VALSTYBĖS DUOMENŲ AGENTŪROS
GENERALINIS DIREKTORIUS

ĮSAKYMAS
DĖL DUOMENŲ NUASMENINIMO IR PSEUDONIMINIMO TVARKOS APRAŠO PATVIRTINIMO

2024 m. sausio 17 d. Nr. DĮ- 24
Vilnius

Siekdama užtikrinti duomenų konfidencialumą sudarant duomenų rinkinius ir atsižvelgdama į 2022 m. gruodžio 19 d. Metodinės komisijos posėdžio protokolą Nr. DP-107:

1. T v i r t i n u Duomenų nuasmeninimo ir pseudoniminimo tvarkos aprašą (pridedama).
2. P a v e d u:

2.1. dirbantiems pagal darbo sutartis Valstybės duomenų agentūros valstybės tarnautojams ir darbuotojams, kurie atlikdami priskirtas funkcijas tvarko duomenis pakartotiniam sveikatos duomenų panaudojimui, panaudojimui mokslo tikslams, rengia atvirų duomenų rinkinius ar atlieka su duomenų tvarkymu susijusius veiksmus, ir leidimą pakartotinai naudoti sveikatos duomenis turintiems asmenims, atliekantiems sveikatos duomenų rezultatų nuasmeninimą, jeigu teisė nuasmeninti rezultatus numatyta leidime, duomenų naudotojams, dirbantiems su pakartotinai naudojamais viešojo sektoriaus duomenimis, vadovautis šio įsakymo 1 punktu patvirtintu Duomenų nuasmeninimo ir pseudoniminimo tvarkos aprašu;

2.2. Metodologijos ir duomenų mokslo grupei, Konfidencialių duomenų valdymo komisijai, sudarytai Lietuvos statistikos departamento generalinio direktoriaus 2022 m. sausio 28 d. įsakymu Nr. DĮ-39 „Dėl Konfidencialių duomenų valdymo komisijos sudarymo ir jos darbo reglamento patvirtinimo“, esant poreikiui konsultuoti Valstybės duomenų agentūros valstybės tarnautojus ir darbuotojus, dirbančius pagal darbo sutartis, duomenų nuasmeninimo bei pseudoniminimo metodų taikymo klausimais.

3. R e k o m e n d u o j u Duomenų nuasmeninimo ir pseudoniminimo tvarkos aprašu vadovautis ir kitoms viešojo sektoriaus institucijoms, teikiančioms pakartotinai naudoti duomenis.

4. P r i p a ž į s t u netekusiu galios Valstybės duomenų agentūros generalinio direktoriaus 2023 m. sausio 17 d. įsakymą Nr. DĮ-13 „Dėl Duomenų nuasmeninimo ir pseudoniminimo tvarkos aprašo patvirtinimo“.

Generalinė direktorė Jūratė Petrauskienė

PATVIRTINTA
Valstybės duomenų agentūros
generalinio direktoriaus
2024 m. sausio 17 d. įsakymu Nr. DĮ-24

DUOMENŲ NUASMENINIMO IR PSEUDONIMINIMO TVARKOS APRAŠAS

I SKYRIUS
BENDROSIOS NUOSTATOS

1. Duomenų nuasmeninimo ir pseudoniminimo tvarkos aprašas (toliau – Aprašas) nustato pakartotinio viešojo sektoriaus duomenų naudojimo, sveikatos duomenų, duomenų, iš kurių rengiami atvirų duomenų rinkiniai, mokslo tikslams skirtų duomenų ir kitų Valstybės duomenų agentūros (toliau – agentūra) duomenų gavėjams teikiamų duomenų (toliau – duomenų) nuasmeninimo ir pseudoniminimo metodus ir tvarką.

2. Aprašo tikslas – užtikrinti duomenų konfidencialumą sudarant duomenų rinkinius.

3. Aprašu vadovaujasi agentūros valstybės tarnautojai ir darbuotojai, dirbantys pagal darbo sutartis (toliau – darbuotojai), ir leidimą pakartotinai naudoti sveikatos duomenis turintys asmenys (toliau – leidimo turėtojas), atliekantys sveikatos duomenų rezultatų nuasmeninimą, jeigu teisė nuasmeninti rezultatus numatyta leidime, duomenų naudotojai dirbantys su pakartotinai naudojamais viešojo sektoriaus duomenimis.

4. Esant poreikiui, agentūra turi teisę pasikviesti atitinkamos srities nepriklausomą (-us) ekspertą (-us).

II SKYRIUS
VARTOJAMOS SĄVOKOS

5. Apraše vartojamos sąvokos:

5.1. **Duomenų atskleidimo kontrolės procesas –** seka veiksmų, atliekamų, siekiant užtikrinti tinkamą duomenų nuasmeninimą ar pseudoniminimą.

5.2. **Duomenų objektas** – fizinis asmuo, šeima, namų ūkis, juridinis asmuo ar jo padalinys, kita organizacija ar jos padalinys, turtas, medžiagos, dokumentai, teisės, gamtos ištekliai ir objektai, reiškiniai, įvykiai, veiksmai, kiti objektai, apie kuriuos pateikiami duomenys.

5.3. **Kvaziidentifikatorius** – duomenų rinkinio kintamasis, kuris sujungtas su kitu (-ais) duomenų rinkinio kintamuoju (-aisiais) (arba kokia nors papildoma, išorine informacija) leidžia tiesiogiai arba su didele tikimybe nustatyti bent vieną duomenų objektą.

5.4. **Tiesioginis identifikatorius** – duomenų rinkinio kintamasis ar grupė kintamųjų, kuris (-ie) tiesiogiai nurodo duomenų objektą.

5.5. Kitos Apraše vartojamos sąvokos suprantamos taip, kaip jos apibrėžtos 2016 m. balandžio 27 d. Europos Parlamento ir Tarybos reglamente (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) (toliau – Reglamentas (ES) 2016/679), Lietuvos Respublikos pakartotinio sveikatos duomenų naudojimo įstatyme, 2022 m. gegužės 30 d. Europos Parlamento ir Tarybos reglamente (ES) 2022/868 dėl Europos duomenų valdymo, kuriuo iš dalies keičiamas Reglamentas (ES) 2018/1724 (Duomenų valdymo aktas) ir kituose teisės aktuose.

III SKYRIUS
TAIKOMI TEISĖS AKTAI

6. Nuasmeninant ir pseudoniminant duomenis, kartu su Aprašu taikomi teisės aktai:

6.1. Reglamentas (ES) 2016/679;

6.2. 2022 m. gegužės 30 d. Europos Parlamento ir Tarybos reglamentas (ES) 2022/868 dėl Europos duomenų valdymo, kuriuo iš dalies keičiamas Reglamentas (ES) 2018/1724 (Duomenų valdymo aktas);

6.3. Lietuvos Respublikos asmens duomenų teisinės apsaugos įstatymas;

6.4. Lietuvos Respublikos pakartotinio sveikatos duomenų naudojimo įstatymas;

6.5. Lietuvos Respublikos teisės gauti informaciją ir duomenų pakartotinio naudojimo įstatymas;

6.6. Lietuvos Respublikos oficialiosios statistikos ir valstybės duomenų valdysenos įstatymas;

6.7. Lietuvos Respublikos civilinis kodeksas.

7. Agentūros darbuotojai prireikus konsultuojasi su agentūros Metodologijos ir duomenų mokslo grupe (dėl nuasmeninimo ir pseudoniminimo metodų taikymo), Duomenų apsaugos skyriumi (dėl Reglamento (ES) 2016/679) nustatytų asmens duomenų apsaugos reikalavimų taikymo), Konfidencialių duomenų valdymo komisija, sudaryta agentūros generalinio direktoriaus, patvirtinta 2022 m. sausio 28 d. įsakymu Nr. DĮ-39 „Dėl Konfidencialių duomenų valdymo komisijos sudarymo ir jos darbo reglamento patvirtinimo“ (toliau – Konfidencialių duomenų valdymo komisija) (dėl Aprašo taikymo ar probleminių duomenų nuasmeninimo ar pseudoniminimo atvejų).

IV SKYRIUS
DUOMENŲ ATSKLEIDIMO KONTROLĖS PROCESO ETAPAI

8. Sudarant duomenų rinkinius, privaloma laikytis duomenų atskleidimo kontrolės proceso etapų.

9. Duomenų atskleidimo kontrolės proceso etapai yra:

9.1. duomenų atskleidimo kontrolės poreikio nustatymas;

9.2. duomenų analizė;

9.3. duomenų atskleidimo rizikos vertinimas;

9.4. nuasmeninimo ir (ar) pseudoniminimo metodų parinkimas ir taikymas;

9.5. duomenų naudingumo ir informacijos praradimo vertinimas;

9.6. pakartotinio duomenų atskleidimo rizikos vertinimas;

9.7. sklaida, duomenų perdavimas leidimo turėtojui arba duomenų naudotojui.

V SKYRIUS
DUOMENŲ ATSKLEIDIMO KONTROLĖS POREIKIO NUSTATYMAS

10. Nustatant duomenų atskleidimo kontrolės poreikį, įvertinama, ar duomenų rinkinyje yra:

10.1. asmens duomenų;

10.2. įmonių (juridinių asmenų) duomenų, duomenų saugomų dėl komercinio konfidencialumo, įskaitant verslo, bendrovės ir profesinę paslaptį;

10.3. valstybės, tarnybos, banko paslaptį sudarančių ar kitų konfidencialių duomenų;

10.4. jautrių duomenų, nurodytų Jautrių duomenų sąraše (Aprašo 1 priedas);

10.5. duomenų, saugomų dėl trečiųjų šalių intelektinės nuosavybės;

10.6. duomenų, kurių teikimą ar atskleidimą riboja įstatymai ar kiti teisės aktai.

11. Jeigu duomenų rinkinyje yra Aprašo 10 punkte nurodytų duomenų, pereinama į antrą duomenų atskleidimo kontrolės proceso etapą, išskyrus šiuos atvejus:

11.1. fizinis asmuo, juridinis asmuo ar jo padalinys, kita organizacija ar jos padalinys, kuriam priklauso duomenys, davė sutikimą šiuos duomenis viešai skelbti arba juos perduoti leidimo turėtojui;

11.2. galiojančiuose teisės aktuose numatyta teisė arba pareiga viešai skelbti duomenų rinkinį sudarančius duomenis.

12. Jeigu duomenų rinkinyje nėra Aprašo 10 punkte nurodytų duomenų, taip pat Aprašo 11.1,11.2 papunkčiuose nurodytais atvejais pereinama į paskutinį (sklaidos arba duomenų perdavimo leidimo turėtojui) etapą.

VI SKYRIUS
DUOMENŲ ANALIZĖ

13. Šio etapo metu nustatoma su duomenų konfidencialumu susijusi informacija: identifikatoriai, kvaziidentifikatoriai, Aprašo 10 punkte nurodyti duomenys. Ji bus naudojama kitame etape – atliekant duomenų atskleidimo rizikos vertinimą.

14. Įvertinama, ar yra ribotai asmenų grupei arba viešai prieinamų duomenų, kurie tapatūs ar panašūs į duomenų rinkinio duomenis arba su jais susiję. Įvertinama tokių duomenų gausa, prieigą prie tokių duomenų turinčių asmenų skaičius ir pobūdis, tokių duomenų jungimo su duomenų rinkinio duomenimis galimybė ir tikimybė. Į vertinimo rezultatus turi būti atsižvelgiama, parenkant taikytiną privatumo modelį duomenų atskleidimo rizikos vertinimo etape, parenkant ir taikant nuasmeninimo ir (ar) pseudoniminimo metodus, taip pat analizuojant atlikto nuasmeninimo ir (ar) pseudoniminimo tinkamumą pakartotinio duomenų atskleidimo rizikos vertinimo etape.

VII SKYRIUS
DUOMENŲ ATSKLEIDIMO RIZIKOS VERTINIMAS

15. Šio duomenų atskleidimo kontrolės proceso etapo metu nustatomos duomenų perdavimo tretiesiems asmenims ar duomenų viešo atskleidimo keliamos rizikos.

16. Atliekant rizikos vertinimą analizuojama, ar duomenų rinkinį sudarantys duomenys leidžia nustatyti arba gali būti panaudoti siekiant nustatyti su konkrečiu duomenų objektu susijusią informaciją.

17. Duomenų atskleidimo riziką galima vertinti atsižvelgiant į šias galimybes:

17.1. išskyrimo galimybė (angl. *singling out*), t. y. galimybė nustatyti duomenų objektą remiantis įrašo verčių retumu, unikalumu arba tiesiogiai iššifruojant pseudonimintas reikšmes;

17.2. susiejimo galimybė (angl. *linkability*), t. y. galimybė susieti bent du įrašus, susijusius su tuo pačiu duomenų objektu arba ta pačia duomenų objektų grupe (tame pačiame duomenų rinkinyje arba skirtinguose duomenų rinkiniuose), naudojantis kvaziidentifikatoriais. Jeigu galima nustatyti (pvz., atliekant koreliacijos analizę), kad du įrašai priskirti tai pačiai duomenų objektų grupei, tačiau negalima iš tos grupės išskirti pavienių duomenų objektų, tai rinkinys yra apsaugotas nuo išskyrimo, bet neužtikrinama apsauga nuo susiejimo;

17.3. išvados padarymo galimybė (angl. *inference*), t. y. galimybė dedukcijos būdu gana tiksliai nustatyti su duomenų objektu susijusios informacijos požymio vertę remiantis įrašo kitų požymių vertėmis.

18. Papildomai duomenų atskleidimo rizika gali būti vertinama ir atsižvelgiant į mozaikos efektą (angl. *Mosaic effect*) – duomenys gali būti jungiami su kita viešai prieinama informacija, tokiu būdu sužinant daugiau su duomenų objektu susijusios informacijos. Todėl būtina įvertinti kitus laisvai prieinamus duomenis ir sąsajas tarp jų ir duomenų rinkinį sudarančių duomenų.

19. Rizikai apibrėžti ir vertinti rekomenduojama pasitelkti šiuos privatumo modelius:

19.1. *k*-anonimiškumas (angl. *k-anonymity*) yra privatumo modelis, kuriuo siekiama panaikinti galimybę išskirti duomenų subjektus, juos grupuojant kartu su ne mažiau kaip *k-1* kitų asmenų. Duomenų rinkinys yra *k*-anonimiškas, jei kiekviena požymių kombinacijų reikšmė matoma bent *k* asmenų. Siekiant užtikrinti *k*-anonimiškumą, kategoriniai kintamieji gali būti sugrupuojami į aukštesnės klasės grupes, kiekybiniai – suskirstomi į intervalus. *k*-anonimiškumas gali būti pasiekiamas naudojant apibendrinimo metodus. Sudarant duomenų rinkinį, būtina atsižvelgti į visus

galimus kvaziidentifikatorius ir parinkti adekvačią *k* vertę. Rekomenduojama *k* vertė skiriasi pagal duomenų objektą – kai duomenų objektas yra asmenys, įmonės ir grupuojama ne pagal jautrius kintamuosius, *k* vertė yra 3. Kai grupuojama pagal specialių kategorijų, jautrius kintamuosius, *k* vertė turėtų būti bent 10.

19.2. *l*-įvairovės (angl. *l-diversity*) metodu išplečiamas *k*-anonimiškumo modelis, siekiant užtikrinti, kad nebūtų atskleidžiami duomenų objekto duomenys pasinaudojant tuo, kad visi *k* grupės duomenų objektai turi vienodą požymio vertę. Duomenų rinkinys yra *l*-įvairus, jei kiekvienoje lygiavertiškumo klasėje kiekvienam požymiui priskirta ne mažiau kaip *l* skirtingų reikšmių.

19.3. *t*-artumas (angl. *t-closeness*) yra *l*-įvairovės patobulinimas, kurio metu yra siekiama sukurti lygiavertes klases, panašias į pradinį pasiskirstymą. Reikalaujama, kad kiekvienoje lygiavertiškumo klasėje kiekvieno kintamojo pasiskirstymas būtų panašus į jo pasiskirstymą visoje populiacijoje (skirtųsi ne daugiau kaip per *t* pasirinkto skirstinių atstumo mato vienetų). Dėl šios priežasties *t*-artumas / tankis yra naudingas norint kaip įmanoma labiau išlaikyti pradinių duomenų struktūrą. Kai požymių vertės gali būti retos, tokį atstumą gali būti lengviau išlaikyti nei *l*-įvairovę.

19.4. *p* % taisyklė (angl. *p* % *rule*) tenkinama, jei kiekvieno požymio vertę bent vienam įrašui galima nustatyti ne tiksliau kaip su *p* procentų paklaida. Dažniausiai taikoma komercinių duomenų kontekste, kai pateikiamas agreguotas požymis, tam tikra suma, bet agreguotoje grupėje dominuoja žinomas stambus ūkio subjektas, o atėmus iš visos sumos dominuojančio subjekto požymio reikšmę, likusi suma yra mažesnė už didžiausios dalies *p* %. Tada grupės suma leidžia apytiksliai nuspėti stambaus subjekto požymio vertę. Rekomenduojama šią taisyklę taikyti su parametru p = 17,65.

19.5. Diferencinis privatumas (angl. *differential privacy*) – matematinis privatumo apibrėžimas, kuris tenkinamas, jei vieno įrašo pašalinimas iš duomenų rinkinio lemia mažą teikiamo rinkinio pokytį (pokytis, matuojant tam tikroje skalėje, turi būti ne didesnis nei ε). Privatumo išsaugojimo laipsnis priklauso nuo pasirinkto privatumo parametro ε, kuris turi būti atsakingai parenkamas, kadangi per didelis privatumo parametras gali netenkinti siekiamos apsaugos.

20. Kaip papildoma priemonė, leidžianti įvertinti duomenų atskleidimo riziką, gali būti naudojama Rizikų matrica (Aprašo 2 priedas).

21. Nustačius atskleidimo riziką 16–19 punktuose išvardytais metodais, atsižvelgiant į duomenų rinkinio turinį, turi būti taikomi nuasmeninimo ar pseudoniminimo metodai.

22. Norint skelbti duomenis netaikant nuasmeninimo ir pseudoniminimo metodų, detaliau nei rekomenduojamos pirmiau nurodytos vertės, reikėtų suderinti skelbimą su konfidencialumo specialistu ir (ar) Konfidencialių duomenų valdymo komisija.

VIII SKYRIUS
NUASMENINIMO IR (AR) PSEUDONIMINIMO METODŲ PARINKIMAS IR TAIKYMAS

23. Nuasmeninimo ir (ar) pseudoniminimo metodų parinkimas priklauso nuo duomenų atskleidimo kontrolės poreikio, duomenų rinkinio struktūros ir kintamųjų tipo. Renkantis nuasmeninimo ir (ar) pseudoniminimo metodus, turi būti atsižvelgiama į metodo taikymo įtaką duomenims, t. y. ar duomenys, pritaikius pasirinktą metodą, atitiks numatomą jų taikymo paskirtį.

24. Taikomi nuasmeninimo metodai yra:

24.1. agregavimas (angl. *aggregation*). Agregavimo metu duomenų rinkinys yra agreguojamas į lenteles;

24.2. apibendrinimas (angl. *generalization*). Naudojant apibendrinimą duomenų objektų požymiai apibendrinami kiek pakeičiant atitinkamą mastelį arba dydžio eilę (pvz., informaciją pateikiant ne miesto, o regiono mastu, pajamas pateikiant intervalais ir pan.). Apibendrinimo metodams priskiriami globalaus perkodavimo, viršaus ir apačios perkodavimo, apvalinimo metodai;

24.3. perstatymo (angl. *permutation*) metodas – duomenų rinkinio vertės sukeičiamos vietomis taip, kad kai kurios iš jų būtų dirbtinai susietos su kitais duomenų objektais. Tokiu

sukeitimu užtikrinama, kad verčių intervalas ir paskirstymas išliktų tokie patys, o verčių ir duomenų objektų ryšiai pasikeistų. Perstatymo metodams priskiriami PRAM (angl. *post-randomization method*) ir įrašų keitimo metodai;

24.4. reikšmių slėpimas (angl. *suppresion*) – metodas, kurio metu yra paslepiama duomenų dalis. Slėpti galima tiek visą duomenų stulpelį, t. y. kintamąjį, tiek tam tikras kintamųjų reikšmes;

24.5. triukšmo įterpimo (angl. *noise addition*) metodas – duomenų įrašai modifikuojami pridedant (arba padauginant, atimant ar pan.) atsitiktines vertes – triukšmą.

25. Detalesni nuasmeninimo metodų aprašymai, naudojimo pavyzdžiai pateikiami Statistinio atskleidimo kontrolės vadove, patvirtintame Lietuvos statistikos departamento generalinio direktoriaus 2022 m. balandžio 26 d. įsakymu Nr. DĮ-107 „Dėl Statistinio atskleidimo kontrolės vadovo patvirtinimo“.

26. Taikomi pseudoniminimo metodai yra:

26.1. šifravimas naudojant slaptą raktą (angl. *encryption with secret key*). Duomenų rinkinyje esantis požymis, galintis atskleisti duomenų objektą, naudojant šifravimą su slaptu raktu yra pakeičiamas. Šiuo atveju raktą turintis asmuo gali nesunkiai atkurti duomenų ryšį su konkrečiu duomenų objektu dešifravęs duomenų rinkinį, nes šifravimas yra grįžtamasis;

26.2. maišos funkcija (angl. *hash function*) – tai funkcija, kuri iš bet kokio dydžio įvesties duomenų (tai gali būti vienas požymis arba požymių rinkinys) parengia nustatyto dydžio išvesties duomenis ir kurios negalima atlikti priešinga kryptimi (t. y. grąžinti pradinę reikšmę ar reikšmes). Tai reiškia, kad nebelieka pakartotinio duomenų ryšio su duomenų objektu nustatymo rizikos, būdingos šifravimui. Tačiau, jeigu yra žinoma maišos funkcijos įvesties verčių aibė, šioms vertėms galima pakartotinai pritaikyti maišos funkciją ir taip gauti tikrąją tam tikro įrašo vertę. Šią riziką galima sumažinti pasinaudojus papildomais įrankiais:

26.2.1. pasinaudojus vadinamąja „druska“: prie duomenų įvesties, kuriai taikoma maišos funkcija, pridedama atsitiktinė vertė, vadinama „druska“;

26.2.2. pasinaudojus saugomu raktu ir maišos funkcija (angl. *keyed-hash function with stored key*) – kartu su maišos funkcija naudojamas papildomas įvesties elementas – slaptas raktas (ši funkcija nuo „druskos“ naudojimu pagrįstos funkcijos skiriasi tuo, kad „druska“ paprastai nėra slaptas elementas ir yra generuojama kiekvienai eilutei atsitiktinai).

27. Gali būti taikomi kiti Aprašo 24, 26 punktuose nenurodyti nuasmeninimo ir (ar) pseudoniminimo metodai, jei metodo taikymui pritaria agentūros Konfidencialių duomenų valdymo komisija, nustačiusi, kad šis metodas leidžia pasiekti mažesnę atskleidimo riziką arba didesnį duomenų naudingumą nei Aprašo 24, 26 punktuose nurodyti metodai.

IX SKYRIUS
DUOMENŲ NAUDINGUMO IR INFORMACIJOS PRARADIMO VERTINIMAS

28. Pritaikius pasirinktą nuasmeninimo ar pseudoniminimo metodą, turi būti įvertinamas duomenų naudingumas (angl. *utility*) ir informacijos praradimas, lyginant pradinį duomenų rinkinį su pakeistu rinkiniu, gautu atlikus duomenų nuasmeninimą ar pseudoniminimą. Didėjant duomenų naudingumui, informacijos praradimas mažėja ir atvirkščiai. Paprastai pakanka įvertinti vieną matą iš dviejų.

29. Duomenų naudingumo ir informacijos praradimo matai yra skirstomi į dvi grupes: skirtus kategoriniams kintamiesiems ir skirtus kiekybiniams kintamiesiems. Pateiktų duomenų naudingumo ir informacijos praradimo matų aprašymai, jų naudojimo pavyzdžiai plačiau aprašyti Statistinio atskleidimo kontrolės vadove, patvirtintame Lietuvos statistikos generalinio direktoriaus 2022 m. balandžio 26 d. įsakymu Nr. DĮ-107 „Dėl Statistinio atskleidimo kontrolės vadovo patvirtinimo“.

30. Kategoriniams kintamiesiems skirti duomenų naudingumo ir informacijos praradimo matai yra:

30.1. trūkstamų reikšmių skaičius;

30.2. pakeistų kintamojo reikšmių skaičius;

30.3. kategorinių kintamųjų reikšmių dažnių lentelių palyginimai, ryšių lentelių palyginimas.

31. Kiekybiniams kintamiesiems skirti duomenų naudingumo ir informacijos praradimo matai yra:

31.1. kiekybinių kintamųjų statistikų (vidurkių, medianų, koreliacijų ir pan.) palyginimai;

31.2. IL1s informacijos praradimo matas;

31.3. pradinio ir pakeisto duomenų rinkinio tikrinių reikšmių palyginimas.

X SKYRIUS
PAKARTOTINIS DUOMENŲ ATSKLEIDIMO RIZIKOS VERTINIMAS

32. Pritaikius nuasmeninimo ir (ar) pseudoniminimo metodus, pakartotinai įvertinama atskleidimo rizika taikant pirmo rizikos vertinimo metu pasirinktas rizikos vertinimo priemones. Jei rizika nėra priimtino lygio, kartojamas etapas, kuriame parenkami ir taikomi nuasmeninimo ir (ar) pseudoniminimo metodai, taikant skirtingus metodus ir (arba) parametrus.

XI SKYRIUS
SKLAIDA, DUOMENŲ PERDAVIMAS LEIDIMO TURĖTOJUI ARBA DUOMENŲ NAUDOTOJUI

33. Parengus konfidencialumo reikalavimus atitinkančių duomenų rinkinį, paskutinis atskleidimo kontrolės proceso etapas yra šio duomenų rinkinio atvėrimas / viešinimas, teikimas mokslo tikslais, duomenų perdavimas duomenų naudotojui arba, sveikatos duomenų atveju, duomenų rinkinio perdavimas arba prieigos prie duomenų suteikimas leidimo turėtojui.

XII SKYRIUS
PRIEDAI

34. Jautrių duomenų sąrašas.

35. Rizikų matrica.

XIV SKYRIUS
BAIGIAMOSIOS NUOSTATOS

36. Pasikeitus Apraše nurodytiems teisės aktams, taikomos galiojančios šių teisės aktų redakcijų nuostatos.

_____

Duomenų nuasmeninimo ir pseudoniminimo tvarkos aprašo
1 priedas

JAUTRIŲ DUOMENŲ SĄRAŠAS

1. **Su fiziniais asmenimis susiję duomenys:**
   1.1. Asmens vardas, pavardė;
   1.2. adresas (gyvenamosios vietos, veiklos vykdymo vietos, turto buvimo vietos ir pan.);
   1.3. telefono numeris;
   1.4. elektroninio pašto adresas;
   1.5. asmens kodas;
   1.6. socialinio draudimo numeris;
   1.7. elektroninės sveikatos istorijos (ESI) identifikacinis numeris;
   1.8. asmeniui išduoto pažymėjimo, leidimo, licencijos ar kito dokumento numeris;
   1.9. asmens naudojamo įrenginio IP adresas;
   1.10. asmens valdomos transporto priemonės registracijos dokumento numeris, valstybinis registracijos numeris, identifikavimo numeris (VIN);
   1.11. asmens valdomo registruotino turto registracijos numeris;
   1.12. mikroschema paženklinto gyvūno augintinio mikroschemos numeris;
   1.13. pajamos, išlaidos, kiti finansinio pobūdžio duomenys;
   1.14. su vaikais susiję duomenys;
   1.15. duomenys apie nusikalstamas veikas, apkaltinamuosius nuosprendžius, kitus teisės pažeidimus.
2. **Su juridiniais asmenimis, kitomis organizacijomis, jų padaliniais susiję duomenys:**
   2.1. pajamos, išlaidos, kiti finansinio pobūdžio duomenys.
3. **Kiti jautrūs duomenys:**
   3.1. kilnojamosios kultūros vertybės adresas;
   3.2. saugomų laukinių gyvūnų, augalų ar grybų rūšių augaviečių ar radaviečių adresai.

Duomenų nuasmeninimo ir pseudoniminimo tvarkos aprašo
2 priedas

RIZIKŲ MATRICA

Poveikis	Asmenims, ūkių subjektams	Bendruomenei, valstybei, VDA reputacijai	Aplinkai, teritorijai	Turtui					
	5 – Labai didelis				Vidutinis	Vidutinis	Aukštas	Labai aukštas	Labai aukštas
	4 – Didelis				Vidutinis	Vidutinis	Aukštas	Aukštas	Labai aukštas
	3 – Vidutinis				Žemas	Žemas	Vidutinis	Aukštas	Aukštas
	2 – Mažas				Žemas	Žemas	Vidutinis	Vidutinis	Vidutinis
	1 – Labai mažas / nėra poveikio				Žemas	Žemas	Žemas	Vidutinis	Vidutinis
					1 – Labai maža	2 – Maža	3 – Vidutinė	4 – Didelė	5 – Labai didelė
					Tikimybė				

Tikimybė	Poveikis
5 – Labai didelė – Labai realu, kad įvykis įvyks Rizikos tikimybė identifikuoti asmenį / ūkio subjektą / kitą objektą yra labai didelė (daugiau nei 90 proc.). Yra kitų viešų panašaus pobūdžio duomenų rinkinių, su kuriais yra lengva sujungti planuojamą atverti / skelbti duomenų rinkinį. *Mikroduomenims:* • Planuojama atverti detaliausiu lygmeniu • Duomenyse yra unikalus asmens / ūkio subjekto / kito objekto kodas • Nėra atlikta jokių pakeitimų *Dažnių lentelėms:* • Yra tokių celių, kuriose yra vienas asmuo / ūkio subjektas / kitas subjektas • Verslo statistika: yra tokių celių, kur ūkio subjektas sudaro daugiau nei 90 proc. celėje esančios reikšmės	**5 – Labai didelis** Labai didelis neigiamas poveikis, turintis labai reikšmingos ir ilgalaikės neigiamos įtakos asmens / ūkio subjekto / kito objekto gyvenimui, veiklai, interesams, reputacijai ir pan.
4 – Didelė – Pakankamai realu, kad įvykis įvyks Rizikos tikimybė identifikuoti asmenį / ūkio subjektą / kitą objektą yra didelė (60–89 proc.).	**4 – Didelis**

Yra kitų viešų panašaus pobūdžio duomenų rinkinių. *Mikroduomenims:* • Duomenyse yra unikalus pseudoniminis asmens / ūkio subjekto / kito objekto kodas / identifikatorius • Įrašai, pateikti žemiausiu lygmeniu, grupuojami (jungiami) pagal aukštesnius klasifikatoriaus / kategorijų rinkinio lygmenis taip, kad bet kurioje grupėje, sudarytoje iš kategorijų rinkinių / klasifikatorių beveik nelieka vienetų • Yra tam tikrų įrašų su išsiskiriančiomis reikšmėmis *Dažnių lentelėms:* • Yra tokių celių, kuriose yra 2–5 asmenys / ūkio subjektai / kiti objektai	Didelis neigiamas poveikis, galintis turėti reikšmingos ir ilgalaikės neigiamos įtakos asmens / ūkio subjekto / objekto gyvenimui, veiklai, interesams, reputacijai ir pan.
3 – Vidutinė – Įvykis gali įvykti Rizikos tikimybė identifikuoti asmenį / ūkio subjektą / kitą objektą yra vidutinė (30–59 proc.). *Mikroduomenims:* • Duomenyse yra unikalus pseudoniminis asmens / ūkio subjekto / kito objekto kodas / identifikatorius • Įrašai, pateikti žemiausiu lygmeniu, grupuojami (jungiami) pagal aukštesnius klasifikatoriaus / kategorijų rinkinio lygmenis taip, kad bet kurioje grupėje, sudarytoje iš kategorijų rinkinių / klasifikatorių beveik nelieka vienetų • Nėra tam tikrų įrašų su išsiskiriančiomis reikšmėmis arba tos reikšmės yra paslėptos naudojant tam tikrus metodus *Dažnių lentelėms:* • Yra tokių celių, kuriose yra 2–5 asmenys / ūkio subjektai / kiti objektai	**3 – Vidutinis** Neigiamas poveikis, galintis turėti pastebimos įtakos asmens / ūkio subjekto / kito objekto gyvenimui, veiklai, interesams, reputacijai ir pan.
2 – Maža – Įvykis nelabai tikėtinas, bet įmanomas Rizikos tikimybė identifikuoti asmenį / ūkio subjektą / kitą objektą yra maža (5–29 proc.). *Mikroduomenims:* • Duomenyse unikalus asmens / ūkio subjekto / kito objekto kodas pakeistas unikaliu eilutės kodu • Nėra tam tikrų įrašų su išsiskiriančiomis reikšmėmis arba tos reikšmės yra paslėptos naudojant tam tikrus metodus *Dažnių lentelėms:* • Populiacijos aprėptis, imties dydis yra nežinomas • Yra tokių celių, kuriose yra 2–5 asmenys / ūkio subjektai / kiti objektai	**2 – Mažas** Trumpalaikis neigiamas poveikis, neturintis labai reikšmingos ir pastebimos įtakos asmens / ūkio subjekto / kito objekto gyvenimui, veiklai, interesams, reputacijai ir pan.
1 – Labai maža – Įvykis labai mažai tikėtinas Rizikos tikimybė identifikuoti asmenį / ūkio subjektą / kitą objektą yra labai maža (mažesnė negu 5 proc.).	**1 – Labai mažas**

Mikroduomenys nėra skelbiami. *Dažnių lentelėms:* • Duomenys smarkiai sustambinti, nėra vienetinių reikšmių, celėse yra didelės asmenų / ūkio subjektų / kitų objektų grupės	Iš esmės jokio poveikio asmenų / ūkio subjektų / kitų objektų gyvenimui, veiklai, interesams, reputacijai nėra.