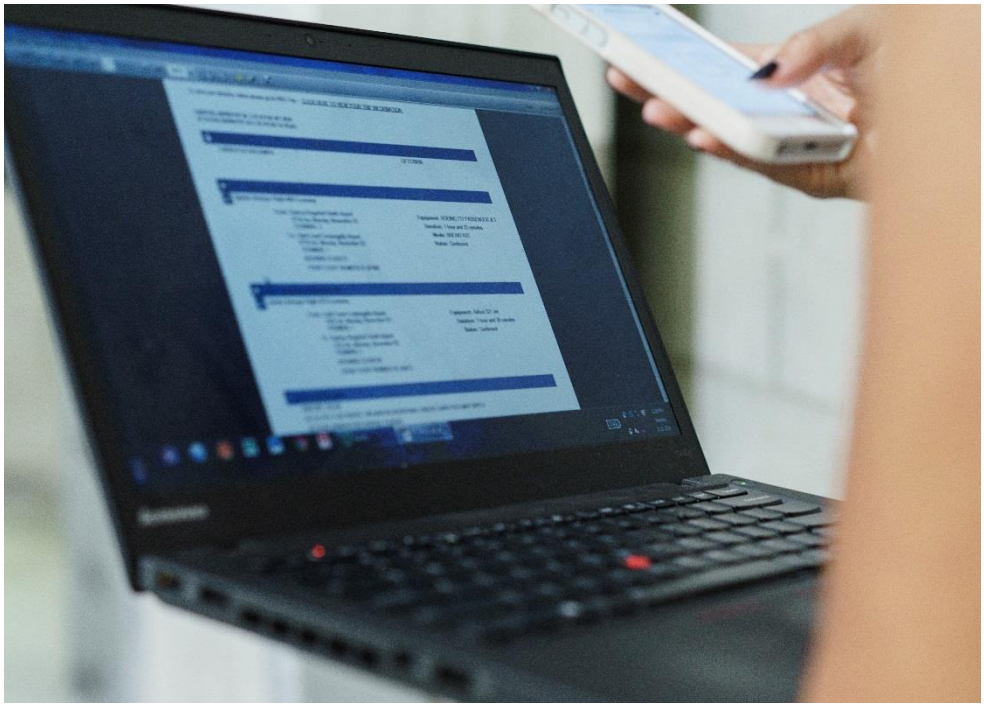
www.pwc.com

Informacinės visuomenės plėtros komitetas prie Susisiekimo ministerijos

Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugos

Duomenų nuasmeninimo ir kitos duomenų naudojimą ribojančios informacijos pašalinimo rekomendacijos ir įgyvendinimo priemonių modelis (R10)

2018 m. rugpjūčio 31 d.
Vilnius



pwc

Dokumento pakeitimai

Lentelė Nr. 1. Dokumento pakeitimų valdymas

Data	Aprašymas	Redaguotos dalys
2018-08-31	Pateiktas pirminė dokumento versija	-

Informacinės visuomenės plėtros komitetui
prie Susisiekimo ministerijos
Gedimino pr. 7,
LT-01103 Vilnius

2018 m. rugpjūčio 31 d.

Projekto „Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugos“ duomenų nuasmeninimo ir kitos duomenų naudojimą ribojančios informacijos pašalinimo rekomendacijos ir įgyvendinimo priemonių modelis

Gerbiamieji,

Šią ataskaitą parengė UAB „PricewaterhouseCoopers“, (toliau – Paslaugų teikėjas arba PwC), remiantis 2017 m. gruodžio 11 d. sudaryta pirkimo sutartimi Nr. 6F-52 (toliau – Sutartis), sudarytą su Informacinės visuomenės plėtros komitetu prie Susisiekimo ministerijos (toliau – Perkančioji organizacija, IVPK).

Darbo apimtis
Duomenų nuasmeninimo ir kitos duomenų naudojimą ribojančios informacijos pašalinimo rekomendacijos ir įgyvendinimo priemonių modelis (toliau – Ataskaita) apimtis buvo apibrėžta pirkimo „Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugos“ techninėje specifikacijoje (toliau – Techninė specifikacija). Ataskaitoje pateikiama informacija yra susijusi su atvirų duomenų naudotojų tikslinėmis grupėmis, jų poreikiais, duomenų rinkinių prioritetais, preliminariu atvertinų duomenų rinkinių sąrašu bei kitomis sritimis, susijusiomis su Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugomis (toliau – Projektas, Paslaugos) ir Techninėje specifikacijoje nurodytais reikalavimais.

Darbo metodika
Ataskaita parengta remiantis Techninėje specifikacijoje nurodytais dokumentais bei reikalavimais, PwC pateiktu techniniu pasiūlymu ir susitikimų metu gauta medžiaga. PwC nesiėmė priemonių šaltinių patikimumui nustatyti ir papildomai netikrino pateiktos informacijos. Todėl PwC nesuteikia tiesioginių ar numanomų garantijų kitiems asmenims (išskyrus IVPK pagal Sutartį) dėl Ataskaitos tikslumo ar išsamumo.
Darbus pagal Sutartį atlikome iki 2018 m. rugpjūčio 24 d. Ataskaitoje neaptariami vėlesnių įvykių ar aplinkybių padariniai, taip pat vėliau paaiškėjusi informacija. Negalime nustatyti, kokią įtaką šie darbai arba tyrimai būtų turėję Ataskaitos parengimui.
Atkreipiame Jūsų dėmesį į Ataskaitoje išdėstytas svarbias mūsų pateikiamas prielaidas. Paslaugų teikėjas neatsako kitiems asmenims (išskyrus IVPK pagal Sutartį) už Ataskaitos rengimą. Be to, Paslaugų teikėjas neprisiima sutartinės, deliktinės ir kitokios atsakomybės (neatsižvelgiant į ieškinio formą) taikytinos teisės leistina apimtimi ir neprisiima atsakomybės už kitų asmenų patirtus padarinius (išskyrus IVPK Sutarties pagrindu) ar kitus pagal šią Ataskaitą priimtus sprendimus ar atsisakymą juos priimti**.**

Dokumento platinimas ir atsakomybė
Asmuo, susipažinęs ir perskaitęs šią Ataskaitą, sutinka ir įsipareigoja laikytis šių sąlygų:

- Ataskaitos skaitytojas supranta, kad Ataskaita pagrįsta ekspertiniu vertinimu;
- Ataskaitos skaitytojas supranta, kad Paslaugų teikėjo darbai yra atlikti pagal mūsų kliento – IVPK – nurodymus ir skirti tik mūsų klientui;
- Ataskaitos skaitytojas pripažįsta, kad Ataskaita parengta mūsų kliento – IVPK – nurodymu, todėl į ją gali būti įtraukta ne viskas, kas svarbu skaitytojui;
- Ataskaitos skaitytojas sutinka, kad PwC, jos partneriai, įgaliotojai, darbuotojai ir atstovai neprisiima sutartinės ar deliktinės atsakomybės (įskaitant aplaidumą ir teisės aktuose

nustatytų įsipareigojimų pažeidimą ir kt.) ir neatsako už žalą, išlaidas ar nuostolius, patirtus skaitytojui vadovaujantis šiuo dokumentu ar su juo susipažinus.

Bendroji informacija
Jeigu kiltų klausimų dėl šios Ataskaitos turinio, prašome kreiptis į Justą Urboną telefonu +37065511056 arba el. paštu justas.urbonas@pwc.com.

Pagarbiai

Audrius Leipus
UAB „PricewaterhouseCoopers“
Konsultacijų skyriaus vyr. projektų vadovas

Turinys

Sąvokos ir sutrumpinimai

Sąvoka	Paaiškinimas
ADP	Atvirų duomenų portalas
Ataskaita	Duomenų nuasmeninimo ir kitos duomenų naudojimą ribojančios informacijos pašalinimo rekomendacijos ir įgyvendinimo priemonių modelis (R10)
IT	Informacinės technologijos
IVPK, Perkančioji organizacija	Informacinės visuomenės plėtros komitetas prie Susisiekimo ministerijos
Kvaziidentifikatoriai	Požymiai, susiję su duomenų subjektu arba duomenų subjektų grupe
Projektas, Paslaugos	Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugos
PwC, Paslaugų teikėjas	UAB „PricewaterhouseCoopers“
Sutartis	Informacinės visuomenės plėtros komiteto prie Susisiekimo ministerijos ir UAB „PricewaterhouseCoopers“ 2017 m. gruodžio 11 d. sudaryta pirkimo sutartis Nr. 6F-52
Techninė specifikacija	Lietuvos atvirų duomenų formavimo metodologinių ir teisinio reglamentavimo priemonių įgyvendinimui bei tam reikalingų valstybės institucijų darbuotojų kompetencijų ugdymo paslaugų techninė specifikacija

1 Daliniai apribojimai

Daliniai apribojimai – tai duomenų audito metu nustatyti apribojimai, kuriuos pašalinus duomenų rinkiniai galėtų būti atveriami, t. y. nepatektų į kategoriją „Nėra atviri duomenys“. Pavyzdžiui, tam tikra liga sergančių pacientų sąrašas negalėtų būti atveriamas, tačiau pašalinus asmens duomenis (vardą, pavardę, kt.), ir paliekant kitus požymius, pvz., lytis, amžiaus grupė, apskritis ir kt., toks duomenų rinkinys galėtų būti atveriamas.

Duomenų valdytojai turi įvertinti tris labai svarbius nuasmeninimo požiūriu rizikos veiksnius:

- **išskyrimo galimybę**, t. y. galimybę išskirti kai kuriuos arba visus įrašus, pagal kuriuos būtų galima nustatyti į duomenų rinkinį įtraukto asmens tapatybę;

- **susiejimo galimybę**, t. y. galimybę susieti bent du įrašus, susijusius su tuo pačiu duomenų subjektu arba ta pačia duomenų subjektų grupe (toje pačioje duomenų bazėje arba dviejose skirtingose duomenų bazėse). Jeigu išpuolio vykdytojas gali nustatyti (pvz., atlikdamas koreliavimo analizę), kad du įrašai priskirti tai pačiai asmenų grupei, tačiau negali iš tos grupės išskirti pavienių asmenų, tai šiuo metodu apsaugoma nuo išskyrimo, bet neužtikrinama apsauga nuo susiejimo;

- **išvados padarymo galimybę**, t. y. galimybę dedukcijos būdu gana tikėtinai nustatyti požymio vertę remiantis kitų požymių rinkinio vertėmis.

Visų pirma, turi būti nustatyti duomenys, kurie negali būti atverti, t. y. priklauso kategorijai „Nėra atviri duomenys“ ir kokie veiksmai turi būti atlikti rengiant duomenų rinkinį. Tai atliekama peržiūrint kiekvieną duomenų bazės lauką, jį priskiriant vienai iš kategorijų ir nustatant veiksmą:

a) Pakeitimai nereikalingi;
b) Pašalinimas;
c) Randomizavimas;
d) Apibendrinimas;
e) Pseudonimų suteikimas;
f) Autentifikacija;
g) Kodavimas;
h) Loginimas.

1.1 Duomenys, kurie negali būti atverti

Toliau esančioje lentelėje pateikiami duomenys, kuriems yra taikomi apribojimai dėl duomenų atvėrimo.

Lentelė Nr. 2. Duomenys, kurie negali būti atverti

Duomenys	Aprašymas	Taikytinos pašalinimo priemonės
Asmens duomenys	Asmens duomenys – bet kuri informacija, susijusi su fiziniu asmeniu – duomenų subjektu, kurio tapatybė yra žinoma arba gali būti tiesiogiai ar netiesiogiai nustatyta pasinaudojant tokiais duomenimis kaip asmens kodas, vienas arba keli asmeniui būdingi fizinio, fiziologinio, psichologinio, ekonominio, kultūrinio ar socialinio pobūdžio požymiai. Asmens duomenų teisinė apsauga yra reguliuojama bendrojo duomenų apsaugos reglamento ir yra reikšmingas apribojimas duomenų atvėrimui Už atvirų duomenų rinkinių nuasmeninimą yra atsakingi asmens duomenų valdytojai, todėl jie prieš nuspręsdami dėl duomenų atvėrimo apimties turi įvertinti asmens duomenų atvėrimo poveikį.	1. Asmens duomenų laukų naikinimas – duomenys, kurie turi požymį „Asmens duomenys“ turi būti ištrinami. 2. Konkrečių įrašų naikinimas – įrašai, kurie be asmens duomenų leidžia identifikuoti asmenį turi būti naikinami pvz. automobiliai „Ferrari“, ligos, kuria

Duomenys	Aprašymas	Taikytinos pašalinimo priemonės
	Asmens duomenų nuasmeninimo poveikio vertinimas turi būti periodinis procesas, kuris turi būti vykdomas kasmet.	serga tik 10 žmonių Lietuvoje, atvejis.
Intelektinė nuosavybė ir patentai	Išskirtinos dvi Intelektinės nuosavybės objektų grupės: • **Pramoninės nuosavybės objektai (patentai, prekių ženklai, dizainai ar topografijos) –** intelektinės nuosavybės teisės į šiuos objektus yra registruojamos – intelektinei nuosavybei atsirasti ir jos apsaugai įtvirtinti reikalingi papildomi registravimo veiksmai Valstybiniame patentų biure ar kitoje intelektinę nuosavybę registruojančioje institucijoje, todėl apsauga jų turiniui atsiranda ir taikoma tik nuo oficialios jų registracijos. • **Autorių ir gretutinių teisių objektai –** intelektinės veiklos rezultatai, į kuriuos teisės kūrėjui atsiranda nuo jų sukūrimo momento, jeigu jie atitinka visus saugotinam objektui keliamus reikalavimus – iš jų išskirtini du saugotini autorių teisių objektai: - **Kūrinys** – originalus kūrybinės veiklos rezultatas literatūros, mokslo ar meno srityje, nepaisant jo meninės vertės, išraiškos būdo ar formos[1]. - **Duomenų bazė** – susistemintas ar metodiškai sutvarkytas kūrinių, duomenų arba kitokios medžiagos rinkinys, kuriuo galima individualiai naudotis elektroniniu ar kitu būdu, išskyrus kompiuterių programas, naudojamas tokių duomenų bazėms kurti ar valdyti[2]. Aukščiau išvardintų objektų kūrėjų (pramoninės nuosavybės savininkų, autorių ar *sui generis* teisių subjektų[3]) ir turtinių teisių į šiuos objektus turėtojų teisės yra saugomos – objektų turinį sudarančių duomenų tolimesniam (pakartotiniam) naudojimui būtinas susitarimas su jų turtinių teisių turėtoju, todėl institucijos negali šių duomenų atverti ir teikti kitiems asmenims.	
Nacionalinio saugumo duomenys	Analogiškai informacijos apsaugai civiliniuose santykiuose, valstybės Institucijų veikloje taip pat yra informacijos, kuriai yra būtina taikyti apsaugą siekiant apsaugoti valstybės interesus. Visa informacija patenkanti į Lietuvos Respublikos valstybės ir tarnybos paslapčių įstatymo sritį neturėtų būti teikiama šios teisės neturintiems asmenims ir negali būti publikuojama.	
Komerciniai duomenys	Komercinės (gamybinės) paslaptys - informacija laikoma komercine (gamybine) paslaptimi, jeigu ji turi tikrą ar potencialią komercinę (gamybinę) vertę dėl to, kad jos	

[1] Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo 2 straipsnio 29 dalis.
[2] Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo 2 straipsnio 7 dalis.
[3] S*ui generis* teisių subjektas – duomenų bazės gamintojas, kuris parinkdamas, sudarydamas, tikrindamas bei pateikdamas duomenų bazės turinį padarė esminių kokybinių ir (ar) kiekybinių (intelektinių, finansinių, organizacinių) investicijų, taip pat fizinis arba juridinis asmuo, kuriam perėjo duomenų bazių gamintojo *sui generis* teisės.

Duomenys	Aprašymas	Taikytinos pašalinimo priemonės
	nežino tretieji asmenys ir ji negali būti laisvai prieinama dėl šios informacijos savininko ar kito asmens, kuriam savininkas ją yra patikėjęs, protingų pastangų išsaugoti jos slaptumą. Informacijos, kuri yra kito ūkio subjekto komercinė paslaptis, naudojimas, perdavimas, skelbimas be šio subjekto sutikimo, taip pat tokios informacijos gavimas iš asmenų, neturinčių teisės šios informacijos perduoti, turint tikslą konkuruoti, siekiant naudos sau arba padarant žalą šiam ūkio subjektui, yra draudžiami nesąžiningos konkurencijos veiksmai[4]. Atitinkamai neteisėtai atskleidus komercinę (gamybinę) paslaptį, kaltas asmuo privalo atlyginti tokias savo veiksmais padarytus nuostolius. Dėl to nesant teisinio aiškumo Institucijoms kyla rizika dėl tokių duomenų atskleidimo galimų teisinių padarinių ir Institucijoms saugiau yra informacijos neteikti. *„Komercinė paslaptis – informacija, atitinkanti visus šiuos reikalavimus:* *a) ji yra slapta ta prasme, kad jos kaip visumos arba tikslios jos sudėties ir sudedamųjų dalių konfigūracijos apskritai nežino arba negali lengvai gauti asmenys toje aplinkoje, kurioje paprastai dirbama su tokia informacija;* *b) ji turi komercinės vertės, nes yra slapta;* *c) ji yra objektas pagrįstų veiksmų, kurių imasi teisėtai tokią informaciją valdantis asmuo, kad tam tikromis aplinkybėmis ją išlaikytų slaptą;[5]“* Komercinio konfidencialumo – tai gali būti informacija, kuriai pagal įmonių vidinius dokumentus (pavyzdžiui konfidencialios informacijos sąrašus) ar įmonių sudarytus sandorius yra nustatytas konfidencialios informacijos statusas. Konfidencialumo apsauga – asmens įpareigojimas jam žinomą konfidencialią informaciją, naudoti tik nustatytais ir apibrėžtais tikslais, užtikrinti, kad ji netaptų žinoma (nebūtų atskleista) tokios teisės neturintiems asmenims ir neatskleisti jos tretiesiems asmenims, nenaudoti minėtos konfidencialios informacijos asmeniniams arba trečiųjų šalių interesams tenkinti, išskyrus kai pareiga atskleisti tokią informaciją yra nustatyta galiojančiuose teisės aktuose.	
Kita konfidenciali informacija	Kiti specialieji įstatymai, reglamentuojantys konkrečių duomenų teikimą, gali įtvirtinti specialius apribojimus (pvz., Lietuvos Respublikos nekilnojamojo turto registro įstatymas[6] apriboja subjektų ratą, kam registro informacija ir dokumentai gali būti teikiami).	

[4] Lietuvos Respublikos konkurencijos įstatymo 15 straipsnio 1 dalies 3 punktą
[5] Direktyvos (ES) 2016/943 2 straipsnio 1 dalis.
[6] Lietuvos Respublikos nekilnojamojo turto registro įstatymo 42 straipsnis.

2 Duomenų nuasmeninimo metodai

2.1 Randomizavimo metodai

Randomizavimas – tai metodų, kuriais keičiamas duomenų tikrumas siekiant panaikinti aiškią duomenų ir asmens sąsają, grupė. Kai duomenys yra ganėtinai nekonkretūs, jų nebegalima susieti su konkrečiu asmeniu. Taikant randomizavimą, atskirų įrašų savitumas nemažėja, nes kiekvienas įrašas vis viena bus išvedamas pagal atskirą duomenų subjektą, tačiau šiuo būdu užtikrinama apsauga nuo išvestinės informacijos gavimo išpuolių ir (arba) rizikos ir jį, siekiant suteikti didesnę privatumo garantiją, galima derinti su apibendrinimu. Norint užtikrinti, kad, remiantis įrašu, nebūtų galima nustatyti pavienio asmens tapatybės, gali prireikti papildomų metodų.

Toliau esančioje lentelėje pateikiami randomizavimo metodai.

Lentelė Nr. 3. Randomizavimo metodai

Metodas	Aprašymas	Taikymo pavyzdys
Iškraipytų duomenų įterpimas (angl. noice addition)	Iškraipytų duomenų įterpimo metodas pirmiausia naudingas tada, kai požymiai gali turėti reikšmingą neigiamą poveikį asmenims. Šio metodo esmė – į duomenų rinkinį įtrauktų požymių pakeitimas sumažinant jų tikslumą, tačiau išsaugant bendrą pasiskirstymą. Tvarkydamas duomenų rinkinį, stebėtojas manys, kad vertės yra tikslios, bet tai bus teisinga tik iš dalies. Jeigu šis metodas bus taikomas veiksmingai, trečioji šalis negalės nustatyti asmens tapatybės ir neturėtų galėti ištaisyti duomenis arba kaip nors kitaip nustatyti, kaip duomenys buvo pakeisti. Iškraipytų duomenų įterpimą paprastai reikia derinti su kitais nuasmeninimo metodais, pvz., su akivaizdžių požymių ir kvaziindentifikatorių pašalinimu. Iškraipymo laipsnis turėtų priklausyti nuo to, kokio lygio informacija yra reikalinga, ir nuo apsaugotų požymių atskleidimo daromo poveikio asmenų privatumui.	Jeigu asmens ūgis iš pradžių buvo išmatuotas centimetrų tikslumu, nuasmenintame duomenų rinkinyje ūgis gali būti nurodomas tik ±10 cm tikslumu, t. y. 174 cm turėtų būti pakeista į 170 cm.
Perstatymas (angl. *permutation*)	Taikant šį metodą, lentelėje esančių požymių vertės sukeičiamos vietomis taip, kad kai kurios iš jų būtų dirbtinai susietos su kitais duomenų subjektais. Tai naudinga, kai svarbu išsaugoti tikslų kiekvieno į duomenų rinkinį įtraukto požymio pasiskirstymą. Perstatymo metodas – tai alternatyva, kurią taikant duomenų rinkinio vertės pakeičiamos tiesiog sumaišant vietomis skirtingų įrašų vertes. Tokiu sukeitimu užtikrinama, kad verčių intervalas ir paskirstymas išliktų tokie patys, o verčių ir asmenų koreliacijos pasikeistų. Jeigu dviem arba daugiau požymių būdingas loginis tarpusavio ryšys arba statistinė koreliacija ir atliekamas nepriklausomas jų perstatymas, toks ryšys sunaikinamas. Todėl gali būti svarbu susijusių požymių rinkinio perstatymą atlikti taip, kad nebūtų pažeistas loginis tarpusavio ryšys, nes kitaip išpuolio vykdytojas galėtų nustatyti sukeistus požymius ir atlikti atvirkštinį perstatymą.	Vieno duomenų lauko įrašai yra apkeičiami vietomis, pvz., jei buvo: A-1 B-2 C-3 pritaikius perstatymo metodą: A-2 B-3 C-1

Metodas	Aprašymas	Taikymo pavyzdys
	Panašiai kaip ir iškraipytų duomenų įterpimo atveju, vien perstatymo pritaikymas gali neužtikrinti nuasmeninimo, todėl jis visada turėtų būti derinamas su akivaizdžių požymių ir (arba) kvaziidentifikatorių pašalinimu.	
Diferencinis privatumas (angl. *differential privacy*)	Diferencinis privatumas priskiriamas randomizavimo metodų grupei, tačiau jis pagrįstas kitokiu principu: iškraipytų duomenų įterpimas taikytinas prieš paskelbiant duomenų rinkinį, o diferencinio privatumo metodas gali būti taikomas, kai duomenų valdytojas parengia nuasmenintus duomenų rodinius, išsaugodamas pirminių duomenų kopiją. Tokie nuasmeninti rodiniai paprastai parengiami naudojant užklausų poaibį, skirtą tam tikrai trečiajai šaliai. Į šį poaibį vėliau sąmoningai įtraukiami atsitiktiniai iškraipyti duomenys. Taikydamas diferencinio privatumo metodą, duomenų valdytojas sužino, kiek iškraipytų duomenų jis turėtų įterpti ir kokiu pavidalu, kad užtikrintų reikiamas privatumo garantijas. Šiuo atveju labai svarbu nuolat stebėti (ne rečiau kaip kiekvienos naujos užklausos atveju), ar neatsirado galimybė nustatyti asmens tapatybę pasinaudojant užklausos rezultatų aibe. Be to, derėtų paaiškinti, kad diferencinio privatumo metodu pirminiai duomenys nepakeičiami, o kol jie išlieka, duomenų valdytojas, atsižvelgdamas į visas galimas pasitelktinas priemones, asmens tapatybę gali nustatyti pasinaudodamas diferencinio privatumo užklausų rezultatais. Šie rezultatai taip pat turėtų būti laikomi asmens duomenimis. Siekiant apriboti išvados padarymo ir susiejimo išpuolių galimybę, būtina sekti subjektų teikiamas užklausas ir stebėti apie duomenų subjektus gautą informaciją; todėl diferencinio privatumo metodu valdomos duomenų bazės neturėtų būti prieinamos viešoms paieškos sistemoms, kuriose nėra užklausas teikiančių subjektų sekimo galimybės.	Metodas taikomas, kai duomenys teikiami pagal užklausas, į juos įterpiant iškraipytus ir apytikslius duomenis. Pvz., įstaiga turi sveikatos duomenis. Duomenų naudotojas kreipiasi dėl ligos X atvejų. Įstaiga, naudodama esamus duomenis, papildo juos iškraipytais duomenimis ir pateikia naudotojui apytikslius duomenis.

2.2 *Apibendrinimo metodai*

Apibendrinimas yra antroji nuasmeninimo metodų grupė. Pagal šį principą duomenų subjektų požymiai apibendrinami arba, kitaip tariant, susilpninami, kiek pakeičiant atitinkamą mastelį arba dydžio eilę (pvz., informaciją pateikiant ne miesto, o regiono mastu, mėnesio, o ne savaitės apimtimi). Nors apibendrinimas, siekiant panaikinti išskyrimo galimybę, ir gali būti veiksmingas, ne visais atvejais šiuo principu užtikrinamas tinkamas nuasmeninimas; pirmiausia, taikant šį principą, būtina pasitelkti specialius sudėtingus kiekybinius metodus, kuriais būtų panaikinta susiejimo ir išvados padarymo galimybė. Tačiau pažymėtina, jog atviri duomenys didžiausią vertę kuria, kai yra pateikiami kaip pirminiai duomenys (angl. raw data), todėl apibendrinimą reikėtų naudoti tik tuomet, kai jis yra būtinas.

Toliau esančioje lentelėje pateikiami apibendrinimo metodai.

Lentelė Nr. 4. Apibendrinimo metodai

Metodas	Aprašymas	Taikymo pavyzdys
Agregavimas ir *k* anonimiškumas (angl. *aggregation* ir *k-anonnymity*)	Agregavimo ir *k* anonimiškumo metodais siekiama panaikinti galimybę išskirti duomenų subjektus, juos grupuojant kartu su ne mažiau kaip *k* kitų asmenų. Šiuo tikslu požymių vertės apibendrinamos tokiu mastu, kad kiekvienam asmeniui būtų priskirta tokia pat vertė. Šie metodai gali būti taikomi tada, kai dėl požymių tikslių verčių koreliacijos gali susidaryti kvaziindikatoriai.	Vietovės mastelį pastambinus nuo miesto iki šalies, bus įtraukta daugiau duomenų subjektų.
***l* įvairovė (angl. *l-diversity*)**	*l* įvairovės metodu išplečiamas k anonimiškumo metodas, siekiant užtikrinti, kad nebebūtų galima rengti determinavimo būdu pagrįstų išpuolių, pasirūpinant, kad kiekvienoje lygiavertiškumo klasėje kiekvienam požymiui būtų priskirta ne mažiau kaip *l* skirtingų verčių. Vienas iš pagrindinių siektinų tikslų – riboti lygiavertiškumo klasių, kurioms būtų būdingas menkas požymių kintamumas, susidarymą, kad bendrųjų žinių apie tam tikrą duomenų subjektą turinčiam išpuolio vykdytojui visada liktų didelių abejonių dėl savo išvadų.	
***t* tankis (angl. *t-closeness*)**	*t* tankio metodas yra patobulintas l įvairovės metodas, nes juo siekiama sudaryti lygiavertiškumo klases, kurioms būtų būdingas panašus į pirminį požymių pasiskirstymas lentelėje. Šis metodas naudingas tada, kai svarbu, kad duomenys būtų kuo panašesni į pirminius; todėl lygiavertiškumo klasei taikomas papildomas apribojimas, pagal kurį kiekvienoje lygiavertiškumo klasėje turėtų būti ne tik mažiau kaip l skirtingų verčių, bet ir kiekviena vertė turi būti pateikta tiek kartų, kiek reikalinga tam, kad būtų atkurtas pirminis kiekvieno požymio pasiskirstymas.	

2.3 Pseudonimų suteikimo metodai

Pseudonimų suteikimas – tai metodas, pagal kurį vienas požymis (paprastai – unikalus) įraše pakeičiamas kitu. Todėl išlieka galimybė netiesiogiai nustatyti fizinio asmens tapatybę; taigi vien pseudonimų suteikimas neužtikrina duomenų rinkinio anonimiškumo.

Taikant pseudonimų suteikimo metodą, sumažinama galimybė duomenų rinkinį susieti su pirmine duomenų subjekto tapatybe; taigi šis metodas yra naudinga saugumo priemonė, bet tai nėra nuasmeninimo metodas. Pseudonimų suteikimo rezultatas gali nepriklausyti nuo pirminės vertės (pvz., jeigu tai atsitiktinis duomenų valdytojo sugeneruotas skaičius arba duomenų subjekto pasirinkta pavardė) arba gali būti sukuriamas naudojantis požymio arba jų grupės pirminėmis vertėmis, pvz., taikant maišos funkciją arba šifravimo sistemą.

Toliau esančioje lentelėje pateikiami pseudonimų suteikimo metodai.

Lentelė Nr. 5. Pseudonimų suteikimo metodai.

Metodas	Aprašymas	Taikymo pavyzdys
Šifravimas naudojant slaptą raktą (angl. *encryption with secret key*)	Šiuo atveju raktą turintis asmuo gali nesunkiai atkurti kiekvieno duomenų subjekto tapatybę dešifravęs duomenų rinkinį, nes asmens duomenys, nors ir užšifruoti, tebėra duomenų rinkinyje. Jeigu buvo pritaikyta pažangi šifravimo sistema, dešifravimas galimas tik žinant raktą.	Organizacija užšifruoja duomenų rinkinį ir raktą suteikia tik tam tikram duomenų rinkinio naudotojui. Toks duomenų rinkinys nėra laisvai prieinamas visiems atvirų duomenų naudotojams.
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; tai reiškia, kad nebelieka pakartotinio tapatybės nustatymo rizikos, būdingos šifravimui. Tačiau, jeigu yra žinomas maišos funkcijos įvesties verčių intervalas, šioms vertėms galima pakartotinai pritaikyti maišos funkciją ir taip gauti teisingą tam tikro įrašo vertę. Kad būtų galima masiškai atkurti didelį verčių, kurioms buvo pritaikyta maišos funkcija, rinkinį, taip pat gali būti parengiamos iš anksto apskaičiuotų verčių lentelės. Taikant *„druskos“ naudojimu pagrįstą maišos funkciją* (angl. *salted-hash function*) (prie požymio, kuriam taikoma maišos funkcija, pridedama atsitiktinė vertė, vadinama „druska“), galima sumažinti įvesties vertės nustatymo tikimybę, tačiau pagrįstomis priemonėmis vis vien gali būti įmanoma apskaičiuoti pirminę požymio vertę, paslėptą sudėtingesnės maišos funkcijos (su „druskos“ elementu) rezultatu.	Tam tikro lauko informacija pakeičiama atsitiktiniu indeksu, pvz., naudojant SHA-256 algoritmą, el. pašto adresas *sarah_example@gmail.com* yra pakeičiamas į *22B28AB920AC727C530D 5C9ADB23C3DB7E26EC5 CC1600B0EDB60B1D6398 D1C6* Taikant „druskos“ naudojimu pagrįstą maišos funkciją, prie algoritmu suformuoto indekso pridedama papildoma atsitikinė vertė.
Saugomo rakto naudojimu pagrįsta maišos funkcija (angl. *keyed-hash function with stored key*)	Tai tam tikra maišos funkcija, kai naudojamas papildomas įvesties elementas – slaptas raktas (ši funkcija nuo „druskos“ naudojimu pagrįstos funkcijos skiriasi tuo, kad „druska“ paprastai nėra slaptas elementas). Duomenų valdytojas, naudodamas slaptą raktą, šią funkciją gali pakartotinai pritaikyti požymiui, tačiau išpuolio vykdytojui tampa gerokai sunkiau pakartoti šią funkciją nežinant rakto, nes mėgintinų variantų skaičius yra per didelis, kad būtų įmanoma tai padaryti.	
Determinavimu pagrįstas šifravimas arba	Pagal šį metodą kiekvienam duomenų rinkinyje esančiam požymiui kaip pseudonimas gali būti parenkamas atsitiktinis	

Metodas	Aprašymas	Taikymo pavyzdys
panaikinamo rakto naudojimu pagrįsta maišos funkcija (angl. *deterministic encryption or keyed-hash function with deletion of the key*)	skaičius, o tada panaikinama atitikties lentelė. Pasitelkus tokį sprendimą, galima sumažinti galimybę duomenų rinkinyje esančius asmens duomenis susieti su kitame duomenų rinkinyje, kuriame naudojamas kitoks pseudonimas, esančiais duomenimis apie tą patį asmenį. Skaičiavimo požiūriu išpuolio vykdytojui, pasitelkusiam net ir pažangų algoritmą, būtų sunku iššifruoti arba pakartoti funkciją, nes, neturint rakto, reikėtų išmėginti kiekvieną galimą raktą.	
Pakaitinių simbolių naudojimas (angl. *tokenization*)	Šis metodas paprastai taikomas (nors gali būti taikomas ir kitur) finansų sektoriuje, siekiant kortelių atpažinimo numerius (angl. ID) pakeisti vertėmis, kurios išpuolio vykdytojui būtų ne tokios naudingos. Šis metodas sukurtas remiantis pirmiau aptartais metodais ir paprastai grindžiamas vienakrypčių šifravimo priemonių taikymu arba eilės numerio ar atsitiktine tvarka sugeneruoto numerio, kuris nėra matematiškai gaunamas iš pirminių duomenų, priskyrimu pasitelkiant indeksavimo funkciją.	

2.4 Duomenų apribojimų šalinimo metodų privalumai ir trūkumai

Toliau lentelėje apibendrinami duomenų apribojimų šalinimo metodų privalumai ir trūkumai.

Lentelė Nr. 6. Duomenų apribojimų šalinimo metodų privalumai ir trūkumai

	Ar išlieka išskyrimo rizika?	Ar išlieka susiejimo rizika?	Ar išlieka išvados padarymo rizika?
Pseudonimų suteikimas	Taip	Taip	Taip
Iškraipytų duomenų įterpimas	Taip	Ne (laikantis tam tikrų sąlygų)	Ne (laikantis tam tikrų sąlygų)
Pakeitimas	Taip	Taip	Ne (laikantis tam tikrų sąlygų)
Agregavimas arba *k* anonimiškumas	Ne	Taip	Taip
***l* įvairovė**	Ne	Taip	Ne (laikantis tam tikrų sąlygų)
Diferencinis privatumas	Ne (laikantis tam tikrų sąlygų)	Ne (laikantis tam tikrų sąlygų)	Ne (laikantis tam tikrų sąlygų)
Maiša arba pakaitinių simbolių naudojimas	Taip	Taip	Ne (laikantis tam tikrų sąlygų)