Ar institucija skatina kaupiti, archyvuoti ir dalintis MTD			
	Ne (%)	Iš dalies (%)	Taip (%)	Kita/nežinau (%)
Per atviras duomenų saugyklas	26,3	0,0	36,8	6,2
Kaip priedas prie publikacijų	56,4	66,7	56,1	37,5
Pagal bendradarbiavimo sutartis	19,2	0,0	38,6**	18,8
Institucijos viduje	33,3	0,0	52,6***	25,0
Tik gavę (asmeninę) užklausą	57,7**	0,0	45,6	18,8
Nesidalina	14,7	66,7	7,0	43,8**

Dalijimasis MTD	Ar institucija skatina kaupiti, archyvuoti ir dalintis MTD			
	Ne (%)	Iš dalies (%)	Taip (%)	Kita/nežinau (%)
Kita	2,6	0,0	0,0	12,5
Atvejų skaičius	156	3	57	16

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): "" p < 0,001, "" p < 0,05.

23 lentelė. Institucinės paskatos.

Dalijimasis MTD	Jei institucija skatina ar iš dalies skatina, koku būdu (%)							
	Finansinės paskatos		Karjeros vertinimo atžvilgiu		Institucinis ar profesinis pripažinimas		Kita/nežino	
	Ne	Taip	Ne	Taip	Ne	Taip	Ne	Taip
Per atviras duomenų saugyklas	27,3	20,0	26,7	33,3	25,9	32,6	27,5	25,0
Kaip priedas prie publikacijų	55,5	40,0	54,4	66,7	55,0	55,8	55,9	50,0
Pagal bendradarbiavimo sutartis	23,8	20,0	23,0	33,3	19,6	41,9	24,5	17,9
Institucijos viduje	36,6	60,0	35,9	53,3	32,3	58,1	39,2	21,4
Tik gavę (asmeninę) užklausą	51,5	40,0	51,2	53,3	54,5	37,2	52,9	39,3
Nesidalina	15,9	0,0	16,1	6,7	18,5	2,3	11,8	42,9
Kita	2,6	0,0	2,8	0,0	3,2	0,0	2,0	7,1
Atvejų skaičius	227	5	217	15	189	43	204	28

24 lentelė. Dalyvavimas tarptautinėje veikloje ir dalijimosi MTD praktikos.

Dalijimasis MTD	Ar dalyvauja tarptautinių MTI ar užsienio mokslo institucijų veikloje, susijusioje su MTD		
	Taip, per bendrus projektus (%)	Taip, bet tik epizodiškai (%)	Ne (%)
Per atviras duomenų saugyklas	34,3	23,7	24,4
Kaip priedas prie publikacijų	61,4	55,3	52,0
Pagal bendradarbiavimo sutartis	34,3""	23,7	17,9
Institucijos viduje	47,1""	39,5	30,1
Tik gavę (asmeninę) užklausą	51,4	52,6	50,4
Nesidalina	10,0	10,5	20,3""
Kita	0,0	7,9	2,4
Atvejų skaičius	70	38	123

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): "" p < 0,001, "" p < 0,05.

25 lentelė. Dalyvavimas tarptautinėje veikloje ir duomenų saugojimo vietos.

Kokiose talpyklose saugo savo MTD	Ar dalyvauja tarptautinių MTI ar užsienio mokslo institucijų veikloje, susijusioje su MTD		
	Taip, per bendrus projektus (%)	Taip, bet tik epizodiškai (%)	Ne (%)
COD	5,7 ^{**}	0,0	0,8
Clarín LT	0,0	0,0	0,0
LiDA	5,7	2,6	1,6
MIDAS	11,4	5,3	13,0
Dryad	0,0	0,0	2,4
Figshare	1,4	0,0	3,3
Harvard Dataverse	2,9	0,0	0,0
NCBI	10,0	5,3	12,2
OSF	10,0 ^{**}	0,0	2,4
PDB	1,4	0,0	2,4
Zenodo	22,9	13,2	12,2
Kita	8,6	10,5	8,1
Nesaugo (nei vienoje iš nurodytų)	51,4	71,1	65,0
Atvejų skaičius	70	38	123

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): ^{***} p < 0,001, ^{**} p < 0,05.

26 lentelė. Dalyvavimas tarptautinėje veikloje ir FAIR principų taikymas.

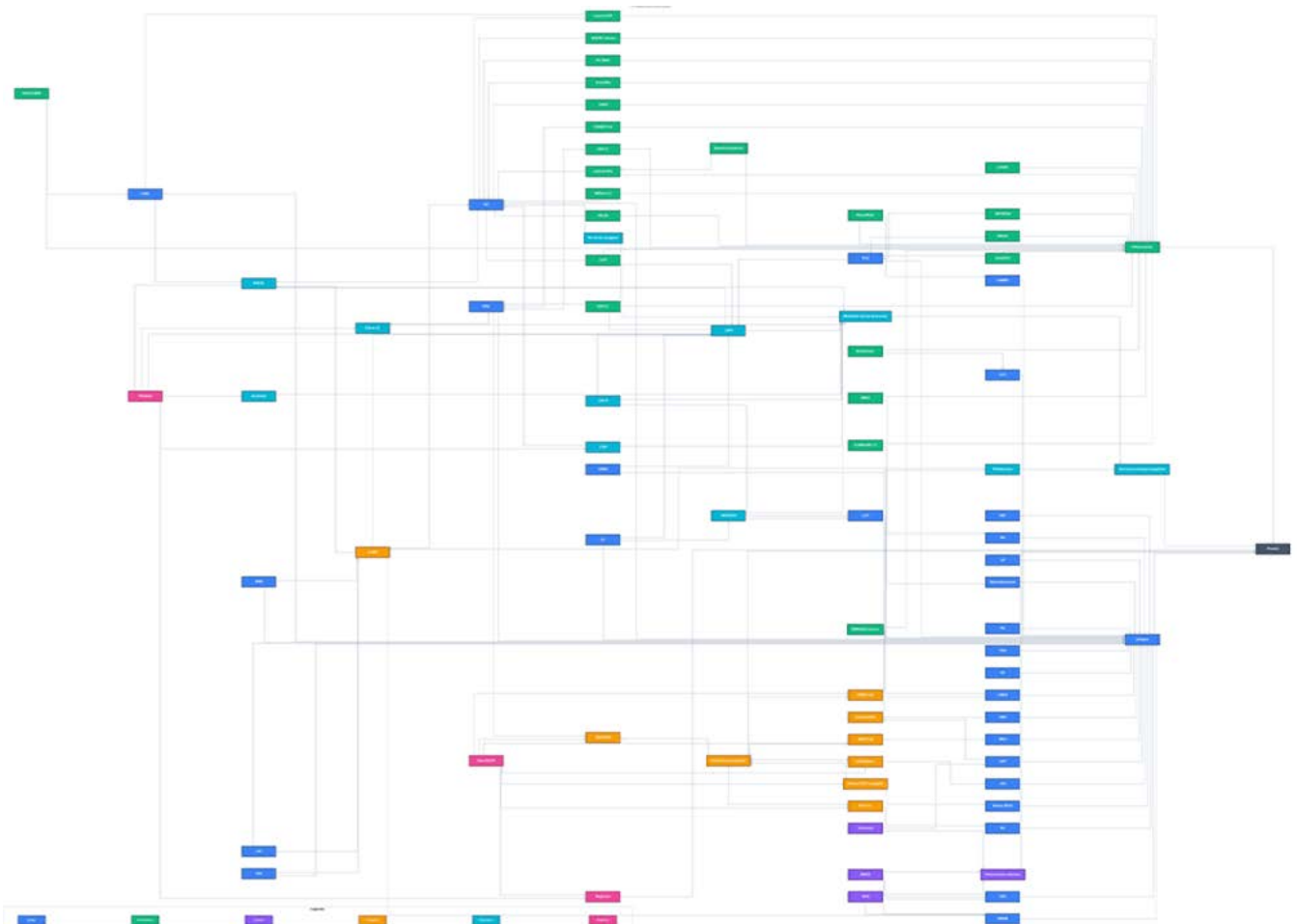
Kaip FAIR principus taiko savo MTD	Ar dalyvauja tarptautinių MTI ar užsienio mokslo institucijų veikloje, susijusioje su MTD		
	Taip, per bendrus projektus (%)	Taip, bet tik epizodiškai (%)	Ne (%)
Surandami	45,7	44,7	30,9
Prieinami	47,1 ^{**}	42,1	29,3
Sąveikūs	42,9	42,1	27,6
Pakartotinai panaudojami	34,3	28,9	21,1
Netaiko	11,4	10,5	22,0 [*]
Atvejų skaičius	70	38	123

Pastaba:

Skirtumo (t. y. kad šioje grupėje didesnė proporcija) statistinio reikšmingo lygmenys (pagal v testą): ^{***} p < 0,001, ^{**} p < 0,05.

8 PRIEDAS

Esamos mokslinių tyrimų duomenų valdymo ekosistemos Lietuvoje analitinis žemėlapis



**BŪTINI POKYČIAI IR INVESTICIJŲ KRYPTYS SIEKIANT LIETUVOS ĮSIJUNGIMO
Į EUROPOS ATVIROJO MOKSLO DEBESĮ: GALIMYBIŲ STUDIJA**

Lietuvos mokslo taryba
Gedimino pr. 3, LT 01103 Vilnius



Lietuvos mokslo taryba



Vilnius, 2026