



**GAIRĖS DĖL PILIEČIŲ RENKAMŲ IR FORMUOJAMŲ DUOMENŲ  
PANAUDOJIMO VYTAUTO DIDŽIOJO UNIVERSITETO VYKDOMUOSE  
MOKSLINIUOSE TYRIMUOSE, JŲ ATVĖRIMO, VALDYMO IR TVARKYMO**

**Rengėjai:**

dr. Saulė Milčiuvienė

dr. Darius Amilevičius

dr. Tomas Berkmanas

dr. Viktorija Kavaliauskaitė-Vilkinienė

**2024**

Įgyvendinant projektą „Piliečių įtraukimo į Vytauto Didžiojo universiteto mokslo kampanijas skatinimas, sukuriant palankią ekosistemą piliečių surinktų duomenų panaudojimui moksliniuose tyrimuose (Vis.DuomUo)“ (Nr. 10-040-T-0001)



**NAUJOS KARTOS  
LIETUVA**



**Finansuoja  
Europos Sąjunga**  
NextGenerationEU

## TURINYS

I SKYRIUS. ĮVADINĖS NUOSTATOS.....	4
II SKYRIUS. REKOMENDACIJOS DĖL ES IR NACIONALINIŲ TEISĖS AKTŲ TAIKYMO VALDANT IR TVARKANT DUOMENIS PILIEČIŲ MOKSLO PROJEKTUOSE.....	7
2.1. Mokslinių tyrimų duomenų valdymui piliečių mokslo projektuose svarbūs ES ir nacionaliniai teisės aktai .....	7
2.2. Mokslinių tyrimų duomenų valdymui piliečių mokslo projektuose svarbios sąvokos, vartojamos ES ir nacionaliniuose teisės aktuose.....	8
2.3. Pareiga atverti mokslinių tyrimų duomenis .....	10
2.4. Mokslinių tyrimų duomenų tvarkymas, atvėrus duomenis ar patalpinus juos talpykloje.....	12
2.5. Mokslinių tyrimų duomenų pakartotinio naudojimo sąlygos .....	13
III SKYRIUS. TYRIMŲ ETIKOS IR KOKYBĖS TOBULINIMO REKOMENDACIJOS ..	15
3.1. Dėl mokslinių tyrimų etikos principų įgyvendinimo (bendra perspektyva) .....	15
3.2. Dėl mokslinių tyrimų atitikties etikai ir profesionalumui vertinimo procedūrų bei atitinkamų infrastruktūrinių sprendimų .....	15
3.3. Dėl informuoto sutikimo ir informacinio lapo reikalavimų ir turinio.....	16
3.4. Dėl duomenų valdymo plano formos ir jo tinkamo užpildymo .....	17
3.5. Dėl atskirų tyrimų specifikos (tyrimų su nepilnamečiais ir pažeidžiamomis grupėmis, biomedicinos tyrimų, tyrimų su gyvūnais) .....	18
3.6. Dėl tyrimų kokybės, profesionalumo ir patikimumo užtikrinimo tyrimo metu (bendra perspektyva).....	18
3.7. Dėl manipuliavimo tyrimo metodologija ir šališkumo tyrimo metu prevencijos priemonių .....	19
3.8. Dėl duomenų fabrikavimo ir falsifikavimo tyrimo metu prevencijos priemonių ....	19
3.9. Dėl tyrėjų aktyvumo, subalansuotumo, reprezentatyvumo ir įtraukties užtikrinimo tyrimo metu.....	20
3.10. Dėl privalumų ir rizikų, susijusių su dirbtinio intelekto (DI) naudojimu tyrimuose .....	21
3.11. Dėl duomenų valdymo juos surinkus ir (ar) apdorojus (bendra perspektyva).....	21
3.12. Dėl duomenų privatumo / konfidencialumo ir saugumo užtikrinimo atlikus tyrimą .....	22
3.13. Dėl duomenų nuosavybės apsaugos ir kontrolės atlikus tyrimą .....	22
3.14. Dėl duomenų saugojimo ir sunaikinimo atlikus tyrimą.....	23

3.15.	Šaltiniai, kuriais remtasi formuluojant tyrimų etikos ir kokybės tobulinimo gaires .....	23
IV SKYRIUS. FAIR PRINCIPAI, DUOMENŲ STANDARTAI IR DUOMENŲ KOKYBĖ 26		
4.1.	FAIR principai .....	26
4.2.	FAIR: sąveikumo ir pakartotinio naudojimo įgyvendinimas.....	27
4.3.	FAIR: surandamumo įgyvendinimas – metaduomenys .....	30
4.4.	Praktinės rekomendacijos metaduomenų formavimui, laikantis FAIR principų .....	32
4.5.	Duomenų kokybė .....	32
4.6.	Praktinės rekomendacijos duomenų kokybei užtikrinti .....	34
V SKYRIUS. BAIGIAMOSIOS NUOSTATOS.....37		

# I SKYRIUS. ĮVADINĖS NUOSTATOS

- 1.1. *Gairės dėl piliečių mokslininkų renkamų ir formuojamų duomenų panaudojimo Vytauto Didžiojo universiteto vykdomuose moksliniuose tyrimuose, jų atvėrimo, valdymo ir tvarkymo* (toliau – Gairės) parengtos įgyvendinant projektą „Piliečių įtraukimo į Vytauto Didžiojo universiteto mokslo kampanijas skatinimas, sukuriant palankią ekosistemą piliečių surinktų duomenų panaudojimui moksliniuose tyrimuose (Vis.DuomUo)“ (Nr. 10-040-T-0001).
- 1.2. Gairių tikslas – pateikti rekomendacijų dėl piliečių mokslininkų renkamų ir formuojamų duomenų panaudojimo Vytauto Didžiojo universiteto vykdomuose moksliniuose tyrimuose, jų atvėrimo, valdymo ir tvarkymo piliečių mokslo projektuose, šiuos procesus vykdant laikantis aukštų etikos ir kokybės standartų.
- 1.3. Įgyvendinant piliečių mokslo projektus / iniciatyvas rekomenduojama vadovautis Europos piliečių mokslo asociacijos išskirtais, gerąją piliečių mokslo praktiką užtikrinančiais principais. Plačiau žr.: <https://www.ecsa.ngo/10-principles/>
- 1.4. Gairės sudaro penki pagrindiniai skyriai. Pirmas skyrius yra skirtas įvadinėms ir aiškinamosioms nuostatomis, kurios papildomai paaiškina kituose skyriuose pateiktą informaciją. Antras skyrius yra skirtas rekomendacijoms dėl Europos Sąjungos ir nacionaliniuose teisės aktuose įtvirtinto reguliavimo taikymo. Trečiame skyriuje aptariami tyrimų etikos ir kokybės klausimai, ketvirtame – FAIR principai, duomenų standartai ir duomenų kokybė. Baigiamosios nuostatos pateikiamos penktame skyriuje.
- 1.5. Dėl labai sparčių skaitmenizacijos procesų tiek teisinis reguliavimas, tiek etikos normos mokslinių tyrimų duomenų valdymo ir tvarkymo klausimais sparčiai kinta.
- 1.6. *Mokslo ir studijų įstatymo* 51 str. įpareigoja Lietuvos mokslo tarybą (toliau – LMT) nustatyti atvirosios prieigos prie mokslinių tyrimų ir eksperimentinės plėtos rezultatų įgyvendinimo tvarką, o *Teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatymas* pats reguliuoja mokslinių tyrimų duomenų tvarkymą juos atvėrus. Dėl to galima teigti, kad Lietuvoje egzistuoja teisės aktų kolizija, kai keli teisės aktai skirtingai reguliuoja tuos pačius santykius. Be to, Lietuvoje nėra patvirtintos nacionalinės mokslinių tyrimų duomenų atvėrimo politikos, kaip tai numato *Atvirųjų duomenų direktyvos* 10 str. Ši teisinio reguliavimo disharmonija lemia teisinio tikrumo trūkumą mokslinių tyrimų duomenų valdymo ir tvarkymo srityje. Dėl šių priežasčių Gairėse pateikiama rekomendacijų dėl teisinio reguliavimo taikymo, valdant ir tvarkant

mokslinių tyrimų duomenis piliečių mokslo projektuose. Toliau Gairėse teikiamos rekomendacijos ir LR teisinis reguliavimas aiškinamas atsižvelgiant į *ES Atvirųjų duomenų direktyvos, ES Duomenų valdymo akto ir ES Duomenų akto* nuostatas.

- 1.7. Pagal *Mokslo ir studijų įstatymo nuostatas*, visi mokslo ir studijų institucijose valstybės biudžeto lėšomis atliekamų mokslinių tyrimų ir eksperimentinės plėtros rezultatai turi būti skelbiami viešai atvirosios prieigos būdais ir priemonėmis (51 str.). *Teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatymas* taikomas mokslinių tyrimų duomenų valdymui ir tvarkymui, jeigu šie tyrimai finansuojami valstybės lėšomis, o tyrėjai, mokslinius tyrimus atliekančios organizacijos arba mokslinius tyrimus finansuojančios organizacijos jau yra juos viešai paskelbę (2 str. 2 d. 8 p.). *Atvirųjų duomenų direktyva* įpareigoja užtikrinti, kad viešosiomis lėšomis finansuojami mokslinių tyrimų duomenys būtų laisvai prieinami (atvirosios prieigos politika), vadovaujantis principu „atviras pagal standartą“ ir užtikrinant suderinamumą su FAIR principais. Visais atvejais reikalavimas atverti duomenis nėra taikomas, jei tai prieštarauja asmens duomenų, intelektinės nuosavybės, profesinių, komercinių ar valstybės ir tarnybos paslapčių apsaugą, nacionalinio saugumo ir gynybos, taip pat teisėsaugos ir viešojo saugumo veiklą reglamentuojantiems teisės aktams.
- 1.8. Antras Gairių skyrius sudarytas sistemiskai sudėliojus nacionalinių teisės aktų ir Europos Sąjungos teisės aktų nuostatas. Šiame skyriuje pateikiamos aktualios sąvokos, vartojamos nacionaliniuose ir Europos Sąjungos teisės aktuose.
- 1.9. *Teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatymas* bei LMT įsakymas *Dėl Atviros prieigos prie mokslinių tyrimų ir eksperimentinės plėtros rezultatų* pateikia skirtingas sąvokos „mokslinių tyrimų duomenys“ apibrėžtis. Rekomenduojame naudoti *Teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatyme* pateiktą „mokslinių tyrimų duomenų“ apibrėžtį, kadangi ji pažodžiui perkelta iš *Atvirųjų duomenų direktyvos*, t. y. apibrėžiama taip: „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 būtiniais mokslinių tyrimų išvadoms ir rezultatams patvirtinti“. Taip pat sąvoka papildyta šiuo *Atvirųjų duomenų direktyvoje* (Konst. d. 27) pateiktu papildomu paaiškinimu: „Mokslinių tyrimų duomenys apima statistinius duomenis, praktiškai atliekamų eksperimentų, matavimų, stebėjimų rezultatus, tyrimų išvadas, pokalbių įrašus ir vaizdo medžiagą. Jiems taip pat priklauso metaduomenys, specifikacijos ir kiti skaitmeniniai objektai. Mokslinių tyrimų

duomenys skiriasi nuo mokslinių straipsnių, kuriuose pristatomos ir komentuojamos jų autorių atliktų mokslinių tyrimų išvados.“

- 1.10. Duomenų valdymo akto II skyrius nustato viešojo sektoriaus įstaigų turimų tam tikrų kategorijų apsaugotų duomenų pakartotinį naudojimą. Tačiau šio akto 3 str. 2 d. c punktas nurodo, kad šis skyrius nėra taikomas kultūros įstaigų ir švietimo įstaigų turimiems duomenims. Iš pateiktų formuluočių nėra visiškai aišku, ar šis skyrius yra taikomas universitetams ir mokslinių tyrimų institutams. Duomenų valdymo akto įgyvendinimo rekomendaciniame dokumente teigiama, kad aktas taikomas viešojo sektoriaus institucijoms, o „viešojo sektoriaus institucija apibrėžiama kaip apimanti valdžios institucijas ir įstaigas, kurių veiklą reglamentuoja viešoji teisė. Į šią apibrėžtį įtrauktos mokslinius tyrimus atliekančios organizacijos, įsteigtos kaip viešojo sektoriaus įstaigos arba reglamentuojamos viešosios teisės.“ Tačiau aiškinant Duomenų valdymo akto II skyrių, taip pat konstatuojama: „Galiausiai šis skyrius taip pat netaikomas tyrėjų keitimuisi duomenimis nekomerciniais mokslinių tyrimų tikslais.“ Todėl nėra visiškai aišku, ar Duomenų valdymo akto II skyrius yra taikomas mokslinių tyrimų duomenims, ir, jei taikomas, tai kokia apimtimi.
- 1.11. Duomenų valdymo akto II skyrius numato pareigą institucijoms skelbti sąlygas, kuriomis duomenys gali būti pakartotinai naudojami, o prašymo dėl tokio pakartotinio naudojimo pateikimo procedūros turėtų būti atliekamos per vieną informacinį punktą.
- 1.12. Duomenų valdymo aktas yra daugiau skirtas atverti viešojo sektoriaus duomenis, reikalingus moksliniams tyrimams, o ne reguliuoti mokslinių tyrimų duomenų atvėrimą, valdymą ir tvarkymą. Duomenų rinkimui ir piliečių mokslo plėtojimui ateityje gali turėti didelės įtakos *Duomenų valdymo akte numatytas* duomenų altruizmo institutas.
- 1.13. Trečiame skyriuje teikiamos gairės tyrimų etikos ir kokybės tobulinimo klausimais remiasi šiomis bendromis prielaidomis:
  - 1.13.1. VDU ketina kurti savo piliečių mokslo strategiją;
  - 1.13.2. Gairės skiriamos institucijai ir yra adaptuotos atsižvelgiant į institucijos infrastruktūros, vidinio reguliavimo bei kitas charakteristikas;
  - 1.13.3. bendrasis mokslinių tyrimų etikos ir kokybės ekosistemos institucijoje tobulinimas *ipso facto* prisidėtų prie palankios ekosistemos panaudoti piliečių surinktus duomenis moksliniuose tyrimuose sukūrimo.

## II SKYRIUS. REKOMENDACIJOS DĖL ES IR NACIONALINIŲ TEISĖS AKTŲ TAIKymo VALDANT IR TVARKANT DUOMENIS PILIEČIŲ MOKSLO PROJEKTUOSE

### 2.1. Mokslinių tyrimų duomenų valdymui piliečių mokslo projektuose svarbūs ES ir nacionaliniai teisės aktai

Mokslinių tyrimų duomenų valdymui piliečių mokslo projektuose yra svarbūs šie ES ir nacionaliniai teisės aktai:

- 2.1.1. 2016 m. balandžio 27 d. Europos Parlamento ir Tarybos reglamentas (ES) 2016/679 dėl fizinių asmenų apsaugos tvarkant asmens duomenis ir dėl laisvo tokių duomenų judėjimo ir kuriuo panaikinama Direktyva 95/46/EB (Bendrasis duomenų apsaugos reglamentas), ELI: <http://data.europa.eu/eli/reg/2016/679/oj>
- 2.1.2. 2019 m. birželio 20 d. Europos Parlamento ir Tarybos direktyva (ES) 2019/1024 dėl atvirųjų duomenų ir viešojo sektoriaus informacijos pakartotinio naudojimo, ELI: <http://data.europa.eu/eli/dir/2019/1024/oj>
- 2.1.3. Europos Parlamento ir Tarybos reglamentas (ES) 2022/868 2022 m. gegužės 30 d. dėl Europos duomenų valdymo, kuriuo iš dalies keičiamas Reglamentas (ES) 2018/1724 (Duomenų valdymo aktas), ELI: <http://data.europa.eu/eli/reg/2022/868/oj>
- 2.1.4. 2023 m. gruodžio 13 d. Europos Parlamento ir Tarybos reglamentas (ES) 2023/2854 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), ELI: <http://data.europa.eu/eli/reg/2023/2854/oj>
- 2.1.5. Lietuvos Respublikos teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatymas (Valstybės žinios, 2000-02-02, Nr. 10-236), <https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.94745/asr>
- 2.1.6. Lietuvos Respublikos valstybės informacinių išteklių valdymo įstatymas (Valstybės žinios, 2011-12-31, Nr. 163-7739), <https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.415499>

- 2.1.7. Lietuvos Respublikos mokslo ir studijų įstatymas (Valstybės žinios, 2009-05-12, Nr. 54-2140), <https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/TAIS.343430/cRlXBRGrEn>
- 2.1.8. Lietuvos mokslo tarybos nutarimas „Dėl Atvirosios prieigos prie mokslinių publikacijų ir duomenų gairių patvirtinimo“ (TAR, 2016, Nr. 2016-04101) <https://www.infolex.lt/ta/355804>
- 2.1.9. Lietuvos mokslo tarybos pirmininko įsakymas „Dėl Atvirosios prieigos prie mokslinių tyrimų ir eksperimentinės plėtros rezultatų tvarkos aprašo patvirtinimo“ (TAR, 2024-09-02, Nr. 2024-15337), <https://www.infolex.lt/ta/971091?ref=5>
- 2.1.10. Lietuvos mokslo tarybos pirmininko įsakymas „Dėl Atvirosios prieigos prie mokslo publikacijų ir duomenų gairių pereinamojo laikotarpio nustatymo“ (TAR, 2024-08-31, Nr. 2024-15267), <https://www.infolex.lt/ta/970782>

## 2.2. Mokslinių tyrimų duomenų valdymui piliečių mokslo projektuose svarbios sąvokos, vartojamos ES ir nacionaliniuose teisės aktuose

- 2.2.1. **Atvirasis formatas** – failo formatas, kuris nepriklauso nuo programinės įrangos ir yra atpažįstamas ir prieinamas visuomenei be jokių apribojimų, trukdančių pakartotinai naudoti duomenis.
- 2.2.2. **Atviroji prieiga** – nemokama ir nevaržoma prieiga prie mokslo publikacijų, mokslinių tyrimų duomenų ir kitos publikuotos mokslinių tyrimų medžiagos, kurią galima nevaržomai skaityti, kopijuoti, pakartotinai naudoti, platinti, atlikti automatinę turinio analizę nepažeidžiant autorių teisių.
- 2.2.3. **Atvirosios prieigos talpykla** – elektroninės mokslinės informacijos infrastruktūra, kurioje saugomi ir atveriami mokslo leidiniai ir (ar) su jais susiję ar atskirai pateikti duomenų rinkiniai, kartu pateikiant metaduomenis ar kitus mokslinių tyrimų rezultatus. Talpykla gali būti tarptautinė, nacionalinė, institucinė, daugiadisciplinė, teminė ar specializuota.
- 2.2.4. **Duomenys** – užfiksuota informacija apie institucijos ar valstybės valdomo subjekto veiklą ar jos dalis, neatsižvelgiant į jos pateikimo būdą, formą (raštu arba elektronine



forma ar kaip garso, vaizdo ar garso ir vaizdo įrašas) ir laikmeną, įskaitant registro duomenis, registro informaciją, registru pateiktus duomenis ir (arba) jų kopijas, valstybės informacinės sistemos duomenis.

- 2.2.5. **Duomenų altruizmas** – savanoriškas dalijimasis duomenimis siekiant visuotinės svarbos tikslų, nesiekiant pelno.
- 2.2.6. **Duomenų rinkinys** – metodiškai susistemintų duomenų, turinčių vienodą konceptualią prasmę, visuma, įskaitant juos apibūdinančius duomenis (t. y. metaduomenis).
- 2.2.7. **Duomenų skelbimas** – duomenų tam tikra forma atskleidimas iš anksto neapibrėžtai asmenų grupei, įskaitant prieigos prie jų sukūrimą tam, kad asmenys galėtų pasiekti duomenis pasirinktoje vietoje reikiamu laiku.
- 2.2.8. **Duomenų teikimas** – duomenų atskleidimas perduodant ar kitu būdu padarant juos prieinamus konkrečiam pareiškėjui ar jo atstovui jų prašymu.
- 2.2.9. **Duomenų valdymo planas** – mokslinių tyrimų duomenų gavimo, kaupimo, saugojimo ir prieinamumo dokumentas, kuriame nurodoma, kaip duomenys bus tvarkomi mokslinio tyrimo metu ir saugomi jam pasibaigus.
- 2.2.10. **FAIR principai** (angl. *Findable, Accessible, Interoperable, Reusable*, FAIR) – nurodo savybes, kuriomis turėtų pasižymėti mokslinių tyrimų duomenys: surandami, prieinami, sąveikūs ir prieinami pakartotiniam naudojimui saugioje ir patikimoje aplinkoje.
- 2.2.11. **Elektroninis formatas** – failo formatas, kurio struktūra yra vientisa ir leidžia kompiuterinėms taikomosioms programoms lengvai nustatyti, atpažinti ir gauti konkrečius duomenis ir jų struktūrą.
- 2.2.12. **Metaduomenys** – struktūriškai apibrėžta informacija, kuri apibūdina, paaiškina dokumentą ar informacinį išteklių, nurodo jo buvimo vietą arba kitokiu būdu palengvina jo suradimą, naudojimą arba valdymą. Tai apima antraštę, autorių ir jo darbovietę, tyrimo finansuotoją, publikavimo laiką ir vietą, leidinio antraštę ir kt.
- 2.2.13. **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 būtinais mokslinių tyrimų išvadoms ir rezultatams patvirtinti.

- 2.2.14. **Nuasmėnimas** – procesas, kai fizinio asmens duomenys pakeičiami anonimais duomenimis, kurie nėra susiję su fiziniu asmeniu, kurio tapatybė yra žinoma arba gali būti nustatyta, arba procesas, kai asmens duomenų anonimiškumas užtikrinamas taip, kad duomenų subjekto tapatybė nebėgali būti nustatyta.
- 2.2.15. **Patikima talpykla** – sertifikuota tarptautiniais sertifikatais ir (ar) tarptautinės mokslo bendruomenės pripažinta daugiadalykė, teminė ar institucinė talpykla, kurios teikiamos paslaugos ir įrankiai užtikrina mokslinių tyrimų duomenų tikslumą, integralumą, autentifikavimą ir prieigą, kartu naudojami nuolatiniai identifikatoriai, mašiniu būdu nuskaitomi, standartizuoti ir išsamūs metaduomenys, įskaitant kilmę ir gavybos būdus.
- 2.2.16. **Piliečių mokslas** – visuomenės į(si)traukimas į mokslinių tyrimų procesą, kai visuomenės nariai dalyvauja įvairiose mokslinių tyrimų veiklose, ypač susijusiose su mokslinių tyrimų duomenimis.

### 2.3. Pareiga atverti mokslinių tyrimų duomenis

- 2.3.1. Rekomenduojama atverti visus mokslinių tyrimų duomenis, kurie yra finansuojami viešosiomis lėšomis, 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. Šiuo atveju mokslinių tyrimų duomenų rinkinių metaduomenys turi būti visiškai atviri, prieinami paieškai ir automatizuotai nuskaitomi, taip pat turi išlikti pasiekiami ir duomenis sunaikinus. Duomenys turi būti saugomi patikimoje duomenų talpykloje įgijus nuolatinį identifikatorių.
- 2.3.2. Duomenų bazių, kuriose yra asmens duomenų arba komerciniu požiūriu konfidencialios informacijos, atveju tai reiškia, kad jais galima dalytis kaip atviraisiais duomenimis, tik kai jie nuasmėninti arba apibendrinti taip, kad joks potencialus pakartotinis naudotojas negalėtų pakartotinai nustatyti asmenų tapatybės ar konfidencialios informacijos. Šiais duomenimis rekomenduojama dalintis laikantis *Duomenų valdymo akte* nurodytų reikalavimų.
- 2.3.3. Asmens duomenys turi būti tvarkomi griežtai laikantis *Bendrojo asmens duomenų apsaugos reglamento* (toliau – BDAR) nuostatų ir VDU patvirtintų *Asmens duomenų*

*tvarkymo taisyklių*. Kilus teisinei kolizijai, pirmenybė teikiama BDAR arba e. privatumo taisyklėmis pagal 2009/136/EB direktyvą.

- 2.3.4. Duomenų turėtojas gali nustatyti mokslinių tyrimų duomenų skirtingus atvėrimo lygius, tačiau atvėrimo lygiai pasirenkami pagal principą „kiek įmanoma atviri, kiek būtina – uždari“ (angl. *as open as possible, as closed as necessary*). Galimi šie mokslinių tyrimų duomenų atvėrimo lygiai:
  - 2.3.4.1. atviri nuo įkėlimo į talpyklą (neatlygintinai prieinami, panaudojami, gali būti modifikuojami ir toliau platinami bet kokiais tikslais pagal nurodytas atvirąsias licencijas);
  - 2.3.4.2 laikinai uždari ir atveriami per duomenų turėtojų nustatytą laikotarpį;
  - 2.3.4.3. atveriami pagal prašymą, įvertinus duomenų naudojimo tikslus.
- 2.3.5. Jei duomenų atverti negalima, turi būti svarstoma galimybė atverti duomenų rinkinių metaduomenis. Metaduomenų atvėrimas neturi prieštarauti asmens duomenų, intelektinės nuosavybės, profesinių, komercinių ar valstybės ir tarnybos paslapčių apsaugą, nacionalinio saugumo ir gynybos, taip pat teisėsaugos ir viešojo saugumo veiklą reglamentuojantiems teisės aktams.
- 2.3.6. Duomenys turi būti atvirai prieinami tuo pačiu metu kaip ir publikacijos. Šie duomenys turi būti prieinami talpyklose ar publikacijų leidėjų nurodytais kitais būdais ir susieti su atitinkamomis publikacijomis.
- 2.3.7. Ruošiantis pradėti vykdyti mokslinius tyrimus turi būti parengiamas duomenų valdymo planas. Į duomenų valdymo planus reikia įtraukti šią informaciją:
  - 2.3.7.1. duomenų aprašymą (nurodyti tipą, formatą, apimtį, duomenų gavimo šaltinį, būdą ir pakartotinio naudojimo galimybes);
  - 2.3.7.2. metaduomenis ir (ar) dokumentaciją, kuri bus pateikiama su duomenimis (nurodyti jų gavimo metodus ir struktūrą), duomenų kokybės užtikrinimo garantijas;
  - 2.3.7.3. duomenų, metaduomenų ir jų kopijų, ypač neskelbtinų duomenų atveju, saugumą užtikrinantį kaupimo būdą;
  - 2.3.7.4. duomenims taikytinus teisinius reikalavimus (intelektinės nuosavybės, asmens duomenų), jei tokių yra, ir akademinės etikos ir sąžiningumo užtikrinimo būdus;

2.3.7.5. dalinimosi duomenimis sąlygas (kada ir kaip bus dalinamasi, ar yra apribojimų, neleidžiančių dalintis, arba priešasčių embargo laikotarpiui), pakartotinio duomenų naudojimo sąlygas (galimybė parsisiųsti ir modifikuoti duomenis), taip pat ilgalaikio duomenų saugojimo sąlygas ir vietą (kuri duomenų dalis bus ilgalaikio saugojimo).

## 2.4. Mokslinių tyrimų duomenų tvarkymas, atvėrus duomenis ar patalpinus juos talpykloje

2.4.1. Paskelbus duomenis, pakartotiniam duomenų naudojimui taikomi šie reikalavimai: duomenų išsamumo, duomenų tikslumo, teisėtumo, objektyvumo, duomenų abipusio naudojimo, nediskriminavimo.

2.4.2. Draudžiama suteikti trečiajam asmeniui išimtines teises pakartotinai naudoti duomenis, išskyrus teisės aktuose nustatytas išimtis.

2.4.3. Duomenys pakartotinai naudoti pareiškėjams teikiami neatlygintinai.

2.4.4. Duomenys turi atitikti sąveikumo kriterijus. Rekomenduojama remtis EOSC plėtojama sąveikumo kriterijais.

2.4.5. Mokslinių tyrimų duomenys turi būti tvarkomi ir atveriami laikantis FAIR principų, t.y.:

2.4.5.1. Duomenų surandamumo – metaduomenys turi būti išsamūs, su nuolatiniais unikaliaisiais identifikatoriais, nuskaitomi automatinėmis ir neautomatinėmis priemonėmis.

2.4.5.2. Duomenų pasiekiamumo – turi būti naudojami nuolatiniai identifikatoriai ir standartiniai žiniatinklio protokolai, gali būti naudojamos ir kitos autentifikavimo ir autorizavimo priemonės.

2.4.5.3. Duomenų sąveikumo – turi būti naudojami tarptautinių mokslo bendruomenių nustatytos ontologijos ir terminų žodynai. Duomenys ir (ar) jų metaduomenys turi būti papildyti nuorodomis į kitus duomenis, kartu paaiškinant santykį su šaltiniais, į kuriuos pateikiamos nuorodos, ir pateikiami naudojant standartinę išteklių aprašymo struktūrą.

2.4.5.4. Pakartotinio duomenų naudojimo – turi būti naudojami išsamūs metaduomenys, standartinės atvirosios turinio ir atvirosios programinės įrangos duomenų naudojimo

licencijos, išsami informacija apie duomenų kilmę ir gavybos būdus. Duomenys ir (ar) jų metaduomenys turi būti aprašomi laikantis specifinėje mokslo srityje priimtų standartų, jei tokie yra sukurti.

2.4.6. Tyreėjams rekomenduojama rinktis talpyklas, kurios:

2.4.6.1. suteikia atvirąją prieigą prie mokslo publikacijų ir mokslinių tyrimų duomenų metaduomenų;

2.4.6.2. turi tarptautinius sertifikatus;

2.4.6.3. taiko FAIR principus;

2.4.6.4. turi aiškiai aprašytas ir viešai paskelbtas paslaugų teikimo ir išteklių saugojimo taisykles, pakartotinio naudojimo sąlygas, nurodo veiklą reglamentuojančius teisės aktus ir etines nuostatas, kuriomis vadovaujasi;

2.4.6.5. suteikia mokslo publikacijoms ir mokslinių tyrimų duomenims nuolatinius identifikatorius, atvirojo turinio licencijas arba kitu būdu nurodo jų naudojimo sąlygas, jei tai nebuvo padaryta mokslo publikacijų ir mokslinių tyrimų duomenų skelbimo metu;

2.4.6.6. naudoja standartizuotus, automatizuotai nuskaitomus metaduomenis ir vykdo metaduomenų kokybės kontrolę;

2.4.6.7. yra pasiekiamos mokslo publikacijų ir (ar) mokslinių tyrimų duomenų paieškos sistemų, turi sąsajas su mokslinių tyrimų duomenų portalais, dalyvauja elektroninių mokslinių tyrimų infrastruktūrų veikloje ir (ar) Europos atvirojo mokslo debesijoje (angl. *European Open Science Cloud, EOSC*) ar kitose atvirojo mokslo platformose.

## 2.5. Mokslinių tyrimų duomenų pakartotinio naudojimo sąlygos

2.5.1. Trečiasis asmuo be universiteto sutikimo gautus duomenis gali:

2.5.1.1. viešai skelbti bet kokių būdu, taip pat padaryti juos viešai prieinamus internete;

2.5.1.2. atgaminti bet kokia forma ar būdu;

2.5.1.3. versti į kitas kalbas;

2.5.1.4. adaptuoti, apdoroti ar kitaip perdirbti;

- 2.5.1.5. platinti ir kitaip teikti (perduoti) tretiesiems asmenims.
- 2.5.2. Gauti duomenys negali būti iškraipomi ir (arba) naudojami siekiant tiesiogiai reklamuoti produktus ar paslaugas arba siekiant sudaryti prielaidą, kad produktas ar paslauga yra universiteto remiami ar kitaip skatinami, taip pat naudojami neteisėtais, apgaulingais, nesąžiningais tikslais arba remiant tokius tikslus ir naudojami taip, kad klaidintų trečiuosius asmenis.
- 2.5.3. Pakartotinai naudodamas iš universiteto gautus duomenis, duomenų gavėjas privalo nurodyti gautų duomenų šaltinį ir datą, kada jie buvo gauti, taip pat užtikrinti, kad nebus pažeistos kitų asmenų teisės ir teisėti interesai. Už adaptuotų, apdorotų ar kitaip perdirbtų duomenų teisingumą atsako duomenų gavėjas. Atsakomybė už tokių duomenų sukeltas teises pasekmes tenka duomenų gavėjui.

## III SKYRIUS. TYRIMŲ ETIKOS IR KOKYBĖS TOBULINIMO REKOMENDACIJOS

### 3.1. Dėl mokslinių tyrimų etikos principų įgyvendinimo (bendra perspektyva)

Rekomenduojama:

- 3.1.1. ieškoti būdų, kaip informaciją apie mokslinių tyrimų etikos principus pateikti piliečiams paprasta, suprantama ir lengvai prieinama forma;
- 3.1.2. reguliariai planuoti ir vesti mokymus, į kuriuos kviesti potencialius piliečius mokslininkus ir supažindinti juos tiek su piliečių mokslo potencialu, tiek su atitinkamais mokslinių tyrimų standartais ir principais, taip iš anksto juos parengiant tyrimams;
- 3.1.3. kaip kiekvieno mokslinio tyrimo plano būtiną dedamąją numatyti privalomą suplanavimą ir paskesnę įgyvendinimą išankstinių profesionalių tyrėjų ir piliečių mokslininkų, kurie ketina dalyvauti tyrime, susitikimų, kurių pagrindinis tikslas būtų aptarti mokslinių tyrimų standartus ir principus;
- 3.1.4. atitinkamus vaidmenis numatyti mokslinių tyrimų etikos komitetams (pvz., minėtus išankstinius susitikimus gali praveisti šie komitetai, dalyvaujant vien tik komitetų nariams ir piliečiams mokslininkams, arba į tai įtraukiant ir profesionalius tyrėjus);
- 3.1.5. pasirengti institucinę pavyzdinę mokslinio tyrimo plano piliečių mokslo projektams formą.

### 3.2. Dėl mokslinių tyrimų atitikties etikai ir profesionalumui vertinimo procedūrų bei atitinkamų infrastruktūrinių sprendimų

Rekomenduojama:

- 3.2.1. ieškoti būdų, kaip tyrimų etikos vertinimo procedūras adaptuoti piliečių mokslo specifikai, tuo pačiu siekiant išlaikyti profesionalumo ir kokybės standartus;

- 3.2.2. kadangi etinės ribos piliečių mokslo srityje dažnai yra dviprasmybių valdymo klausimas, nenustatantis aiškių ribų tarp neetiškumo ir etiškumo, siekti tokias neaiškias situacijas kiek įmanoma neutralizuoti (ypač kai jos gali vesti prie labiau neigiamų padarinių), taip pat atskirais atvejais jas išnaudoti kūrybiškai kuriant kuo atviresnę ir lankstesnę etikos ir kokybės sąveiką piliečių moksle;
- 3.2.3. įsteigti atskirą piliečių mokslui ir jo vertinimui skirtą konsultacinį / informacinį centrą / pareigūną, kurie galėtų padėti padaliniuose veikiančioms mokslinių tyrimų atitikties etikai ir profesionalumui vertinimo komitetams vertinti piliečių mokslo projektus;
- 3.2.4. tyrimų etikos vertinimo procedūrą piliečių mokslo projektams skirstyti į du etapus: 1) kai pirmame etape atliekamas išankstinis tyrimo vertinimas mokslinių tyrimų atitikties etikai ir profesionalumui vertinimo komitete, šiam komitetui labiau bendradarbiaujant su mokslininkais (profesionaliais ir piliečiais mokslininkais), piliečius labiau įtraukiant į bendrakūros procesus (angl. *co-creation, co-design*) (pvz., tyrimo idėjos / hipotezės, metodologijos, mokslinio tyrimo plano (angl. *study protocol*) bendrakūra); šiam etapui taip pat gali būti parengiami ir naudojami atskiri, tam adaptuoti, informuotas sutikimas ir informacijos apie tyrimą lapas; 2) kai antrame etape atliekamas tradiciniu tapęs tyrimo vertinimas.

### 3.3. Dėl informuoto sutikimo ir informacinio lapo reikalavimų ir turinio

Rekomenduojama:

- 3.3.1. dėl skirtingų dalyvaujančių tyrėjų profilių piliečių mokslo projektuose kartu su informuotu sutikimu pasirašyti bendradarbiavimo / bendros mokslinės veiklos susitarimą (arba informuotą sutikimą apskritai pakeisti tokiu susitarimu, kurio integrali dalis būtų tiek toks sutikimas, tiek informacinis lapas), kuris *inter alia* numatytų visų pasirašančiųjų įsipareigojimus, siejamus su skirtingais vaidmenimis, etikos standartų pripažinimu / supratimu, galimų interesų konfliktų deklaravimu, autorystės politika, tyrime renkamais duomenimis ir t. t.;
- 3.3.2. informacijos apie tyrimą lapą piliečių mokslo projektuose papildyti informacija apie piliečių mokslininkų, kaip labiau pasyvių tyrimo objektų / subjektų (duomenų teikėjų),



ir piliečių mokslininkų, kaip aktyvių tyrimo dalyvių (įskaitant duomenų rinkėjų), dalyvavimo specifikas;

- 3.3.3. informuotame sutikime ir informaciniame lape, kurie siejami su piliečių mokslo projektais, išskirti svarbiausią informaciją, aktualiausiose vietose reikalaujant dėti atskirus parašus, kartu su informuotu sutikimu reikalauti pasirašyti ir informacijos apie tyrimą lapą.

### 3.4. Dėl duomenų valdymo plano formos ir jo tinkamo užpildymo

Rekomenduojama:

- 3.4.1. sukurti ir patvirtinti individualizuotą (ypač lyginant su LR Akademinės etikos ir procedūrų kontrolieriaus patvirtinta pavyzdine forma) institucinę pavyzdinę duomenų valdymo plano formą (individualizuota forma gali būti papildoma tokiais klausimais: „kokie resursai (pavyzdžiui, finansiniai ir laiko) bus skiriami duomenų tvarkymui ir užtikrinimui, kad duomenys atitiktų FAIR reikalavimus?“, „kaip bus užtikrinamas unikalus ir pastovus identifikatoriaus (kaip DOI) suteikimas ir palaikymas kiekvienam duomenų masyvui?“, „kas institucijoje bus atsakingas už duomenų tvarkymą (pvz., duomenų tvarkytojai (angl. *data stewards*))?“ ir kt.;
- 3.4.2. suderinti patvirtintą duomenų valdymo plano formą su esama duomenų tvarkymo ir saugojimo infrastruktūra bei atitinkamos paramos / pagalbos teikimo tyrėjams galimybėmis institucijoje;
- 3.4.3. duomenų valdymo plane svarstyti atskirti pirminius (angl. *raw*) ir apdorotus (angl. *aggregated*) duomenis, kiek tai susiję su jų saugojimu ir FAIR principų užtikrinimu;
- 3.4.4. kiek įmanoma plėsti institucijos paramą / pagalbą tyrėjams pildant duomenų valdymo planus, rengti duomenų tvarkytojus (angl. *data stewards*), kurie galėtų teikti tokią paramą / pagalbą, integruoti kompetencijų, siejamų su tinkamu duomenų plano užpildymu, teikimą į institucijoje realizuojamas šviečiamojo pobūdžio priemones.

### 3.5. Dėl atskirų tyrimų specifikos (tyrimų su nepilnamečiais ir pažeidžiamomis grupėmis, biomedicinos tyrimų, tyrimų su gyvūnais)

Rekomenduojama:

- 3.5.1. jeigu į planuojamą mokslinį tyrimą, kuriame numatytas nepilnamečių ar asmenų, priklausančių pažeidžiamoms grupėms, dalyvavimas, ketinama *aktyviau* įtraukti piliečius mokslininkus, numatyti (pvz., tyrimo plane), kaip tokie piliečiai mokslininkai bus supažindinami su tokių tyrimų specifika ir jiems (įskaitant duomenų rinkimą) keliamais reikalavimais;
- 3.5.2. numatyti (pvz., mokslinio tyrimo plane), kaip nepilnamečių ar asmenų, priklausančių pažeidžiamoms grupėms, informuoto sutikimo gavime ir supažindinime su informaciniame lape esančia informacija dalyvaujantys asmenys (pvz., tėvai, globėjai, tretieji asmenys) bus supažindinami su numatomu tyrimu (jiems žinotiniais jo elementais, įskaitant duomenų rinkimą) ir užtikrinamas atitinkamų veiksmų realumas;
- 3.5.3. numatyti (pvz., mokslinio tyrimo plane), kaip į biomedicinos tyrimus ir tyrimus su gyvūnais įtraukiami piliečiai mokslininkai bus supažindinami su tokiems tyrimams (įskaitant juose vykdomą duomenų rinkimą) keliamais reikalavimais;
- 3.5.4. siekti, kad kiekviena konkreti piliečių mokslo bendruomenė ir susiję partneriai nustatytų geriausias specifikacijas, pritaikytas kiekvieno atskiro projekto unikalioms aplinkybėms, taip siekdami, kad piliečių mokslo tyrimai galėtų išnaudoti visą savo potencialą.

### 3.6. Dėl tyrimų kokybės, profesionalumo ir patikimumo užtikrinimo tyrimo metu (bendra perspektyva)

Rekomenduojama:

- 3.6.1. prižiūrėti (sukurti to mechanizmą), kaip į tyrimus įtraukiami piliečiai mokslininkai (priklausomai nuo jų įsitraukimo lygmens) laikosi pareigos užtikrinti tyrimų kokybę,

profesionalumą ir patikimumą, ypač turint omenyje duomenų rinkimą / teikimą, apdorojimą, panaudojimą, vertinimą ir t. t.;

- 3.6.2. siekti, kad visi piliečių mokslo projektuose dalyvaujantys subjektai sklandžiai bendradarbiautų, ypač atsižvelgdami į informuoto asmens sutikimą, taip pat projektų ypatybes / specifikaciją bei duomenų rinkimo ir tvarkymo praktikos skaidrumą (t. y., siekdami, kad duomenys būtų kokybiški ir tinkami naudoti pagal paskirtį).

### 3.7. Dėl manipuliavimo tyrimo metodologija ir šališkumo tyrimo metu prevencijos priemonių

Rekomenduojama:

- 3.7.1. imtis šviečiamojo pobūdžio priemonių supažindinti institucijos mokslininkus su piliečių mokslo atveriamomis galimybėmis / priemonėmis mažinti manipuliavimo tyrimo metodologija ir šališkumo (įskaitant nesąmoningą) rizikas reguliariuose profesionalių mokslininkų atliekamuose tyrimuose (priemonių pavyzdžiai: visiškai atviro tyrimų recenzavimo leidimas (angl. *peer review*), bendras platesnis tyrėjų įtraukimas, validavimas antriniais duomenų šaltiniais (angl. *validation with a secondary data source*), taip išplečiant duomenų arealą, kt.);
- 3.7.2. institucijos piliečių mokslo strategijoje numatyti priemones, kaip institucija ir jos mokslininkai prisidės neutralizuojant galimas manipuliacijas ir šališkumus pačiame piliečių moksle, ypač kaip sumažins ar neutralizuos su piliečių mokslu siejamą grupės mąstymo (angl. *groupthink*) riziką (priemonių pavyzdžiai: kolegų įsitraukimas ir mokymas (angl. *peer learning*), atitinkamos komunikacijos tarp tyrėjų (profesionalų ir piliečių mokslininkų) priemonės, kt.).

### 3.8. Dėl duomenų fabrikavimo ir falsifikavimo tyrimo metu prevencijos priemonių

Rekomenduojama:

- 3.8.1. imtis šviečiamojo pobūdžio priemonių supažindinti institucijos mokslininkus su priemonėmis, padedančiomis išvengti duomenų fabrikavimo ar falsifikavimo tyrimų metu ar sumažinti tokių reiškinių atsiradimo galimybes (priemonių pavyzdžiai: kolegų ir ekspertų verifikacija (angl. *peer and expert verification*), automatinė duomenų kokybės patikra (angl. *automatic quality assessment*), modelinė duomenų kokybės patikra (angl. *model-based quality assessment*), pokalbių robotai ir automatizuotos bendravimo platformos (angl. *chatbots and conversational agents*), kt.);
- 3.8.2. pagal galimybes, institucijai dalyvauti kuriant ir tyrėjų bendruomenei pasiūlant tam tikras duomenų kokybės užtikrinimo priemones / sprendimus (ar duomenų validavimo / patikros praktikas), kuriais siekiama išvengti duomenų fabrikavimo ir falsifikavimo (priemonių pavyzdžiai: automatinės patikros sistemos, pokalbių robotai (angl. *chatbots*), kt.).

### 3.9. Dėl tyrėjų aktyvumo, subalansuotumo, reprezentatyvumo ir įtraukties užtikrinimo tyrimo metu

Rekomenduojama:

- 3.9.1. imtis šviečiamojo pobūdžio priemonių supažindinti institucijos mokslininkus su priemonėmis, kaip piliečius mokslininkus įtraukti ir angažuoti juos pasilikti tam tikrame piliečių mokslo projekte / tyrime, taip pat atskirais atvejais sudaryti sąlygas tokias priemones realizuoti (priemonių pavyzdžiai: išankstiniai interviu ar apklausos, siekiant išsiaiškinti piliečių angažuotumą projektui / veiklai ir specialius poreikius, naudotojui draugiškų ir įtraukiančių platformų sukūrimas, tam tikrų apdovanojimų ir pripažinimo piliečiams mokslininkams siūlymas, grįžtamojo ryšio teikimas, didesnė laisvė piliečiams mokslininkams pasirinkti įsitraukimo į tyrimą būdus, tyrimo sutapatinimas su žaidimu (angl. *gamification*), kt.);
- 3.9.2. imtis šviečiamojo pobūdžio priemonių supažindinti institucijos mokslininkus su priemonėmis, kaip suvaldyti / subalansuoti atskirų piliečių mokslininkų įsitraukimo lygmenis į piliečių mokslo projektus / tyrimus, ypač kai susidaro disbalansas tarp tyrėjų, kurie įsitraukia į tyrimą menkai / per mažai, ir vadinamų „superdalyvių“, kurie įsitraukia itin aktyviai ir gali lemti tyrimo rezultatų / duomenų šališkumą (priemonių

pavyzdžiai: virtualūs botai (angl. *virtual bots*) arba virtualūs recenzentai (angl. *virtual peers*));

- 3.9.3. institucijos piliečių mokslo strategijoje numatyti reikalavimus ir priemones, kaip užtikrinti tyrėjų reprezentatyvumą ir įtrauktį piliečių mokslo projektuose / tyrimuose, ypač reaguojant į tai, kad juose vis dar esama lyčių disbalanso, didesnio jaunesnių, labiau išsimokslinusių ar geriau ekonomiškai aprūpintų žmonių įsitraukimo, kas siejama su vadinamąją skaitmenine atskirtimi (angl. *digital divide*), žmonių su negalia per menko įtraukimo ir t. t. Tai turėtų būti realizuojama laikantis reguliavimo ekosistemos prieigos, t. y. nuostatas strategijoje susiejant su kitais vidiniais sureguliuojamais (pvz., lyčių lygybės planu, negalios politika ir pan.).

### 3.10. Dėl privalumų ir rizikų, susijusių su dirbtinio intelekto (DI) naudojimu tyrimuose

Rekomenduojama:

- 3.10.1. institucijos bendroje mokslo strategijoje numatyti siekį balanso tarp mokslo efektyvumo, naudojant DI, ir poreikio labiau atverti mokslą, į jį įtraukiant vis daugiau žmonių;
- 3.10.2. institucijos piliečių mokslo strategijoje numatyti DI naudojimo piliečių moksle saugiklius siekiant mokslo / tyrimų kokybės (pvz., sukonkretinant, koku mastu piliečiai mokslininkai būtų prileidžiami prie DI kūrimo, vystymo ar naudojimo moksliniuose tyrimuose, kaip būtų užtikrinama, kad DI duos moksliskai validžius / tikslius rezultatus ir kt.).

### 3.11. Dėl duomenų valdymo juos surinkus ir (ar) apdorojus (bendra perspektyva)

Rekomenduojama:

- 3.11.1. institucijos piliečių mokslo strategijoje numatyti FAIR principų laikymosi prievolę, duomenų, surinktų ir (ar) apdorotų piliečių mokslo projektuose / veiklose, atžvilgiu

(plačiau žr. Gairių ketvirtą skyrių); remiantis Europos piliečių mokslo asociacijos nustatytu vienu iš piliečių mokslo principu, užtikrinti, kad piliečių mokslo projektų duomenys ir metaduomenys būtų pateikiami visuomenei prieinama forma, o rezultatai, kai tai yra įmanoma, publikuojami atviros prieigos formatu;

3.11.2. institucijos piliečių mokslo strategijoje numatyti, kaip institucija prisidės prie geresnio duomenų sąveikumo ir daugkartinio naudojimo principų realizavimo per įvairių piliečių mokslo iniciatyvų ir resursų integraciją (atitinkamai, įgalinant informacijos judėjimą tarp daugybės duomenų bazių), ypač globaliame kontekste (tai siejama taip pat su atitinkamais infrastruktūriniais, ypač duomenų saugojimo, sprendimais).

### 3.12. Dėl duomenų privatumo / konfidencialumo ir saugumo užtikrinimo atlikus tyrimą

Rekomenduojama:

3.12.1. institucijos piliečių mokslo strategijoje akcentuoti esamas priemones duomenų privatumui / konfidencialumui užtikrinti (ypač identifikatorių mechanizmą), taip pat numatyti atitinkamų papildomų technologinių priemonių naudojimo ir plėtojimo (įskaitant institucijos įsitraukimo) galimybes;

3.12.2. realizuoti papildomas priemones duomenų privatumui / konfidencialumui ir saugumui užtikrinti situacijose, kuriose tam kyla didesnės rizikos (pvz., naudojant technologijas, kurios leidžia nustatyti tyrimo dalyvių buvimo vietą, įvertinti jų politiškai jautrias nuostatas, priėti prie neuroduomenų (angl. *brain data*) ir pan.).

### 3.13. Dėl duomenų nuosavybės apsaugos ir kontrolės atlikus tyrimą

Rekomenduojama:

3.13.1. įvertinti poreikį institucijai įsteigti nepelno organizaciją (-as), siekiant dalyvauti sprendimų priėmimo dėl potencialiai komercializuotinų piliečių mokslo kontekste

vykdomų tyrimų rezultatų; tuo pačiu siekti, kad komercinė nauda netaptų pagrindiniu piliečių mokslo rėmuose vykdomų tyrimų tikslu;

- 3.13.2. įvertinti poreikį institucijai dalyvauti su didžiaisiais duomenimis susijusiuose procesuose (t. y. jų rinkime, saugojime, sisteminime), pvz., duomenų kooperatyvų ar duomenų *trastų* (angl. *data trust*) veikloje, pradedant nuo galimo jų steigimo. Tokiu būdu sudaryti palankesnes sąlygas piliečių mokslo plėtrai institucijoje, kartu neapsiribojant oportunistiniu duomenų rinkimu, o skatinant surinktų duomenų sistemine analizę / apdorojimą ir jų mokslinį-praktinį pritaikomumą.

### 3.14. Dėl duomenų saugojimo ir sunaikinimo atlikus tyrimą

Rekomenduojama:

- 3.14.1. suvienodinti duomenų saugojimo praktikas institucijoje, *inter alia*, piliečių mokslo projektuose, kuriuose dalyvauja institucijos tyrėjai, užtikrinti duomenų saugojimo kokybę, tyrėjus įpareigoti naudotis institucijoje esamais duomenų saugojimo sprendimais / infrastruktūra;
- 3.14.2. įvertinus institucijos infrastruktūros galimybes, svarstyti galimybes duomenis saugoti ilgiau nei pagal dabartinį reguliavimą nustatyta minimali riba (t. y. dešimt metų).

### 3.15. Šaltiniai, kuriais remtasi formuluojant tyrimų etikos ir kokybės tobulinimo gaires

- 3.15.1. Citizen science and citizen engagement, Achievements in Horizon 2020 and recommendations on the way forward [2020] // <https://op.europa.eu/en/publication-detail/-/publication/c30ddc24-cbc6-11ea-adf7-01aa75ed71a1/language-en>.
- 3.15.2. CORDIS results pack on ethics and integrity in research [2024] // <https://op.europa.eu/en/publication-detail/-/publication/e3b4050a-1bde-11ef-a251-01aa75ed71a1/language-en>.

- 3.15.3. Data-validation solutions for citizen science data on invasive alien species [2021] // <https://op.europa.eu/en/publication-detail/-/publication/7c793276-2fbb-11ec-bd8e-01aa75ed71a1/language-en>.
- 3.15.4. ECSA Characteristics of Citizen Science: Explanation Notes [2020] // [https://zenodo.org/records/3758555/files/ecsa\\_characteristics\\_of\\_citizen\\_science\\_explanation\\_notes\\_-\\_v1\\_final.pdf?download=1](https://zenodo.org/records/3758555/files/ecsa_characteristics_of_citizen_science_explanation_notes_-_v1_final.pdf?download=1).
- 3.15.5. ESCA 10 piliečių mokslo principų [2015] // [https://zenodo.org/record/5127534/files/ECSA\\_Ten\\_principles\\_of\\_CS\\_Lithuanian.pdf?download=1](https://zenodo.org/record/5127534/files/ECSA_Ten_principles_of_CS_Lithuanian.pdf?download=1).
- 3.15.6. Ethics and integrity in research, Building a culture of trust and excellence [2024] // <https://op.europa.eu/en/publication-detail/-/publication/8a4e03c7-279a-11ef-a195-01aa75ed71a1/language-en>.
- 3.15.7. Europos Komisijos rekomendacija dėl Europos mokslininkų chartijos ir dėl Mokslininkų priėmimo į darbą elgesio kodekso [2005] // <https://eur-lex.europa.eu/legal-content/LT/TXT/PDF/?uri=CELEX:32005H0251&qid=1399637238252&from=LT>.
- 3.15.8. Exploration of citizen engagement methodologies in European R&I projects 3.0 [2023] // <https://op.europa.eu/en/publication-detail/-/publication/2222ec21-6ef9-11ee-9220-01aa75ed71a1/language-en>.
- 3.15.9. Ficorilli, Antonella, *et al.* Investigating the process of ethical approval in citizen science research: the case of Public Health, *JCOM Journal of Science Communication* [2021] // [https://jcom.sissa.it/article/pubid/JCOM\\_2006\\_2021\\_A04/](https://jcom.sissa.it/article/pubid/JCOM_2006_2021_A04/).
- 3.15.10. Leach, Brandi, *et al.* Emerging developments in citizen science, Reflecting on areas of innovation, *RAND Europe* [2020] // [https://www.rand.org/pubs/research\\_reports/RR4401.html](https://www.rand.org/pubs/research_reports/RR4401.html).
- 3.15.11. LR akademinės etikos ir procedūrų kontrolieriaus tarnybos patvirtintos Atitikties mokslinių tyrimų etikai vertinimo gairės [2020] // <https://www.e-tar.lt/portal/lt/legalAct/2c05c6603d4011ec992fe4cdfceb5666/asr>.
- 3.15.12. Mutual learning exercise, Citizen science initiatives: policy and practice: final report [2023] // <https://op.europa.eu/en/publication-detail/-/publication/63baa6bb-d359-11ed-a05c-01aa75ed71a1/language-en>.



- 3.15.13. Rainey, Stephen. *Philosophical Perspectives on Brain Data* (Cham: Palgrave Macmillan, 2023).
- 3.15.14. Ridley, Julie, *et al.* Methodological Framework for Data Collection and Analysis, *YouCount* [2022] // [https://www.youcountproject.eu/resources/project-reports/graffiti-tag-culture/download/15\\_d7c3c3b5730cff1a759d67a101bb4c78](https://www.youcountproject.eu/resources/project-reports/graffiti-tag-culture/download/15_d7c3c3b5730cff1a759d67a101bb4c78).
- 3.15.15. Survey report, Data management in Citizen Science projects [2016] // <https://op.europa.eu/en/publication-detail/-/publication/b6799013-1d83-11e6-ba9a-01aa75ed71a1/language-en>.
- 3.15.16. VDU Etikos kodeksas [2024] // [https://www.vdu.lt/wp-content/uploads/2024/03/Etikos-kodeksas\\_2024-1.pdf](https://www.vdu.lt/wp-content/uploads/2024/03/Etikos-kodeksas_2024-1.pdf).
- 3.15.17. VDU Mokslinių tyrimų atitikties pagrindiniams tyrimų profesionalumo ir etikos principams vertinimo nuostatai [2021] // <https://www.vdu.lt/wp-content/uploads/2021/04/Moksliniu-tyrimu-atitikties-pagrindiniams-tyrimu-profesionalumo-ir-etikos-principams-vertinimo-nuostatos.pdf>.

## IV SKYRIUS. FAIR PRINCIPAI, DUOMENŲ STANDARTAI IR DUOMENŲ KOKYBĖ

### 4.1. FAIR principai

- 4.1.1. FAIR principai<sup>1</sup> – surandamumas, prieinamumas, sąveikumas ir pakartotinis naudojimas (angl. *Findable, Accessible, Interoperable, Reusable*) – yra esminiai užtikrinant, kad moksliniai duomenys būtų lengvai randami, prieinami, sąveikūs ir pakartotinai naudojami. Šie principai ne tik skatina duomenų kokybę, bet ir yra pagrindinė Europos Sąjungos atvirojo mokslo strategijos ir EOSC (*European Open Science Cloud*) platformos bei nacionalinės atvirųjų mokslo duomenų erdvės dalis. FAIR principai taikomi trimis esybėms: duomenims (ar bet kokiam skaitmeniniam objektui), metaduomenims (informacijai apie tą skaitmeninį objektą) ir infrastruktūrai. Pavyzdžiui, F4 principas (žr. toliau) apibrėžia, kad tiek metaduomenys, tiek duomenys turi būti registruoti arba indeksuoti paieškai tinkamame šaltinyje (infrastruktūros komponente).
- 4.1.2. *Surandamumo principas* (angl. *Findable*), kuriam keliami pagrindiniai reikalavimai pateikti F1–F4. Pirmasis žingsnis naudojant (arba pakartotinai naudojant) duomenis yra juos surasti. Metaduomenys ir duomenys turėtų būti lengvai randami tiek žmonėms, tiek kompiuteriams. Mašininio būdu nuskaityti metaduomenys yra būtini automatiniams duomenų rinkinių ir paslaugų atradimui, todėl tai yra esminė FAIR principų įgyvendinimo proceso dalis.
- 4.1.2.1. **F1.** (Meta)duomenims priskiriami globaliai unikalūs ir nuolatiniai identifikatoriai.
- 4.1.2.2. **F2.** Duomenys aprašomi išsamiais metaduomenimis.
- 4.1.2.3. **F3.** Metaduomenys aiškiai ir nedviprasmiškai nurodo aprašomų duomenų identifikatorių.
- 4.1.2.4. **F4.** (Meta)duomenys registruojami arba indeksuojami paieškos sistemose.

---

<sup>1</sup> Plačiau žr.: <https://www.go-fair.org/fair-principles/>.

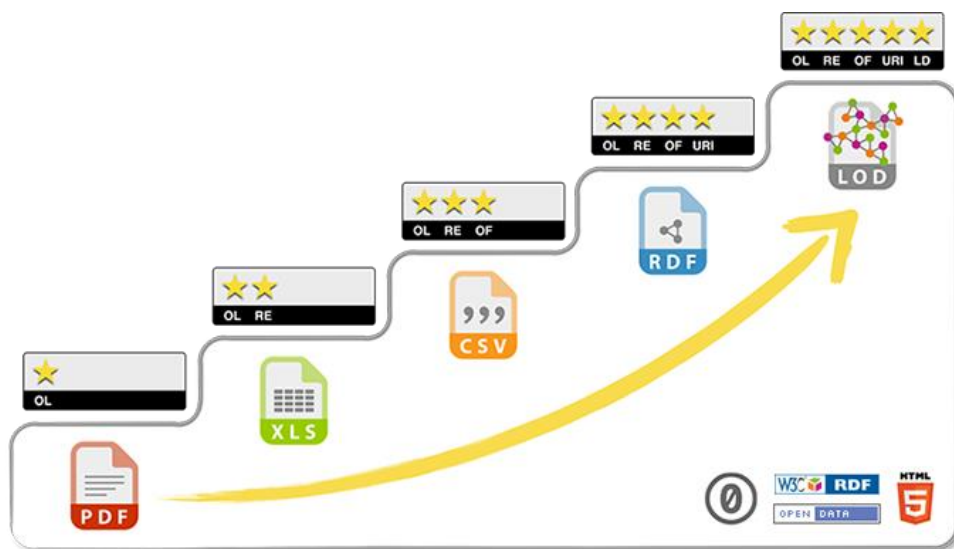
- 4.1.3. *Prieinamumo principas* (angl. *Accessible*), , kuriam keliami pagrindiniai reikalavimai pateikti A1 ir A2. Kai naudotojas suranda reikalingus duomenis, jis turi žinoti, kaip prie jų galima prieiti, įskaitant galimus autentifikavimo ir autorizacijos procesus.
- 4.1.3.1. **A1.** (Meta)duomenys pasiekiami per jų identifikatorių, naudojant standartizuotą komunikacijos protokolą:
- 4.1.3.1.1. **A1.1.** Protokolas yra atviras, nemokamas ir universaliai įgyvendinamas.
- 4.1.3.1.2. **A1.2.** Protokolas, jei reikia, leidžia autentifikavimo ir autorizacijos procedūras.
- 4.1.3.2. **A2.** Metaduomenys išlieka prieinami net ir tada, kai duomenys nebepasiekiami.
- 4.1.4. *Sąveikumo principas* (angl. *Interoperable*), , kuriam keliami pagrindiniai reikalavimai pateikti I1–I3. Duomenys dažnai turi būti integruojami su kitais duomenimis. Be to, duomenys turi būti suderinami su programomis ar darbo procesais, skirtais analizei, saugojimui ir apdorojimui.
- 4.1.4.1. **I1.** (Meta)duomenys naudoja formalizuotą, prieinamą, bendrą ir plačiai taikomą žinių reprezentavimo kalbą.
- 4.1.4.2. **I2.** (Meta)duomenys naudoja žodynus, atitinkančius FAIR principus.
- 4.1.4.3. **I3.** (Meta)duomenys turi kvalifikuotas nuorodas į kitus (meta)duomenis.
- 4.1.5. *Pakartotinio naudojimo principas* (angl. *Reusable*), , kuriam keliami pagrindiniai reikalavimai pateikti R1.1–R1.3. Pagrindinis FAIR principų tikslas yra optimizuoti duomenų pakartotinį naudojimą. Tam pasiekti metaduomenys ir duomenys turėtų būti gerai aprašyti, kad juos būtų galima atkurti ir (ar) derinti skirtinguose kontekstuose.
- 4.1.5.1. **R1.** (Meta)duomenys yra išsamiai aprašyti daugybe tikslų ir aktualių atributų:
- 4.1.5.1.1. **R1.1.** (Meta)duomenys pateikiami su aiškia ir prieinama naudojimo licencija.
- 4.1.5.1.2. **R1.2.** (Meta)duomenys turi išsamią kilmės informaciją.
- 4.1.5.1.3. **R1.3.** (Meta)duomenys atitinka konkrečios srities bendruomenės standartus.

## 4.2. FAIR: sąveikumo ir pakartotinio naudojimo įgyvendinimas

- 4.2.1. Laikantis FAIR *sąveikumo* (angl. *Interoperability*) principo, itin svarbu įgyvendinti Tim Berners-Lee penkių žvaigždučių sistemos „Linked Data“ reikalavimus<sup>2</sup>, nes tai užtikrina, kad duomenys būtų lengvai suprantami, naudojami ir susiejami su kitais duomenimis.

---

<sup>2</sup> Plačiau žr.: <https://5stardata.info/en/>.



**Tim Berners-Lee penkių žvaigždučių sistema.** Šaltinis: <https://5stardata.info/en/>

- 4.2.2. ★ *Duomenys pateikiami kaip atviri duomenys.* Duomenų teikėjas suvokia, kad duomenų laikymas uždaro arba ribotos prieigos saugykloje nėra veiksmingas, nes duomenys tikrai duoda naudą tik tada, kai kiti gali juos pasiekti ir pakartotinai naudoti. Todėl duomenų kūrėjas nusprendžia dalytis savo projekto duomenimis su išoriniu pasauliu pagal atvirą licenciją. Tačiau gali būti, kad duomenų kūrėjas aktualiai duomenis pateikia mažai optimaliu, nemasiniu būdu nuskaitomu formatu (pvz., lentelės rezultatai kaip nuskaitytas PDF failas arba PDF su matavimų sąrašu). Dėl šios priežasties taip atverti duomenys vertinami viena žvaigždute.
- 4.2.3. ★★ *Struktūruoti ir mašininio būdu nuskaitomi duomenys (angl. machine readable).* Duomenų kūrėjas savo projekto duomenis pateikia ne tik pagal atvirą licenciją, bet ir struktūrizuotu (taigi ir mašininio būdu nuskaitomu) formatu. Pavyzdžiui, tekstas pateikiamas „MS Word“ dokumente, skaičiai – „MS Excel“ failuose, o aukštos kokybės nuotraukos pateikiamos RAW formatu. Tai leidžia pasiekti dviejų žvaigždučių duomenų vertinimą. Vartotojai dabar gali vizualizuoti duomenis arba atlikti su jais skaičiavimus, jei turi tinkamą programinę įrangą šiems duomenims apdoroti.
- 4.2.4. ★★★ *Saugomi atvirais duomenų formatais.* Tai žingsnis link didesnio duomenų prieinamumo ir naudojimo. Šio lygmens duomenys pateikiami atvirais duomenų formatais pateikiami tokiu formatu, kuris yra atviras, nemokamas ir nepriklauso nuo jokios konkrečios programinės įrangos (ypač turinčios komercinės licencijos ribojimus). Pavyzdžiui, CSV, JSON, XML yra atviri formatai, o „Excel“ (XLS)

neatviras, nes jis priklauso nuo „Microsoft“ programinės įrangos. Atvirieji formatai yra plačiai palaikomi formatai, t. y. duomenys turi būti suderinami su daugybe įrankių, leidžiančių vartotojams juos atidaryti, redaguoti ir analizuoti. Šio lygmens duomenys turi aukštesnio lygio prieinamumą, nes vartotojams nereikia pirkti specifinės programinės įrangos ar susidurti su formatais, kurių jų turimi įrankiai nepalaiko. Šio lygmens naudos: 1) prieinamumas (duomenys tampa plačiau prieinami, nes jų naudojimui nereikia konkrečių įrankių ar licencijų); 2) pakartotinis naudojimas (duomenis galima lengvai apdoroti skirtingose sistemose ir pritaikyti įvairioms reikmėms); 3) bendradarbiavimo skatinimas (atviri formatai leidžia, kad daugiau vartotojų gali dalyvauti analizuojant ir naudojant duomenis).

4.2.5. ★★★★★ *Globaliai unikalūs identifikatoriai (DOI, URI)*. Pasiekus keturių žvaigždučių lygmenį, duomenys tampa ypač vertingi. Nuo šio lygmens bet kas gali vienareikšmiškai susieti savo duomenis su konkrečiu duomenų egzemplioriumi kitame duomenų rinkinyje. Pvz., tyrėjas biologas norėtų susieti savo tyrimus su 100 įrašų iš tam tikro duomenų rinkinio, tačiau jis supainiojo vieną duomenų rinkinį su kitu, todėl padarė neteisingas išvadas. Ateityje to būtų galima išvengti naudojant keturių žvaigždučių lygmens duomenis, kurie turi globaliai unikalų, nuolatinį identifikatorių. Surinktų duomenų reikšmę taip pat gali nustatyti kitos sistemos, net visiškai automatiškai, naudojant duomenų URI adresus. Nuo šio lygmens duomenų rinkinys jau gali būti įtrauktas į „Linked Open Data“ (LOD) tinklą. Kiekvienas duomenų rinkinys ar jo elementas gali būti pasiekiamas naudojant nuolatinį identifikatorių, o tai užtikrina patikimumą ir galimybę susieti duomenis su kitais ištekliais. Šiame etape sukuriamos nuorodos į kitus duomenų rinkinius, suteikiant semantinį kontekstą ir didinant duomenų pritaikomumą. Šio lygmens naudos: 1) vienareikšmiškumas (globaliai unikalūs identifikatoriai padeda išvengti duomenų painiavos ar dubliavimo); 2) sąveikumas (duomenys tampa suderinami su kitais šaltiniais, leidžiant juos lengvai integruoti ir analizuoti); 3) pasitikėjimas (aiškus identifikavimas užtikrina duomenų šaltinio patikimumą).

4.2.6. ★★★★★ *Susieti su išoriniais šaltiniais*. Penkių žvaigždučių lygmuo yra aukščiausias „Linked Open Data“ standartų lygis. Jis siekia užtikrinti, kad duomenys būtų ne tik atviri, lengvai naudojami ir globaliai identifikuojami, bet ir susieti su kitais duomenų šaltiniais, sukuriant semantiškai turtingą ir integruotą duomenų ekosistemą. Duomenys turi būti susieti su kitais atvirais duomenų rinkiniais, naudojant semantinius standartus,

tokius kaip RDF (angl. *Resource Description Framework*). Susiejimas suteikia papildomą kontekstą ir leidžia integruoti duomenis su globaliu „Linked Open Data“ tinklu. Šio lygmens duomenų elementai turi aiškiai apibrėžtas nuorodas į kitų duomenų šaltinių elementus. Pavyzdžiui, tam tikras asmuo ar vieta viename duomenų rinkinyje gali būti susieta su papildoma informacija iš kitų šaltinių. Šiame lygmenyje pasiekiamas semantinis suderinamumas. Naudojami bendri ontologijų ir žodynų standartai, kurie užtikrina, kad duomenų prasmė yra vienodai suprantama skirtingose sistemose. Galima SPARQL paieška duomenyse. Duomenys tampa prieinami ir užklausomi naudojant standartizuotą SPARQL kalbą, leidžiančią vartotojams atlikti sudėtingas užklausas ir gauti išplėstinius rezultatus. Šio lygmens naudos: 1) visapusiška integracija (duomenys tampa prasmingai susiję, todėl vartotojai gali atlikti sudėtingas analizes ir gauti daugiau naudingų įžvalgų); 2) pritaikomumas (susieti duomenys gali būti lengvai naudojami įvairiose srityse, pavyzdžiui, moksliniuose tyrimuose, vyriausybės sprendimuose ar verslo procesuose); 3) didžiausios vertės kūrimas (penkių žvaigždučių duomenys ne tik yra lengvai pasiekiami ir naudojami, bet ir prisideda prie globalios atvirų duomenų ekosistemos kūrimo).

### 4.3. FAIR: surandamumo įgyvendinimas – metaduomenys

- 4.3.1. Metaduomenų vaidmuo užtikrinant duomenų surandamumą yra esminis FAIR principų įgyvendinimo aspektas. Metaduomenys, apibūdinantys duomenų rinkinį, leidžia juos lengvai identifikuoti, aprašyti ir rasti tiek žmonėms, tiek kompiuteriams. Pagrindinis tikslas – užtikrinti, kad duomenys būtų lengvai prieinami ir randami naudojant tinkamus paieškos įrankius ar sistemas. Be išsamiai aprašytų metaduomenų, net pats vertingiausias duomenų rinkinys gali likti nepastebėtas ar nepritaikytas pakartotiniam naudojimui.
- 4.3.2. Metaduomenys suteikia informaciją apie duomenų šaltinį, turinį, struktūrą, laikymo vietą ir licencijavimo sąlygas. Tai leidžia vartotojams greitai suprasti, ar duomenys atitinka jų poreikius, ir užtikrina, kad duomenys galėtų būti pasiekti pagal nustatytus standartus. Standartizuoti metaduomenys taip pat padeda duomenis registruoti ar indeksuoti specializuotose paieškos sistemose ir kataloguose, kur jie tampa prieinami platesnei vartotojų grupei. Šio proceso rezultatas – efektyvesnė duomenų paieška ir geresnis jų valdymas.

- 4.3.3. EOSC ir nacionalinė atvirųjų duomenų erdvė, kaip platforma, skirta mokslinių duomenų ir paslaugų dalijimuisi, remiasi FAIR principais, kad užtikrintų visų duomenų rinkinių, įkeltų į platformą, aukštą kokybę, prieinamumą ir sąveikumą.
- 4.3.4. EOSC ir nacionalinės atvirųjų duomenų erdvės vaidmuo *surandamumo* principo įgyvendinime yra glaudžiai susijęs su standartizuotų metaduomenų naudojimu. Visi duomenys, įkeliami į EOSC arba nacionalinę atvirųjų duomenų erdvę, privalo turėti išsamius, standartizuotus ir mašininio būdu nuskaitomus metaduomenis. Šie metaduomenys leidžia duomenis lengvai integruoti į EOSC arba nacionalinių atvirųjų duomenų erdvės katalogus ir užtikrina, kad juos galima rasti naudojant paieškos įrankius, tokius kaip EOSC arba nacionalinės atvirųjų duomenų erdvės portalas. Pavyzdžiui, naudojant tokias metaduomenų schemas, kaip „DCAT-AP“ ar „Dublin Core“, galima aprašyti duomenų turinį, jų kūrėjus, prieigos sąlygas ir kitą svarbią informaciją.
- 4.3.5. EOSC ir nacionalinė atvirųjų duomenų erdvė skatina *sąveikumo* ir *surandamumo* principų įgyvendinimą ne tik tarp skirtingų duomenų rinkinių, bet ir tarp nacionalinių ir teminių duomenų infrastruktūrų. Tai reiškia, kad kiekvienas duomenų rinkinys, integruotas į EOSC arba nacionalinę atvirųjų duomenų erdvę, gali būti susietas su kitais duomenimis ar paslaugomis, sudarydamas vientisą duomenų tinklą. Metaduomenys, leidžiantys automatizuotai susieti duomenis ir juos rasti, yra šio tinklo pagrindas.
- 4.3.6. FAIR principų, ypač *surandamumo*, integracija EOSC ir nacionalinėje atvirųjų duomenų erdvėje padidina duomenų pakartotinio naudojimo galimybes. Mokslininkai, įmonės ir visuomenė gali lengviau rasti reikalingus duomenis, užtikrindami, kad moksliniai tyrimai būtų efektyvesni, ekonomiškesni ir skaidresni. Be to, standartizuoti metaduomenys užtikrina, kad EOSC platformoje esantys duomenys galėtų būti naudojami įvairiais būdais, įskaitant automatizuotas analizes ir dirbtinio intelekto taikymus.
- 4.3.7. Mokslo duomenys yra labai nevienalypiai, todėl vieno standarto taikymas visiems duomenų tipams nėra tinkamas sprendimas, kaip tai yra, pavyzdžiui, ES viešajame sektoriuje naudojant vieningą DCAT-AP bazinį standartą. Mokslinių duomenų įvairovė kyla dėl skirtingų disciplinų, tyrimo metodikų, eksperimentų pobūdžio ir įvairių formatų, kurie reikalingi specifiniams analizės ar apdorojimo poreikiams. Mokslinės bendruomenės kuria specializuotus standartus, atitinkančius konkrečių sričių poreikius (pvz., NetCDF klimatologijoje, BioSchemas biologijoje).

4.3.8. Užuoat siekus universalaus standarto, FAIR principai siūlo bendrą struktūrą, kuri gali būti pritaikyta įvairių sričių standartams. Tai reiškia, kad visi standartai turi užtikrinti duomenų surandamumą, prieinamumą, sąveikumą ir pakartotinį naudojimą. FAIR principai taip pat palaiko modulinę metaduomenų struktūrą: galima kurti standartus, kuriuos būtų lengva išplėsti arba pritaikyti skirtingiems disciplininiais poreikiams.

## 4.4. Praktinės rekomendacijos metaduomenų formavimui, laikantis FAIR principų

Rekomenduojama:

4.4.1. Užtikrinti surandamumą:

4.4.1.1. Kiekvienas duomenų rinkinys turi būti priskirtas globaliai unikaliam identifikatoriui (pvz., DOI, URI).

4.4.1.2. Parengti metaduomenis pagal pripažintus standartus (pvz., „Dublin Core“ arba „DataCite“) ir užtikrinti, kad jie būtų mašininio būdu nuskaitomi.

4.4.2. Užtikrinti prieinamumą:

4.4.2.1. Naudoti atvirus prieigos protokolus (pvz., HTTPS, FTP) ir aiškiai nurodyti duomenų licencijas (pvz., „Creative Commons“).

4.4.2.2. Užtikrinti, kad metaduomenys liktų prieinami net tada, kai duomenys yra ištrinami.

4.4.3. Užtikrinti sąveikumą:

4.4.3.1. Pasirinkti bendras žinių reprezentavimo kalbas, tokias kaip RDF, JSON-LD ar OWL.

4.4.3.2. Pritaikyti disciplinos specifines ontologijas ir žodynus (pvz., AGROVOC žemės ūkyje ar SNOMED CT medicinoje).

4.4.4. Skatinti pakartotinį naudojimą:

4.4.4.1. Įtraukti aiškias licencijų nuostatas ir duomenų kilmės informaciją.

4.4.4.2. Naudoti standartus, kurie leidžia išsaugoti svarbias disciplinines savybes ir metaduomenis.

## 4.5. Duomenų kokybė



- 4.5.1. Duomenų kokybė yra kritiškai svarbi atvirųjų mokslinių duomenų atveju, ypač įgyvendinant FAIR principus ir integruojant juos į EOSC arba nacionalinę atvirųjų duomenų erdvės platformą. Duomenų kokybė ne tik lemia mokslinių rezultatų patikimumą, bet ir užtikrina jų tinkamumą pakartotiniam naudojimui, sąveikumui bei platesnei sklaidai tarp disciplinų.
- 4.5.2. FAIR principai yra orientuoti į duomenų prieinamumą ir pakartotinį naudojimą, o tam būtina aukšta duomenų kokybė. Konkretūs aspektai:
- 4.5.2.1. *Surandamumas* (angl. *Findable*). Duomenys turi būti aprašyti išsamiais ir tiksliais metaduomenimis, kurie padeda vartotojams ir kompiuteriams juos lengvai surasti. Netikslūs ar nepilni metaduomenys gali lemti, kad duomenys bus nepastebėti ar netinkamai interpretuoti.
- 4.5.2.2. *Prieinamumas* (angl. *Accessible*). Duomenys turi būti pateikiami patikimose ir standartizuotose saugyklose, kurios užtikrina jų ilgaamžiškumą ir prieinamumą. Žemos kokybės duomenys gali trukdyti analizei arba reikalauti papildomų išteklių juos paruošti naudoti.
- 4.5.2.3. *Sąveikumas* (angl. *Interoperable*). Aukštos kokybės duomenys leidžia juos integruoti su kitais šaltiniais, naudojant standartizuotas struktūras, formatus ir ontologijas. Prastai parengti duomenys gali sukelti konfliktus tarp skirtingų sistemų.
- 4.5.2.4. *Pakartotinis naudojimas* (angl. *Reusable*). Kad duomenys būtų tinkami pakartotiniam naudojimui, jie turi būti išsamūs, nuoseklūs, turėti aiškias licencijas ir aprašytus kilmės šaltinius. Žemos kokybės duomenys riboja jų pritaikomumą naujiems tyrimams.
- 4.5.3. EOSC siekia sukurti visą Europą apimančią atvirųjų mokslinių duomenų platformą, o duomenų kokybė yra esminė jos veikimo sąlyga. Kokybiški duomenys yra būtini, siekiant:
- 4.5.3.1. *Tarpdisciplininio bendradarbiavimo*. Mokslininkai iš skirtingų sričių gali dalytis duomenimis ir juos integruoti į naujas studijas tik tuo atveju, jei duomenys yra nuoseklūs ir patikimi.
- 4.5.3.2. *Duomenų sąveikumo palaikymo*. EOSC remiasi standartizuotais metaduomenimis ir formatais, todėl prastai paruošti duomenys gali nesuderėti su platformos reikalavimais.
- 4.5.3.3. *Pakartotinio naudojimo skatinimo*. Aukštos kokybės duomenys, kurių kilmė yra aiški ir kurie pateikiami standartizuotais formatais, skatina mokslininkus pasitikėti platforma ir naudotis jos resursais.

## 4.6. Praktinės rekomendacijos duomenų kokybei užtikrinti

Duomenų kokybei užtikrinti rekomenduojamas:

### 4.6.1. Standartų taikymas.

4.6.1.1. Tikslųjų ir gamtos mokslų srityje rekomenduojama laikytis šių tarptautinių standartų: ISO 19157 (geografinių duomenų kokybei, apibrėžia geografinių duomenų kokybės elementus, matavimo kriterijus ir metodus, tinka erdviniams (GIS) duomenims, užtikrinant tikslumą ir geografinį suderinamumą); MIAME (angl. *Minimum Information About a Microarray Experiment*) (biologijoje naudojamas standartas *mikroarray* eksperimentų duomenims aprašyti ir vertinti, užtikrina, kad biologiniai duomenys būtų pakankamai išsamūs ir pritaikomi pakartotiniams tyrimams); CF (angl. *Climate and Forecast*) NetCDF (klimatologijos ir meteorologijos duomenų kokybei ir formato valdymui skirtas standartas, užtikrinantis suderinamumą ir nuoseklumą tarp klimatologinių duomenų rinkinių); DICOM (angl. *Digital Imaging and Communications in Medicine*) (medicininiems vaizdų duomenų sistemoms skirtas standartas, užtikrinantis duomenų struktūros, formato ir kokybės suderinamumą) ir kt.

4.6.1.2. Humanitarinių ir socialinių mokslų srityse duomenys dažnai būna tekstiniai, vaizdiniai ar garso, todėl jų kokybės užtikrinimui ir valdymui būtina laikytis specifinių standartų. Šių sričių duomenys yra labai įvairūs, todėl reikalingi lankstūs, tačiau struktūrizuoti metodai, kurie padėtų užtikrinti jų tikslumą, nuoseklumą, pakartotinį naudojimą ir sąveikumą, pritaikomi specifiniams disciplininiais poreikiams. Naudojant tokius standartus, kaip TEI, DDI ar CIDOC CRM, užtikrinama duomenų kokybė, jų sąveikumas ir pritaikymas skirtinguose kontekstuose. Tai ypač svarbu, siekiant suderinti tradicinius mokslinių tyrimų procesus su šiuolaikinėmis skaitmeninių duomenų valdymo praktikomis.

4.6.2. Metaduomenų tobulinimas. Rekomenduojama rinkti ir pateikti visą reikalingą informaciją apie duomenų kilmę, surinkimo būdus ir apdorojimą.

4.6.3. Automatinių kokybės tikrinimo priemonių naudojimas.

4.6.3.1. Piliečių mokslo projektuose, kur dažnai dirbama su dideliais duomenų kiekiais, kuriuos pateikia įvairūs tyrimo dalyviai, labai svarbu užtikrinti duomenų kokybę. DI technologijos gali būti labai naudingos automatizuojant kokybės tikrinimą, duomenų analizę ir jų tikslumo užtikrinimą.

- 4.6.3.2. Naudojant DI įrankius, svarbu ne tik pasirinkti tinkamą platformą, bet ir užtikrinti, kad duomenų kokybės kriterijai būtų aiškiai apibrėžti projekto pradžioje. Paminėtini šie naudingi įrankiai ir platformos:
- 4.6.3.3. *OpenRefine*<sup>3</sup> – duomenų valymui ir tikrinimui skirtas įrankis. Tinka įvairių tipų duomenims struktūrizuoti, valyti, tikrinti duomenų suderinamumą ir šalinti pasikartojimus. Plėtiniai ir DI modeliai leidžia aptikti anomalijas arba numatyti trūkstamus duomenis. Lengvai naudojamas ir pritaikomas piliečių mokslo duomenims, ypač didelėms duomenų bazėms.
- 4.6.3.4. *DataRobot*<sup>4</sup> – DI platforma automatizuotai duomenų analizei ir kokybei tikrinti. Aptinka anomalijas, užpildo trūkstamus duomenis ir generuoja duomenų kokybės ataskaitas. Integruojama su įvairiomis platformomis, palaiko sudėtingus statistinius ir prognozavimo modelius.
- 4.6.3.5. *Trifacta (Alteryx Designer Cloud)*<sup>5</sup> – duomenų kokybės valdymo ir DI modelių kūrimo platforma. Valo ir tikrina duomenis, aptinka pasikartojimus, formatuoja įvairius duomenų šaltinius. Vizualus valdymas, tinka didelėms duomenų bazėms piliečių mokslo projektuose.
- 4.6.3.6. *Great Expectations*<sup>6</sup> – duomenų testavimo ir tikrinimo platforma. Sukuria „lūkesčius“, kurie tikrina, ar duomenys atitinka nustatytus kokybės kriterijus (pvz., duomenų diapazoną, formatą). Galima integruoti su DI sprendimais anomalijoms aptikti ir „lūkesčių“ modeliui pritaikyti. Integravimas su DI sprendimais gali užtikrinti tiek automatizuotą duomenų kokybės kontrolę, tiek lengvesnį duomenų apdorojimą ir interpretaciją. Lanksti ir pritaikoma įvairiose disciplinose.
- 4.6.3.7. *FAIR Evaluator*<sup>7</sup> – internetinė platforma, leidžianti automatiškai vertinti duomenų ir metaduomenų atitikimą FAIR principams. Naudojant globalius identifikatorius (pvz., DOI arba URI), vertinami metaduomenų elementai, tokie kaip prieinamumas, licencijos ir sąveikumas. Suteikia detalizuotas ataskaitas apie atitikimą FAIR kriterijams ir pateikia rekomendacijas, kaip tobulinti duomenis.

---

<sup>3</sup> Plačiau žr.: <https://openrefine.org/>

<sup>4</sup> Plačiau žr.: <https://www.datarobot.com/>

<sup>5</sup> Plačiau žr.: <https://asana.com/apps/trifacta> <https://www.alteryx.com/about-us/trifacta-is-now-alteryx-designer-cloud>

<sup>6</sup> Plačiau žr.: <https://greatexpectations.io/>

<sup>7</sup> Plačiau žr.: <https://egonw.github.io/cookbook-dev/content/recipes/assessing-fairness/fair-assessment-fairevaluator.html>

4.6.3.8. *F-UJI FAIR Data Assessment Tool*<sup>8</sup> – automatinis įrankis, skirtas FAIR duomenų vertinimui, remiantis *FAIR Data Maturity Model*. Naudojant duomenų rinkinį ar jo metaduomenis (pvz., DOI), atliekamas automatinis atitikties vertinimas. Greita analizė, pateikianti aiškias rekomendacijas tobulinimui.

---

<sup>8</sup> Plačiau žr.: <https://www.f-uji.net/>

## V SKYRIUS. BAIGIAMOSIOS NUOSTATOS

- 5.1. Šios parengtos Gairės atspindi asmeninę Gairių rengėjų nuomonę, bet nėra oficiali Vytauto Didžiojo universiteto pozicija. Gairės nėra patvirtintos Vytauto Didžiojo universiteto vidaus teisės aktu.
- 5.2. Gairių rengėjai rekomenduoja šias gaires patvirtinti Vytauto Didžiojo universiteto vidaus teisės aktu.