

VILNIAUS UNIVERSITETAS  
EKONOMIKOS FAKULTETAS  
EKONOMINĖS INFORMATIKOS KATEDRA

Karolis Laucė

*Strateginio informacinių sistemų valdymo magistro programa*

**MAGISTRO BAIGIAMASIS DARBAS**

**Duomenų atvėrimo poreikio nustatymas Lietuvos viešajam ir  
privačiajam sektoriui**

**Determining the Need for Data Opening for the Lithuanian Public and  
Private Sectors**

Leidžiama ginti \_\_\_\_\_

(parašas)

Katedros vedėjas prof. dr. **R. Skyrius**

Magistrantas \_\_\_\_\_

(parašas)

Darbo vadovas \_\_\_\_\_

(parašas)

Asist. dr. **M. Kazimianec**

Darbo įteikimo data:

Registracijos Nr.

**Vilnius, 2020**

## TURINYS

TURINYS.....	2
LENTELIŲ SĄRAŠAS.....	4
SANTRAUKA .....	5
ĮVADAS.....	6
1. ATVIRAI PRIEINAMŲ DUOMENŲ TEORINIAI ASPEKTAI .....	9
1.1 Atvirų duomenų samprata ir raida .....	9
1.2 Atvirai prieinamų duomenų istorija.....	12
1.3 Atvirai prieinamų duomenų rūšys ir populiarumas .....	13
1.4 Atvirai prieinamų vyriausybinių duomenų gyvavimo ciklas.....	16
1.5 Susietieji duomenys ir atvirai prieinami duomenys.....	18
1.6 Atvirai prieinamų duomenų skatinimo valstybinės programos .....	21
1.7 Atvirai prieinamų duomenų teikiama nauda bei ypatumai .....	23
2. ATVIRAI PRIEINAMŲ DUOMENŲ PANAUDOJIMO ATVEJŲ TYRIMAS.....	30
2.1 Tyrimo metodai ir darbo eiga .....	30
3. ATVIRAI PRIEINAMŲ DUOMENŲ RINKINIŲ KLASIFIKAVIMAS .....	31
3.1 Atvirai prieinamų duomenų rūšys.....	31
3.2 Atvirai prieinami duomenys ir jų rūšys Lietuvoje .....	33
3.3 Duomenų panaudojimo atvejų pasaulyje klasifikavimas.....	40
3.4 Lietuvos atvirų duomenų panaudojimo atvejai.....	49
3.5 Asociacijų paieškos tyrimas.....	57
IŠVADOS IR REKOMENDACIJOS .....	66
LITERATŪRA .....	68
Summary.....	72
PRIEDAI .....	74

## PAVEIKSLŲ SĄRAŠAS

1 pav. Duomenų rūšys .....	9
2 pav. Atvirai prieinamų duomenų rūšys.....	14
3 pav. Populiariausi duomenų rinkiniai.....	16
4 pav. Atvirai prieinamų vyriausybinių duomenų gyvenimo ciklas.....	17
5 pav. Susietųjų duomenų tinklas .....	20
6 pav. Ryšys tarp atvirai prieinamų, vyriausybinių ir susietųjų duomenų .....	21
7 pav. Duomenų vertės kūrimo grandinė .....	24
8 pav. Prognozuojamas atvirai prieinamų duomenų rinkos dydis Europos sąjungoje .....	26
9 pav. Prognozuojamas darbo vietų skaičius darbui su atvirai prieinamais duomenimis Europos sąjungoje.....	27
10 pav. 2020 metų prognozuojamas viešojo sektoriaus išlaidų sumažėjimas įtakotas atvirai prieinamų duomenų paskelbimo.....	28
11 pav. Atvirų duomenų teikiama nauda.....	29
12 pav. „Creditinfo“ duomenų gavimo ir dalijimosi grafinis atvaizdavimas .....	52
13 pav. SPSS Modeler asociacijų paieškos funkcijų žemėlapis.....	59

## LENTELIŲ SĄRAŠAS

1 lent. Valstybių priimtos atvirų duomenų skatinimo programos .....	22
2 lent. ODINE programos šalių reitingas iš ODINE 2017 m. ataskaitos .....	29
3 lent. Atvirų duomenų rinkinių rūšiavimas .....	32
4 lent. Lietuvos atvirų duomenų rinkmenų sąrašas .....	34
5 lent. Lietuvos atvirų duomenų, kurių nėra <a href="http://www.opendata.gov.lt">www.opendata.gov.lt</a> .....	38
6 lent. Atvirų duomenų panaudojimas žemės ūkio srityje.....	40
7 lent. Atvirų duomenų panaudojimas kultūros ir turizmo srityje.....	41
8 lent. Atvirų duomenų panaudojimas verslo srityje.....	41
9 lent. Atvirų duomenų panaudojimas švietimo srityje.....	42
10 lent. Atvirų duomenų panaudojimas energijos ir klimato srityje.....	43
11 lent. Atvirų duomenų panaudojimas finansų ir draudimo srityje.....	44
12 lent. Atvirų duomenų panaudojimas valstybės valdymo srityje.....	45
13 lent. Atvirų duomenų panaudojimas sveikatos srityje.....	46
14 lent. Atvirų duomenų panaudojimas statybų ir NT srityje.....	47
15 lent. Atvirų duomenų panaudojimas juos atvaizduojant žemėlapyje.....	47
16 lent. Atvirų duomenų panaudojimas žiniasklaidos ir ryšių srityje.....	48
17 lent. Atvirų duomenų panaudojimas transporto srityje.....	49
18 lent. Projekto „Kur gyvenu?“ naudojami atviri duomenys.....	50
19 lent. Kredito biuro „Creditinfo“ naudojami duomenys.....	53
20 lent. Projekto „Credit reform“ naudojami duomenys.....	56
21 lent. Ne Lietuvoje dažniausiai naudotos duomenų rūšys pagal aukščiausią „Paramos rodiklį“ .	59
22 lent. Ne Lietuvoje naudotos duomenų rūšys pagal aukščiausią „Paramos rodiklį“ .....	60
23 lent. Ne Lietuvoje naudotos duomenų rūšys pagal aukščiausią „Taisyklės paramos rodiklį“ .....	60
24 lent. Kiti ne Lietuvoje statistiškai reikšmingai dažnai naudotos duomenų rūšys pagal aukščiausią „Taisyklės paramos rodiklį“ .....	61
25 lent. Ne Lietuvoje naudojamų atvirų duomenų rūšys nebuvo panaudotos su kitomis duomenų rūšimis pagal aukščiausią „Įdiegiamumo rodiklį“ .....	62
26 lent. Lietuvoje naudotos duomenų rūšys pagal aukščiausią „Taisyklės paramos rodiklį“ .....	63
27 lent. Lietuvoje naudojamų atvirų duomenų rūšys nebuvo panaudotos su kitomis duomenų rūšimis pagal aukščiausią „Įdiegiamumo rodiklį“ .....	64

## SANTRAUKA

Atvirai prieinamų duomenų panaudojimo mastai kasmet auga dideliu greičiu. Skaičiuojama, kad atvirai prieinamų duomenų panaudojimas gali padėti kasmet pasaulyje bendrai sugeneruoti ir sutaupyti 3 trilijonus JAV dolerių, o 2020 metais turėtų sukurti ~13 000 naujų laisvų darbo vietų darbui su atvirais duomenimis. Šiuo metu atvirai prieinamų duomenų panaudojimo tema yra ypač aktuali.

Magistrinis darbas sudarytas iš trijų dalių: literatūros analizės, metodikos bei tyrimo, tyrimo rezultatų, išvadų ir rekomendacijų.

Pirmoje darbo dalyje, remiantis moksline literatūra, aprašomi atvirieji duomenys, atvirai prieinamų duomenų teikiama nauda ir duomenų atvėrimo skatinimo programos.

Antroje darbo dalyje aprašoma metodika ir darbo eiga.

Trečioje darbo dalyje analizuojamos pasaulyje atvertų duomenų rūšys, aptariami surinkti ir suklasifikuoti Lietuvoje atvertų duomenų šaltiniai ir rinkiniai, analizuojami ir aprašomi Lietuvoje esantys atvirai prieinamų duomenų panaudojimo atvejai, išanalizuojami Lietuvos ir pasaulio šalių atvirai prieinamų duomenų atvejai, jų naudojamos atvirų duomenų rūšys, ieškoma duomenų rūšių panaudojimo asociacijų.

Rezultatai atskleidė, kad Lietuvoje dažniausiai yra naudojama tik kelių rūšių duomenų rinkiniai: statistiniai, geografiniai ir transporto, todėl ir kuriamų projektų įvairumas labai menkas, priešingai nei kitose pasaulio šalyse. Remiantis asociacijų tyrimo rezultatais galima manyti, kad Lietuvoje nėra užtikrinamas atvirų duomenų panaudojimas – duomenų vertė yra per žema, todėl dalis atvertų duomenų nėra pakartotinai panaudojami.

SAP valdybos nario, Juergeno Muellero (2019) teigimu duomenų vertė nustatoma naudojantis formule:

**Duomenų kiekis \* Duomenų kokybė \* Duomenų panaudojamumas = Duomenų vertė**

Tikėtina, kad mažiausiai vienas iš duomenų vertės dedamųjų yra nepakankamai geros kokybės, todėl visa duomenų vertė tampa žema. Žemos vertės atviri duomenys yra žymiai rečiau arba visai nenaudojami.

## IVADAS

Šiuolaikinėje informacinėje visuomenėje duomenys tapo neatsiejama daugelio prekių ir paslaugų dalimi. Kasdien pasaulyje yra sugeneruojami milžiniški duomenų kiekiai, informacija apie tai ką žmonės perka, skrydžių duomenys, finansinės transakcijos, sveikatos įrašai, socialinių tinklų įrašai ir kt. Šis duomenų srautas kasdien darosi vis svarbesnis ir tampa kertine pasaulinės ekonomikos dalimi ir kitų sektorių, tokių kaip medicina, transportas ar prekyba, kurie ieško galimybių pasinaudojant duomenų analize.

Nuolat didėjant generuojamų duomenų kiekiui, stipriai populiarėja viešai prienami duomenų rinkiniai. Viešo duomenų skelbimo tikslas yra skatinti, kad duomenys būtų kuo daugiau naudojami ir analizuojami, taip pasiekiant socialinių tikslų, tokių kaip skaidrumo didinimas, institucijų atskaitingumas, skurdo mažinimas ir inovacijų skatinimas. Viešai prieinamų duomenų iniciatyva paskatino naujų organizacijų atsiradimą, tokių kaip „Viešai prieinamų duomenų institutas“ („Open Data Institute“)<sup>1</sup> ir „Atvirų žinių fondas“ („Open Knowledge Foundation“)<sup>2</sup>. Šios ir panašios organizacijos nuolat skatina naujų atvirai prieinamų duomenų atsiradimą, taip inicijuodami didelius pokyčius viešai prieinamų duomenų srityje, tokius kaip naujų taisyklių ar įstatymų atsiradimą.

Tai yra tik pirmieji žingsniai, tačiau atvirai prieinamų duomenų judėjimas jau paskatino daugybės duomenų rinkinių atsiradimą. Atvirai prieinami duomenų rinkiniai tinkami kurti naujus produktus ar paslaugas, įgyti konkurencinį pranašumą ar kitaip daryti įtaką skirtingoms sritims ar veikloms. Tam puikiai galima pasinaudoti vyriausybinių duomenų portalo duomenų rinkiniais, tokias kaip įvairios apžvalgos, orų registru, kriminalinių nusikaltimų registru, piliečių išitraukimo į sprendimų priėmimą duomenų rinkiniu ar šalies auto įvykių per ataskaitinį laikotarpį įrašų rinkiniu. Viešai prieinamus duomenis gali teikti institucijos, kompanijos ar pavieniai asmenys, tačiau vyriausybės ar valstybinės institucijos įprastai yra didžiausi duomenų viešintojai. Tačiau beveik nesvarbu ar duomenys yra geo-lokaciniai, metrologiniai, transporto, statistiniai ar kt. jie turi socialinę ir komercinę vertę. 2015 metais Wendy Carrara ir kt. apskaičiavo bendrą atvirai prieinamų duomenų Europos sąjungos šalyse vertę 2016-iems metams, kuri buvo 193 mlrd. eurų ir numatomas augimas iki 256 mlrd. € 2020 metais. Manyika (2013) ir kt. taip pat įvertino, kad atvirai prieinami duomenys kasmet pasaulyje gali sugeneruoti ir padėti sutaupyti apie 3 trilijonus JAV dolerių. Šios vertės sukūrimas pagrįstas duomenų analizės pagalba kuriant naujus tobulesnius ir patrauklesnius produktus ar paslaugas, didinant efektyvumą ar mažinant išlaidas. Pavyzdžiui Londone įsikūrusi didžiųjų duomenų analitikos kompanija Mastodon C naudoja atvirus duomenys tam, kad nustatyti

---

<sup>1</sup> <https://theodi.org/> (peržiūrėta: 2018-12-18)

<sup>2</sup> <https://okfn.org/> (peržiūrėta: 2018-12-18)

nereikalingas išlaidas skiriamas receptiniams vaistams. Dėl šios priežasties Didžioji Britanija sutaupo didelę dalį pinigų skirtą sveikatos apsaugai. Nors iki galo kiekybiškai nėra įvertinta, tačiau atvirai prienami duomenys leidžia kurti daugybę socialinių naudų, tokių kaip gerinti verslo konkurencingumą, tobulinti viešojo transporto paslaugas, gerinti sveikatos apsaugą, mažinti korupciją ir didinti skaidrumą ar skatinti piliečių socialinį, politinį įsitraukimą.

Jorn Berends ir kt. apibrėžę atvirai prienamų duomenų vertės sukūrimo grandinę, kurioje matyti, kad patys duomenys neturi jokios vertės kol jie nėra panaudojami. Duomenys teikia didžiausią naudą juos modifikuojant, apdorojant ar pakartotinai naudojant. Todėl vertės sukūrimas apdirbant duomenis tampa vis aktualesnis ir svarbesnis. Duomenys ir gebėjimas teisingai juos analizuoti tampa žinių ekonomikos ašimi.

Visa ko svarbiausias elementas, tai pati duomenų kuriama nauda, apdirbant ir analizuojant atvirai prienamus duomenis, European data portal ataskaitoje, gaunama nauda yra skirstoma į tris grupes:

- Veiklos – Viešojo sektoriaus efektyvumo gerinimas ir kokybės gerinimas
- Ekonominė nauda – Inovatyvių paslaugų kūrimas ir naujų verslo modelių kūrimas
- Socialinė nauda – skaidrumo bei atskaitomybės gerinimas ir gyventojų įsitraukimo didinimas.

**Tyrimo aktualumas, naujumas** – Atvirai prienami duomenys yra palyginti naujas dalykas ir potencialas yra neišnaudotas, todėl poreikis atverti vis daugiau duomenų nuolat didėja. Šiuo metu atvirų duomenų prienamumas ir duomenų rinkinių kiekis yra ganėtinai menkas, visas atvirų duomenų judėjimas yra dar tik vystymosi stadijoje, ypač Lietuvoje. Atvirai prienami duomenys dažnai yra nemokamas ir tinkamose rankose puikus įrankis gerinant efektyvumą, skatinant visuomenės įsitraukimą ar kuriant verslą, taip pat iš apdorotų duomenų generuojant pajamas.

**Tyrimo tikslas** – Identifikuoti dažniausiai Lietuvoje ir kitose pasaulio šalyse naudojamų atvirų duomenų rūšis ir atlikus naudojamų duomenų rūšių asociacijų tyrimą nustatyti poreikį duomenų atvėrimui Lietuvoje.

**Tyrimo objektas** – Atvirų duomenų rūšių naudojimas.

**Tyrimo uždaviniai:**

1. Išanalizuoti mokslinę literatūrą ir apibūdinti atvirų duomenų sampratą, atvirai prienamų duomenų rūšis ir teikiamą naudą.
2. Surinkti ir susisteminti Lietuvos svarbiausius atvirų duomenų šaltinius. T.y. valstybinių įstaigų ir didžiųjų Lietuvos miestų savivaldybių skelbiamus atvirus duomenis.

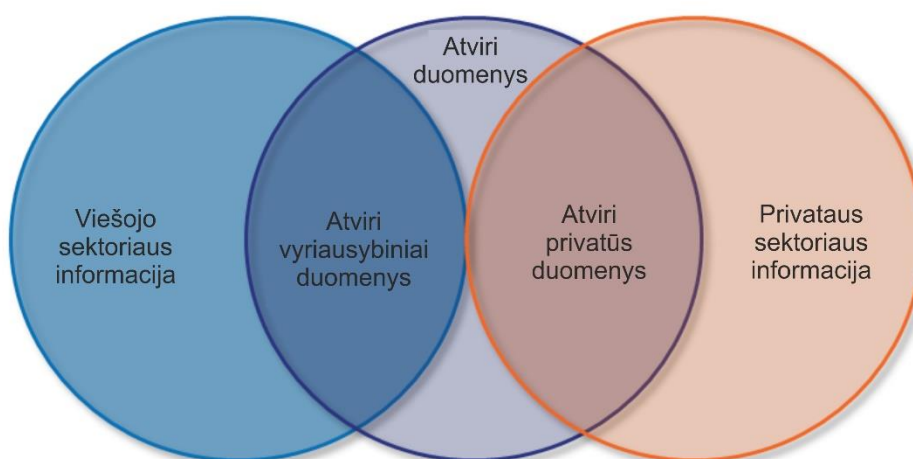
3. Išskirti sėkmingiausius Lietuvos ir pasaulio atvirų duomenų projektus ir išanalizuoti jų naudojamas atvirų duomenų rūšis.
4. Naudojantis asociacijų paieška nustatyti ryšius tarp naudojamų duomenų rinkinių rūšių, remiantis rezultatais nustatyti poreikį duomenų atvėrimui Lietuvoje, pateikti išvadas ir rekomendacijas.



# 1. ATVIRAI PRIENAMŲ DUOMENŲ TEORINIAI ASPEKTAI

## 1.1 Atvirų duomenų samprata ir raida

Siekiant apibrėžti atvirų duomenų sampratą, pradžioje reikėtų aptarti kas yra patys duomenys. Plačiai pripažintą apibrėžimą pateikia T. Davenport ir L. Prusak (1998) teigdami, kad duomenys yra diskretiškų ir objektyvių faktų rinkinys apie praeities įvykius. Organizacijos viduje duomenys įprastai apibūdinami kaip struktūriniai sandorių įrašai. Pasak T. Davenport ir L. Prusak (1998) visos organizacijos turi ir tvarko duomenis, dalies organizacijų veikla yra tiesiogiai priklausoma nuo valdomų duomenų.



1 pav. Duomenų rūšys

(J. Berends, 2017)

J. Berends (2017) sudarytoje lentelėje matome, kad bendrai duomenys turi du šaltinius, tai yra gali būti privataus sektoriaus, verslo ar privačių suinteresuotų asmenų surinkti duomenys arba viešojo sektoriaus renkama informacija. Tie duomenys, kuriais viešasis arba privatusis sektorius laisvai pasidalina be apribojimų galima vadinti atvirais duomenimis. Šiai dienai privatusis sektorius labiau linkęs būti duomenų vartotoju, todėl daugiausiai yra atvirai prieinamų vyriausybinių duomenų rinkinių.

Mokslinėje literatūroje sutinkami keli skirtingi atvirai prieinamų duomenų apibrėžimai, dažniausiai apibūdinimo skirtumai atsiranda atsižvelgiant į duomenų teikėją, duomenų tipą, duomenų naudotojus, duomenų formatą ir duomenų kilmės tipą.

Pasak Lauros James (2013) – Atvirai prienami duomenys, tai duomenys, kuriais be apribojimų gali naudotis, dalintis ir kurti, bet kas, nepriklausomai kurioje pasaulio vietoje, bet koku tikslu. Ali M. Al-Khouri (2014) atvirai prieinamus duomenis apibrėžia kaip duomenų dalį ar turinį, kuriuo gali bet kas pasinaudoti, naudoti pakartotinai ar dalintis. Duomenys yra laisvai prienami su galimybe juos perduoti savo nuožiūra be jokių autoriaus teisių apribojimų, patentų ar kitų kontrolės mechanizmų. Oficialiame JAV atvirai prieinamų duomenų tinklalapyje Data.gov<sup>3</sup> galime rasti, kad atvirai prienami duomenys apibrėžiami kaip didelės vertės duomenys, paruošti darbui kompiuteriu, kuriuos kuria vyriausybė, kurie neturi nacionalinio saugumo apribojimų ir yra patalpinti viešoje duomenų saugykloje. Pasaulinės atvirų duomenų skatinimo organizacijos „The Open Data Barometer“ ataskaitoje<sup>4</sup> (2015) viešai prienami duomenys apibrėžiami kaip duomenys, kurie atitinka laisvos prieigos kriterijus, neatsižvelgiant į vartotojo tapatybę ar tikslus, skaitmeniniame, darbui kompiuteriu paruoštame formate, be jokių apribojimų naudoti ar dalintis. Kitokį požiūrį galima rasti Europos komisijos 2003 metų atvirų duomenų ataskaitoje<sup>5</sup>. Atvirai prienami duomenys, tai duomenys, kuriuos viešasis sektorius renka, apdirba, daugina ir platina.

Visi išvadinti apibrėžimai turi ir panašumų ir skirtumų, tačiau pagrindiniai punktai išlieka visuose apibrėžimuose vienodi: duomenys turi būti laisvai prienami visiems, neturėti apribojimų su jais dirbti ar platinti ir duomenys turėtų būti paruošti importavimui į kompiuterines programas.

Žodis „atviras“ (Open) apibrėžiamas ir turi atitikti 15 punktų<sup>6</sup>:

1. Darbas turi būti viešai prienamas arba suteikta prieiga su atvira licencija (kaip apibrėžta 5 punkte). Bet kokios papildomos sąlygos, pridedamos prie darbo (pvz., naudojimo sąlygos, licencijos ar išdavėjo turimi patentai), neturi prieštarauti viešo naudojimo darbo sričiai ar licencijos sąlygoms.
2. Darbas turi būti prienamas pilno turinio su galimybe atsisiųsti internetu, taip pat su darbu turi būti pridėta visa papildoma licencijos atitikties informacija, tokia kaip autorių vardai.
3. Duomenys turi būti paruošti darbui kompiuteriu, duomenys turi būti pateikti formatu, kuriuo galima be papildomo apdorojimo pradėti darbą. Atskiros darbo dalys turi būti lengvai prienamos ir modifikuojamos.

---

<sup>3</sup> <https://www.data.gov/> [žiūrėta 2018-12-19]

<sup>4</sup> <https://opendatabarometer.org/assets/downloads/Open%20Data%20Barometer%20-%20Global%20Report%20-%202nd%20Edition%20-%20PRINT.pdf> [žiūrėta 2018-12-19]

<sup>5</sup> [http://www.europarl.europa.eu/RegData/etudes/BRIE/2018/628312/EPRS\\_BRI\(2018\)628312\\_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2018/628312/EPRS_BRI(2018)628312_EN.pdf) [žiūrėta 2018-12-20]

<sup>6</sup> <https://opendefinition.org/od/2.1/en/> [žiūrėta 2018-12-19]

4. Duomenys turi būti pateikti atviru formatu su galimybe juos apdoroti naudojant nemokamą programinę įrangą.
5. Darbo licencijos turi būti atviros ir neprieštarauti jokioms kitoms licencijoms. Licencijos turi atitikti sekančius reikalavimus:
  - Licencija turi leisti laisvai perskirstyti licencijuotą darbą, įskaitant kelių apjungtų rinkinių pardavimą.
  - Licencija turi leisti kurti licencijuojamo kūrinio išvestines priemones ir leisti platinti tokius išvestinius produktus tokiomis pačiomis sąlygomis kaip originalus licencijuotas darbas.
  - Licencija turi leisti bet kokią darbo dalį laisvai naudoti, platinti ir modifikuoti atskirai nuo bet kokios kitos darbo dalies.
  - Licencija turi leisti platinti licencijuotą darbą kartu su kitais darbais, netaikant apribojimų atliktiems darbams.
  - Licencija neturi diskriminuoti jokio asmens ar grupės.
  - Teisės, susijusios su darbu, turi būti taikomos visiems, nereikalaujant papildomo sutikimo su teisinėmis sąlygomis.
  - Licencija turi leisti naudoti, perskirstyti, modifikuoti ir kompiliuoti bet kokiam tikslui. Licencija negali apriboti duomenų naudojimo taikant jokia konkrečiai sričiai.
  - Licencija negali numatyti jokio mokesčio, autorinio atlygio, kompensacijos ar kito piniginio atlygio.

2007 gruodžio 7-8 dieną posėdžiaudami Kalifornijoje, JAV iš 30 advokatų sudarytą atvirai prieinamų duomenų darbo grupė buvo išsikėlusį tikslą sudaryti 10 punktų atvirai prieinamų duomenų principų rinkinį. Carl Malamud vadovaujama darbo grupė po dviejų dienų sudarė aštuonių punktų rinkinį. Remiantis punktų rinkiniu (2007) <sup>7</sup>atvirai prieinami duomenys turi būti:

### **1. Užbaigti**

Visi vieši duomenys yra prieinami. Vieši duomenys – tai duomenys, kuriems netaikomi galiojantys privatumo, saugumo ar privilegijų apribojimai.

### **2. Pirminiai**

Duomenys yra renkami šaltinyje, turintys aukščiausią įmanomą kokybę ir tikslumą, ne apdoroti ar kitaip pakeisti.

---

<sup>7</sup> <https://opengovdata.io/2014/8-principles/> [žiūrėta 2020-01-04]

### **3. Laiku**

Duomenys pateikiami kuo greičiau, kad būtų išsaugota duomenų vertė.

### **4. Prieinami**

Duomenys yra prieinami plačiajai naudotojų grupei, turintiems daugybę tikslų.

### **5. Paruošti kompiuteriniam apdorojimui**

Duomenys yra racionaliai struktūrizuoti, kad būtų galima automatizuotai apdoroti kompiuteriu.

### **6. Nediskriminaciniai**

Duomenys yra prieinami visiems, nereikalaujant registracijos ar kitų duomenų ar informacijos.

### **7. Nekomerciniai**

Duomenys yra prieinami tokiu formatu, kuris neribotų duomenų panaudojimo visiems ar nesuteiktų pranašumo tam tikrai grupei žmonių.

### **8. Be licencijos**

Duomenims netaikomi jokie autorių teisių, patentų, prekių ženklų ar komercinės paslapties reglamentai. Gali būti leidžiami pagrįsti privatumo, saugumo ir privilegijų apribojimai.

- Taisyklių rinkinio punktų laikymasis turi būti peržiūrimas.
- Turi būti paskirtas kontaktinis asmuo, atsakantis į žmonių, bandančius naudoti duomenis, klausimus.
- Turi būti paskirtas kontaktinis asmuo, atsakantis į skundus dėl principų pažeidimų.
- Administracinis arba bendrosios kompetencijos teismas turi turėti jurisdikciją patikrinti, ar agentūra tinkamai taikė šiuos principus.

## **1.2 Atvirai prieinamų duomenų istorija**

Atvirai prieinami duomenų ir duomenų viešinimo svarba pastebėta tik palyginti visai neseniai, todėl šiuo metu tai yra visame pasaulyje sparčiai auganti sritis. Turint galvoje, kad terminas „atvirai prieinami duomenys“ pirmą kartą buvo paminėtas 1995 metais Amerikos mokslų agentūros dokumente. Dokumentas buvo susijęs su geofizinių ir aplinkos duomenų atskleidimu. Bendros gerovės žinių dalijimosi idėja atsirado gerokai prieš interneto atsiradimą. Amerikietis, sociologijos

mokslų specialistas Robert K. Merton 1942 metais plačiai aiškino tyrimo rezultatų populiarinimo ir dalijimosi svarbą ir ragino visus mokslininkus ir tyrėjus dalintis savo atliktų darbų rezultatais, taip prisidedant savo indeliu į bendrą „mokslo puodą“ ir atsisakyti intelektualinės nuosavybės teisių, kad žinios galėtų laisvai judėti į priekį.

Tačiau tik 2007 metų gruodį Kalifornijoje, JAV 30 atvirai prieinamų duomenų pradininkų sudarė aštuonių atvirai prieinamų duomenų principų rinkinį, kuriuo punktus atitinkančius duomenis galima klasifikuoti kaip atvirai prieinamus duomenis.

Informacinės technologijos taip pat įnešė naujų vėjų šios bendrosios filosofijos atžvilgiu. Savo tyrimuose 2009 m. Nobelio premijos laureatė Elinor Orstrom parodė informacijos bendrumą, siekį dalintis ir teigė, kad tai kaip viešosios gėrybės, vieno žmogaus vartojimas neįtakoja kito žmogaus vartojimo ir nemažina duomenų vertės. Tai yra naujos rūšies viešosios gėrybės, jų vartojimas ne tik ne mažina vertės, bet tik praturtina.

Anksčiau tokie duomenys buvo prieinami tik mokslininkams, tai jie pirmieji ir išvelgė ir ragino platinti duomenis, taip visiems gaunant naudą. Mokslininkai ir tyrėjai formavo atvirų duomenų sąvoką, tokią kokią mes ją žinome šiandien.

### **1.3 Atvirai prieinamų duomenų rūšys ir populiarumas**

Atvirai prieinamų duomenų populiarumas nuolat didėja ir turi beveik neribotas panaudojimo galimybes, tačiau gana svarbu žinoti aiškų atvirai prieinamų duomenų klasifikavimą. Pasak Lauros James (2013), atvirai prieinami duomenys yra skirstomi į aštuonias grupes pagal tai kokią informaciją duomenų rinkiniai saugo.



2 pav. Atvirai prieinamų duomenų rūšys

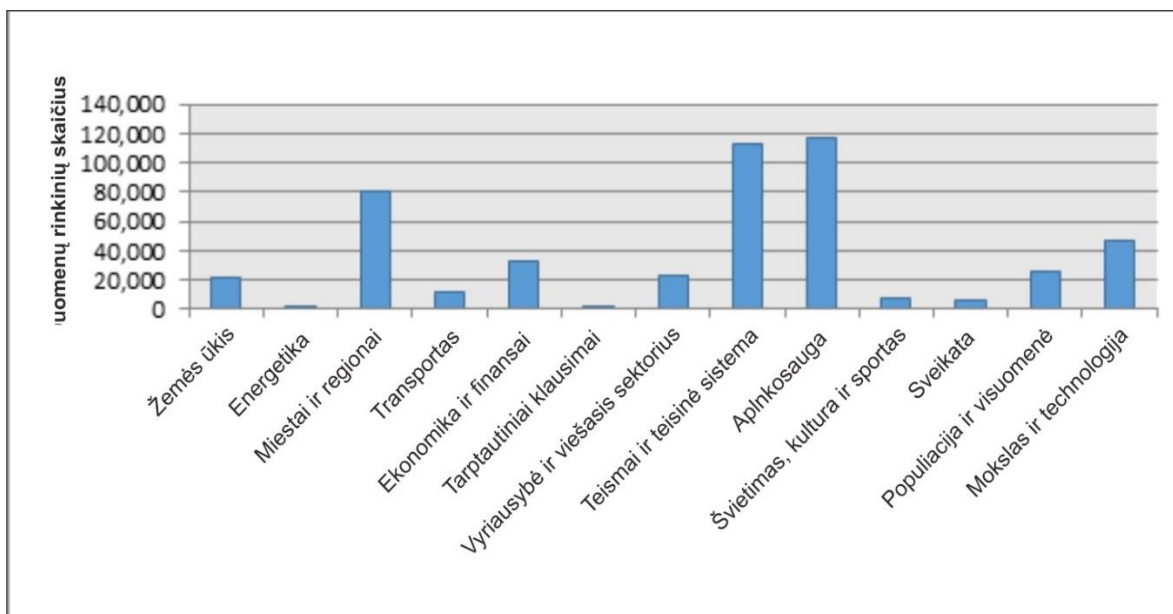
(Laura James, 2013)

- Geografiniai duomenys – Tai duomenys skirti sudaryti žemėlapius, naudojant duomenis, kuriuose nurodytos kelių ar pastatų vietos, taip pat naudojama kompleksiniams projektams, tokiems kaip nusikaltimų žemėlapis ir pan. pritaikant prie kitų duomenų rinkinių ir atvaizduojant norimą informaciją žemėlapyje.
- Kultūros duomenys – Duomenys apie kultūrinę veiklą ir artefaktus, pavyzdžiui informacija apie knygas – knygos autorius, pavadinimas ir informaciją apie leidinį. Kultūros informacija įprastai disponuoja galerijos, bibliotekos, archyvai ir muziejai.
- Moksliniai duomenys – Duomenys dar dalis duomenų, kurie buvo sugeneruoti mokslinio tyrimo metu, pavyzdžiui astronominiai, zoologiniai ar pan. duomenys
- Finansiniai duomenys – Visi valstybės biudžeto, išlaidų duomenys, taip pat finansų rinkų duomenų rinkiniai, akcijų biržų, obligacijų ar kiti finansiniai duomenys
- Statistiniai duomenys – Visi statistiniai duomenys, kaip žmonių skaičius, bedarbių statistiniai duomenys, įmonių statistika ir pan. Daugiausiai tokių duomenų rinkinių valdo statistikos departamentas.
- Metrologiniai duomenys – Visi duomenys susiję su orais, tai valandiniai orų įrašai, lietaus, debesuotumo, ar žemės drebėjimų įrašai padedantys prognozuoti ar padėti nuspėti galimą orą ar net stichinę nelaimę.

- Aplinkos duomenys – Duomenys padedantys saugot aplinką, atkreipti dėmesį į taršos susitelkimą ar jos didžiausius sukėlėjus, duomenų rinkiniai su informaciją apie upių ar vandens telkinių užterštumą, taršos pasiskirstymą pagal vietovę, oro kokybės duomenys ir pan.
- Transporto duomenys – Visi duomenys su viešojo transporto tvarkaraščiais, maršrutais, vėlavimo statistika, taip pat kelių srauto ar spūsčių duomenys. Su šiais duomenimis galima tobulinti viešojo transporto kokybę ar mažinti miesto spūsčių problemas.

Europos duomenų portalas renka viešojo sektoriaus informacijos metaduomenis, skelbiamus tiek nacionaliniu, tiek regioniniu, tiek vietos portaluose visoje Europoje. Šiuo metu renkamas 71 katalogas, iš kurių Europos duomenų portale galima rasti daugiau kaip 616 tūkst. Duomenų rinkinių. Šie duomenų rinkiniai suskirstyti į 13 duomenų kategorijų. Duomenų kategorijos yra duomenų rinkinių, susietų su bendra tema kategorizavimas. 13 duomenų kategorijų, kurias galima rasti Europos duomenų portale, atitinka kategorijas, kurias Europos Komisija nustatė persvarstant Europos duomenų katalogo taisyklių rinkinio (DCAT) taikomąją informaciją.

Bet kokios yra duomenimis turtingiausios duomenų kategorijos? Žvelgiant į duomenų rinkinių prieinamumą Europos duomenų portale kiekvienai kategorijai, galima pastebėti, kad kategorija „Aplinka“ (> 117 tūkst. Duomenų rinkinių) ir „Teisingumas, teisinė sistema“ (> 112 tūkst. Duomenų rinkinių) turi daugiausia duomenų. Po jų seka kategorija „Regionai ir miestai“, siūlanti daugiau nei 81 tūkst. Duomenų rinkinių. Tačiau, tokios kategorijos kaip energetika, tarptautiniai klausimai ir sveikata apima mažiau nei 10 tūkst. Duomenų rinkinių. Tai rodo, kad esama didelių duomenų rinkinių kiekių skirtumų tarp kategorijų, tai galima būtų paaiškinti, kad šalys išskirtinai orientuojasi į konkrečių duomenų išleidimą taip apleisdami kitų duomenų viešinimą arba tiesiog tam tikros sritys generuoja daugiau kokybiškų duomenų, dėl galimai stipresnės organizacinės veiklos populiariausiose srityse.



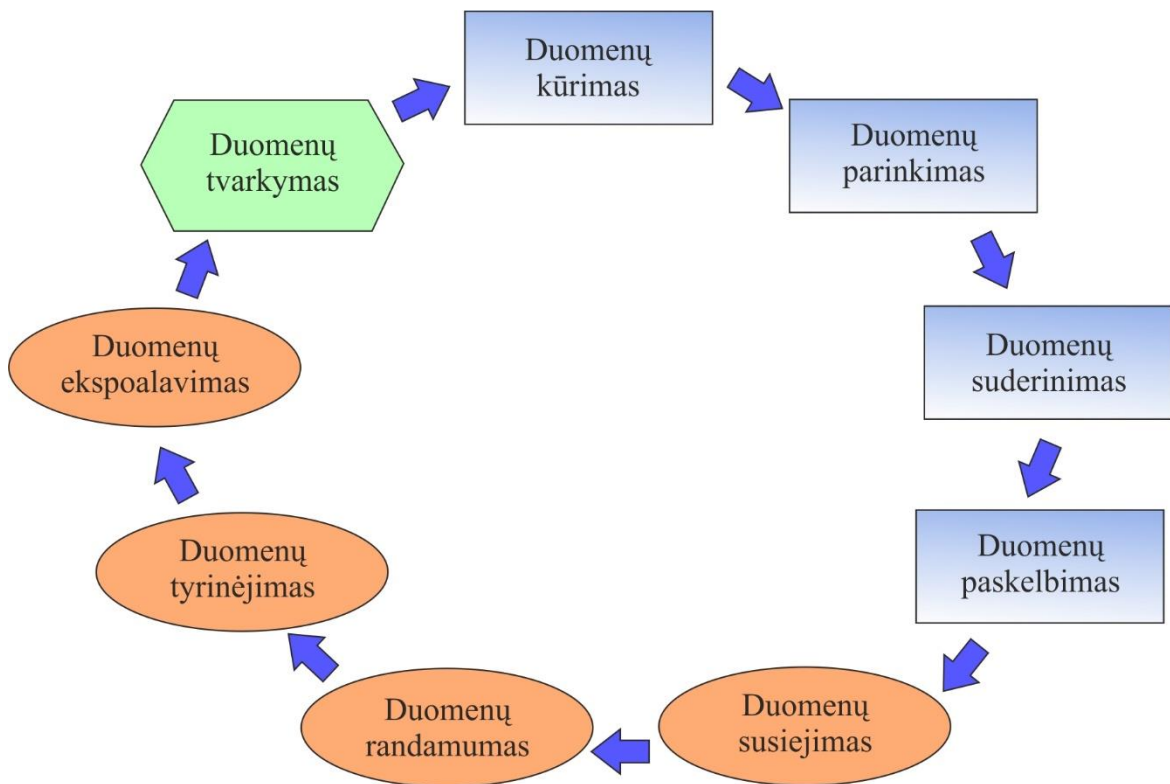
3 pav. Populiariausi duomenų rinkiniai

(Open data portal, 2016)

## 1.4 Atvirai prieinamų vyriausybinių duomenų gyvavimo ciklas

Nepaisant didelio atvirai prieinamų duomenų kiekio, bene didžiausią jų dalį sudaro vyriausybinių atviri duomenys, todėl svarbu apžvelgti atvirų vyriausybinių duomenų gyvavimo ciklą. Juddie Attard ir kt. (2016) pateikia bendrą atvirai prieinamų vyriausybinių duomenų gyvavimo ciklą, kaip pagrindą visų bazinių procesų, kuriais galima vadovautis. Siūlomas gyvavimo ciklas pavaizduotas 4 paveikslėlyje, kuris sudarytas iš trijų dalių, kur duomenų apdorojimas yra stačiakampiuose, eksploatacinė dalis – ovalo formos, o priežiūros skiltis šešiakampiuose.





4 pav. Atvirai prieinamų vyriausybinių duomenų gyvenimo ciklas

(Judie Attard, 2016)

**Duomenų kūrimas** – Atvirai prieinamų valstybinių duomenų gyvavimo ciklas prasideda nuo duomenų kūrimo. Valstybinėse institucijose duomenų kūrimo procedūra yra įprasta kasdienio darbo dalis.

**Duomenų parinkimas** – Tai procesas, kuriuo metu parenkami duomenis paskelbimui. Tam būtina pašalinti asmeninius ar asmens duomenis, taip pat nustatyti kur ir kokiomis sąlygomis duomenis bus paskelbti.

**Duomenų suderinimas** – Šis etapas skirtas duomenų paruošimui, tvarkymui prieš paviešinant laikantis duomenų tvarkymo standartų.

**Duomenų paskelbimas** – Tai procedūra kaip duomenys yra paskelbiami, patalpunami į viešai prieinamas duomenų bazines.

**Duomenų susiejimas** – Duomenų susiejimo etapas leidžia duomenų rinkiniams suteikti papildomą vertę. Susijusių duomenų priskyrimas leidžia paskelbtiems duomenis sudaryti kontekstą ir palengvinti informacijos teisingą suvokimą.

**Duomenų randamumas** – Vien duomenų paskelbimo neužtenka, duomenų vartotojai ir analitikai turi žinoti apie duomenų egzistavimą arba gebėti lengvai rasti norimus duomenis. Todėl svarbu populiarinti viešai prieinamus duomenis ir jų teikiamą naudą.

**Duomenų tyrinėjimas** – Šis žingsnis bene svarbiausias visame duomenų cikle. Vartotojas duomenis tyrinėja juos atvaizduodamas ar struktūruodamas, taip suteikdamas aiškų vaizdą turimiems duomenims.

**Duomenų eksploatavimas** – Šis būdas yra pažangesnis būdas vartoti duomenis. Šiame etape vartotojas analizuoja, atlieka prognozes, pakartotinai naudoja ar platina turimus duomenis.

**Duomenų tvarkymas** – Nors šis esamas ne visada įvykdomas, tačiau jis yra labai svarbus norint išlaikyti paskelbtų aukštą duomenų kokybę. Šis etapas apima duomenų atnaujinimą, duomenų praturtinimą, patikslinimą ir kt.

## 1.5 Susietieji duomenys ir atvirai prieinami duomenys

Siekiant praturtinti ir padidinti duomenų vertę labai svarbus vienas iš Judie Attard (2016) paminėtų procesų – duomenų susiejimas. Susietieji duomenys, tai apibrėžimas, kurį sukūrė Tim Berners-Lee (2006). Tim Berners-Lee (2006) viešai pareiškė, kad duomenų talpinimas į viešai prieinamus internetinius puslapius yra tik pirmas žingsnis, toliau būtina duomenis susieti su kitais talpinamais duomenų rinkiniais, taip atveriant visą duomenų aplinką ir lengvinant darbą su duomenimis.

Susietieji duomenys, tai duomenų paskelbimo metodas, kaip jau nurodo pavadinimas, yra vienas iš būdų, kuriais duomenys susiejami su kitų žmonių duomenimis. Tai turi tam tikras technines taisykles, kuriuos Tim Berners-Lee pristatė TED konferencijoje 2009 metais:

1. „Viskas turi savo unikalų vardą ir jis prasideda HTTP“
2. „Jei aš gaunu informaciją, aš ją paverčiu duomenimis, standartizuotu formatu, tinkamu naudoti kompiuterinėmis programomis, kas yra gana patogiu norint daugiau žinoti apie kokį nors dalyką ar įvykį“
3. „Kai gaunu informaciją, tai gauni ne žmogaus ūgį, svorį ir kada jis gimė, tai turi ryšį. Tai visa susijusi informacija turi turėti savo HTTP vardą priskirtą pateiktai informacijai. Pavyzdžiui žmogus gimė Berlyne, turime nuorodą į duomenis apie Berlyną. Berlynas Vokietijoje – nuorodą į duomenis apie Vokietiją ir pan. Tai ir yra priskirtieji duomenys.“

Taip pat 2006 metais Tim Berners-Lee pateikė keturis pagrindinius susietųjų duomenų principus:

1. Pavadinimuose naudokite URI siekiant identifikuoti elementus.
2. Naudokite HTTP URI, kad jūsų elementai pagal pavadinimą būtų lengviau randami ir jūsų pateikta informacija pasiektų kuo daugiau suinteresuotų asmenų.
3. Kai kažkas ieško pagal URI, pateikite naudingos informacijos naudojantis tokiais standartais kaip RDF ir SPARQL
4. Pateikite nuorodas į kitus URI, kad suinteresuoti asmenys galėtų lengvai surasti susijusias, naudingos informacijos.

Remiantis Open data portas susietųjų duomenų apžvalga būtina apibrėžti panaudotas sąvokas:

**URI** – Vieningas išteklių identifikatorius (angl. Uniform Resource Identifier – URI) yra kompaktiška raidžių seka, kuri identifikuoja abstraktų arba fizinį išteklių.”

Pvz.: Žmogus, Artūras Karnišovas • [https://en.wikipedia.org/wiki/Artūras\\_Karnišovas](https://en.wikipedia.org/wiki/Artūras_Karnišovas)

**RDF** (angl. Resource Description Framework) – tai sintaksė skirta duomenų ir išteklių apibūdinimui žiniatinklyje, kurios konstrukciją sudaro trys elementai:

**Subjekto:** dalyko, kuris identifikuojamas URI pagalba

**Teiginio:** Teiginio, kuris apibūdina ryšį/

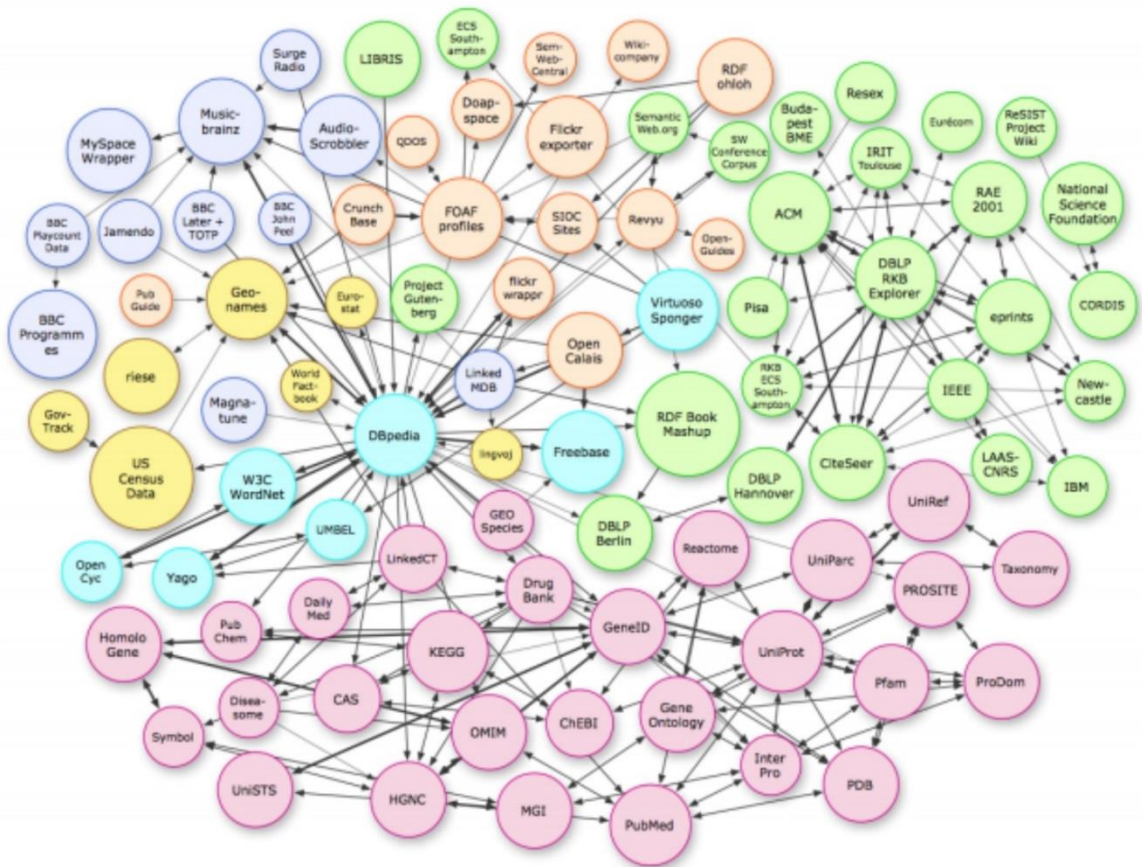
**Objekto:** Dalyko susijusio su subjektu.

<http://dbpedia.org/page/Vilnius> is the capital of “Lithuania”.

Arba

<http://dbpedia.org/page/Vilnius> is the capital of <http://dbpedia.org/page/Lithuania> .

**SPARQL** – duomenų aprašytų RDF standartu užklausų kalba.

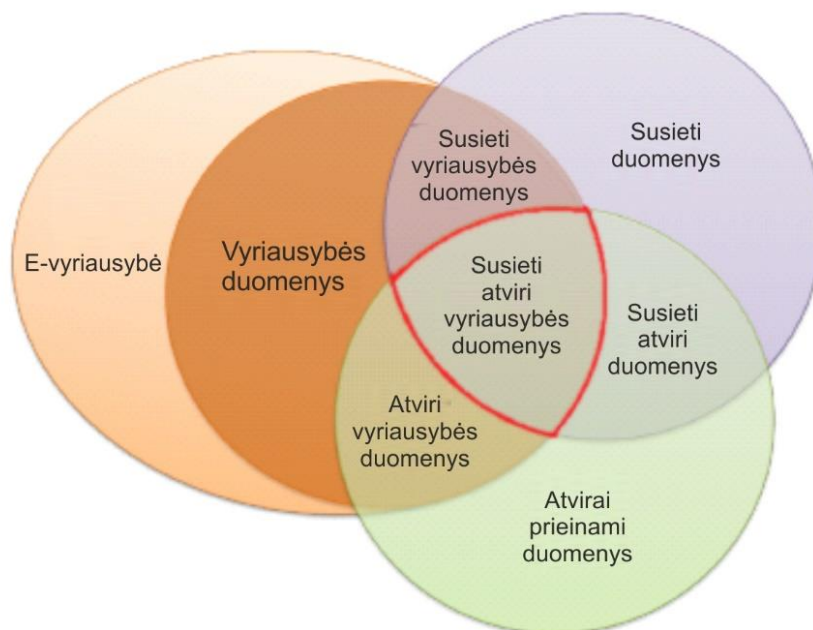


5 pav. Susietųjų duomenų tinklas

(Tim Berners-Lee, 2009)

Taip pat Tim Berners-Lee (2009) pateikė 5 žvaigždučių schemą, kaip galima būtų vertinti duomenų kokybę:

- ★ Duomenys yra internete, nepaisant formato, bet turi atvirą licenciją ir priėjimą visiems žmonėms.
- ★ ★ Duomenys yra pritaikyti ir pateikti kompiuteriui suprantamu formatu. (pvz.: Excel, nuotraukos su tekstine informacija).
- ★ ★ ★ Duomenys pateikti patogesniu formatu (pvz.: CSV vietoje Excel).
- ★ ★ ★ ★ Visi anksčiau išvardinti ir panaudotas atvirumo standartas iš W3C (RDF, SPARQL) skirtas identifikuoti pateiktus duomenis, kad kiti galėtų jais dalintis.
- ★ ★ ★ ★ ★ Visi anksčiau išvardinti ir papildomai susieti savo duomenis su kitų žmonių susijusiais duomenimis, kad sukurti kontekstą ir leisti detaliau matyti ir pasinaudoti duomenimis.



6 pav. Ryšys tarp atvirai prieinamų, vyriausybinių ir susietųjų duomenų

(Judie Attard ir kiti, 2015)

## 1.6 Atvirai prieinamų duomenų skatinimo valstybinės programos

Bene svarbiausias dalykas siekiant išlaikyti ir populiarinti atvirus duomenis yra atvirų duomenų skatinimo programos valstybiniu mastu. Atvirų duomenų galima nauda, populiarumo didinimo galimybės, tendencijos ir kita informacija yra plačiai aptarinėja, todėl įdomu kokiomis programos valstybės skatina atvirų duomenų rinkinių daugėjimą ir kur įžvelgia naudą. Remiantis Noor Huijboom (2011) ir kitų atlikta viešai prieinamų duomenų skatinimo valstybinių programų palyginimu ir autoriaus Lietuvos programos analize matome, kad programų tikslai skiriasi. Šalių vadovai skirtingai mato galimą viešai prieinamų duomenų naudą. JAV išskiria siekį skatinti skaidrumą ir pilietiškumą, kai tuo tarpu Danija mato galimybę skatinti naujų produktų ir paslaugų atsiradimą, Ispanija siekia ekonominės naudos vietos verslams, o Lietuva siekia piliečių aktyvumo ir valstybės valdymo skaidrumo.

Šalis	Pradžios data	Atsakinga institucija	Tikslai
Australija	2010 m.	Vyriausybės informacijos valdymo biuras	Siekiami ekonominės naudos gyventojams ir verslams, inovacijų skatinimas.
Danija	2010 m.	Mokslo, technologijų ir inovacijų ministerija	Tikimasi naujų verslų atsiradimo informacinių technologijų srityje skaitmeninių paslaugų plėtojime.
Didžioji Britanija	2009 m.	Vyriausias išduos sekretorius	Sustiprinti pilietiškumą, atveriant duomenis didinti skaidrumą, duomenų pagalba didinti informuotumą, taip padedant piliečiams ir verslams turėti ekonominės naudos. Taip pat skatinti gyventojų įsitraukimą mažinti nusikalstamumą
Ispanija	2010 m.	Industrijos, turizmo ir komercijos ministerija	Viešinant vyriausybės duomenų rinkinius bus skatinama ekonominė nauda šalies verslams ir piliečiams. Taip pat siekiama paskatinti naujų produktų ir paslaugų kūrimą.
JAV	2009 m.	Baltieji rūmai ir valstijos departamentas	Sukurti beprecedentį vyriausybės atvirumo lygį. Atvirumas turėtų sustiprinti demokratiją, skatinti vyriausybės naudingumą ir efektyvumą, skaidrumas parodyti piliečiams ką vyriausybė veikia, taip didinant pasitikėjimą.
Lietuva	2011 m.	Informacinės visuomenės plėtros komitetas, sveikatos apsaugos, transporto, teisingumo ir finansų ministerijos	Skaitmeninės valdžios skatinimas, informacijos prieinamumo ir visuomenės dalyvavimo valdyme skatinimas, didesnės atskaitomybės ir skaidrumo bei kovos su korupcija siekis.

1 lent. Valstybių priimtos atvirų duomenų skatinimo programos

(sudaryta autoriaus remiantis Noor Huijboom, 2011)

Analizuojat sudarytą lentelę matome, kad vyrauja trys pagrindinės viešai prieinamų duomenų panaudojimo motyvacijos:

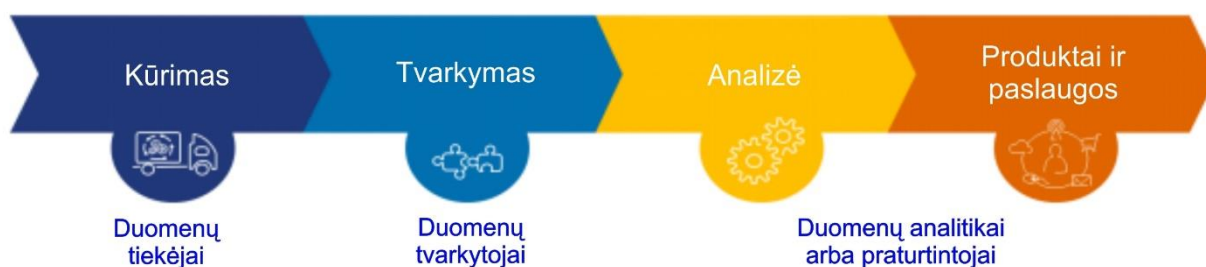
1. Politinio įsitraukimo ir demokratijos skatinimas. Viešinant duomenis siekiama skatinti gyventojų susidomėjimą vyriausybės veikla, taip didinant pasitikėjimą ir skatinant skaidrumą. Didžiosios Britanijos vyr. izdo sekretorius pareiškė: „kuo labiau gyventojai yra informuoti, tuo daugiau galios jie turi. Modernioje visuomenėje pilietis turi žinoti kur jo pinigai buvo išleisti ir koks rezultatas pasiektas.“
2. Naujų paslaugų ir produktų skatinimas. Danijos vyriausybė teigia, kad viešai prieinami duomenys turėtų paskatinti inovacijų atsiradimą. Gyventojai panaudoję turimus duomenis ir pasitelkę išradingumą gali juos praktiškai pritaikyti ir spręsti kasdieniais problemas. Taip pat viešai prieinami duomenys gali generuoti ekonominę naudą, skatinti verslų atsiradimą ir pan.
3. Teisėsaugos stiprinimas. Viešinant vyriausybės duomenis siekiama skatinti žmonių įsitraukimą mažinant nusikalstamumą. Lietuvos informatikos ir ryšių departamentas įgyvendino projektą „Nusikalstamų veikų Lietuvoje žemėlapis“, tai virtualus žemėlapis su visais įvykusiais nusikaltimais Lietuvoje, jo aprašymu, tiksliu laiku ir panašiais duomenimis. Taip skatinama pažinti savo kaimynystę, gebėti pažinti nusikaltimą ar žinoti kur jo galima tikėtis, taip apsaugant save ar savo artimą žmogų.

## **1.7 Atvirai prieinamų duomenų teikiama nauda bei ypatumai**

Nuo pat 2007 gruodžio, kai Sebastopolyje, kai Kalifornijoje susirinko 30 atvirų duomenų pionierių ir surinko aštuonių atvirų duomenų principų rinkinį, kuris pradėjo naują demokratinės naujovės ir ekonominių galimybių erą, visada lydi klausimas: kokią naudą teikia atvirų duomenų paskelbimas kiekvienam žmogui asmeniškai, verslui, visuomenei, vyriausybės institucijoms ir šaliai bendrai. Europa žengia sparčiu žingsniu atvirai prieinamų duomenų viešinimo brandos link, matydami didelį potencialą. Naujausi tyrimai rodo, kad didesnės ekonominės naudos tikimasi iš atvirų duomenų tiek makroekonominiu, tiek mikroekonominiu lygiu. Duomenys turi didžiulį potencialą visuomenei ir ekonomikai. Atskleidus duomenis, galima paskatinti ekonomikos augimą, geresnį sprendimų priėmimą, didesnę vyriausybių skaidrumą ir veiksmingumą, taip pat aukštesnę gyvenimo kokybę ir labiau įsitraukusią visuomenę. Daugelyje Europoje ir pasaulyje atliktų tyrimų matuojamas duomenų poveikis ekonominiam, politiniam ir visuomeniniam vystymuisi, keletas tyrimų nurodė atvirų duomenų svarbą ekonomikos augimui. Atviri duomenys laikomi ekonomikos

augimo varikliu ir padeda užtikrinti skaidrumą ir atskaitomybę, taip pat inovacijų ir žinių skatinimą. 2017 metų gruodį paskelbtame atvirų duomenų ekonominės naudos pranešime<sup>8</sup> nagrinėjami iki šiol turimi įrodymai ir pateikiami išsamūs įvairūs Europoje atliktų vertinimų ir tyrimų rinkiniai.

Open data portal 2017 metų ataskaitoje pateikta kaip neapdirbti duomenis panaudojami ir išgaunama nauda. Duomenų kūrimą atlieka tiekėjai, vėliau duomenys renkami ir kaupiami bendrai veikiančių duomenų tvarkytojų/administratorių. Praturtintojai naudoja duomenis naujoms programoms, produktams ar paslaugoms kurti, o analitikai naudoja duomenis, kad gautų naujų ir (arba) geresnių įžvalgų iš duomenų analizės.



7 pav. Duomenų vertės kūrimo grandinė

(Jorn Berends ir kt., 2017)

Remiantis European data portal ataskaita (2017), 38 % organizacijų naudojančių atvirai prieinamus duomenis turėtų būti klasifikuojami kaip duomenų tvarkytojai. Tai reikštų, kad atvirai prieinamų duomenų nauda tik papildo jau jų teikiamą paslaugą, todėl duomenys yra tik paslauga papildanti informacija. Tik 12% organizacijų naudoja atvirai prieinamus duomenys kaip jų pagrindinį veiklos šaltinį plėtojant teikiamą paslaugą internetiniame puslapyje ar mobiliojoje programėlėje, tai tokie modeliai kai paslaugos turinys yra klientui reikiamo atsakymo pateikimas, reikiamos informacijos pateikimas arba reikiamų duomenų pateikimas realiu laiku. Atvirai prieinami duomenys ne visada yra tiesiogiai siejami su generuojamu pelnu. Apklausus atvirai prieinamus duomenis naudojančias organizacijas paaiškėjo, kad 47 % vadovų paminėjo inovacijas, o 26% vadovų teigia pradėję naudoti atvirai prieinamus duomenis sumažino išlaidas ir padidino veiklos efektyvumą.

Spartus duomenų ekonomikos augimas paskatino tolimesnį šio sektoriaus įmonių atsiradimą. Tai reiškia, kad atsirado nemažai įmonių, kurių verslas paremtas duomenų vertės kūrimo grandinės modeliu. Bendrovės analizuoja ir apdoroja informaciją gautą iš viešų ir (arba) privačių

<sup>8</sup> [https://www.europeandataportal.eu/sites/default/files/analytical\\_report\\_n9\\_economic\\_benefits\\_of\\_open\\_data.pdf](https://www.europeandataportal.eu/sites/default/files/analytical_report_n9_economic_benefits_of_open_data.pdf) [žiūrėta 2018-12-31]



šaltinių ir kuria pridėtinę vertę teikiančius produktus ar paslaugas, kurios leidžia tretiesiems asmenims ar visuomenei palengvinti ir pagreitinti sprendimų priėmimo procesą.

Ispanijos daugiasektorinės informacijos asociacija „ASEDIE“<sup>9</sup> kasmet skelbia ataskaitą apie infomedijos sektoriaus padėtį. 2017 m. penktas paskelbtas ataskaitos leidimas, kuriame pateikiama išsami 636 įmonių analizė, kurių veikla grindžiama viešosios ir (arba) privačios informacijos pakartotiniu naudojimu kuriant pridėtinę vertę trečiosioms šalims ir (arba) piliečiams. Šių įmonių veikla apima įvairius sektorius, pradedant nuo kultūros, ekonominė ir finansinė, leidyba, rinkos tyrimai, geografinė ar meteorologinė informacija ir (arba) turizmas, šių įmonių bendra apyvarta siekia 1,7 mlrd. 2015 m. ir jose dirba apie 19.300 žmonių.

Didžiausias augimas buvo ekonomikos ir finansų sektoriuose – 219 darbuotojų daugiau nei 2014 m., po to sekė geografinės informacijos pramonė, kurioje dirba 215 darbuotojų daugiau lyginant su 2014 m. Tyrimas taip pat parodė, kad atvirai prieinamų duomenų naudojimas skatina naujovių diegimą viešajame sektoriuje kuriant modelius geresniam sprendimų priėmimui, gerinant viešąją politiką, taip pat gerinant duomenų kokybę.

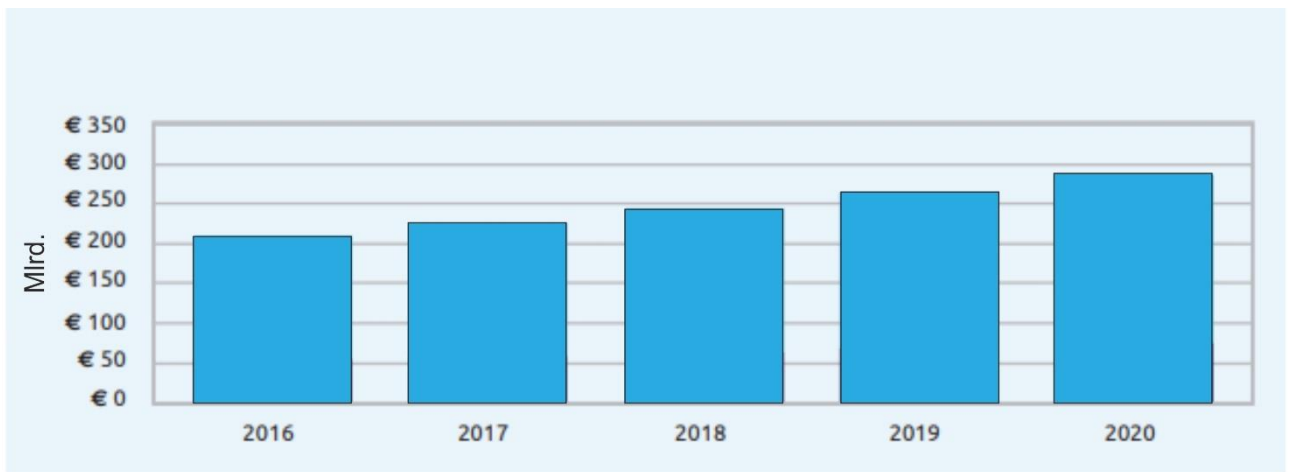
2017 metais pateiktoje Europos sąjungos duomenų rinkos ataskaitoje<sup>10</sup> nagrinėjama bendrovių, dirbančių ne tik su atvirai prieinamais duomenimis, bet su duomenimis bendrai. Ataskaitoje pateikta informacija apie įmonių skaičių, kurių pagrindinė veikla susijusi su duomenų apdirbimu ir panašių skaitmeninių produktų ir paslaugų teikimu. Tokių įmonių 2016 metais Europos sąjungoje buvo 254 850, tai yra 2,47 % visų Europos sąjungos įmonių. Tikimasi, kad iki 2020 tokių įmonių skaičius turėtų didėti nuo 1% iki 8,9%.

Atvirai prieinamų duomenų teikiama nauda išryškėja stebint atvirai prieinamų duomenų rinkos augimą, taip pat įvardinant pardavimo apimtis sugeneruotas naudojantis atvirai prieinamais duomenimis. 2016 metais D. W. Carrara apskaičiavo atvirai prieinamų duomenų rinkos dydį, kuris buvo tarp 193 mlrd. € ir 209 mlrd. € ir prognozuojamas augimas 2020 metais iki nuo 265 mlrd. € iki 286 mlrd. €.

---

<sup>9</sup> <http://www.asedie.es/assets/asedie-infomediary-sector-report-2017-.pdf> [žiūrėta 2019-01-01]

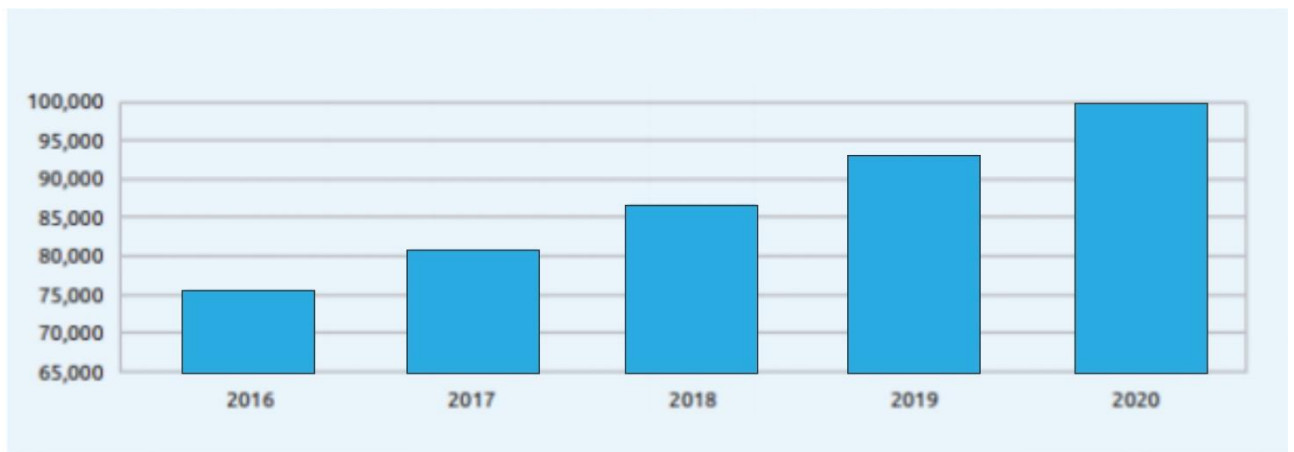
<sup>10</sup> [https://sites.google.com/a/open-evidence.com/download/repository/SMART20130063\\_Final%20Report\\_030417\\_2.pdf?attredirects=0&d=1](https://sites.google.com/a/open-evidence.com/download/repository/SMART20130063_Final%20Report_030417_2.pdf?attredirects=0&d=1) [žiūrėta 2019-01-01]



8 pav. Prognozuojamas atvirai prieinamų duomenų rinkos dydis Europos sąjungoje  
(D. W. Carrara ir kt., 2015)

Manoma, kad atvirai prieinami duomenys gali pagerinti ne tik viešųjų paslaugų kokybę, bet ir efektyvumą. Atvirai prieinami duomenys leidžia geriau paskirstyti išteklius, pavyzdžiui duomenų išleidimas, gali reikšti, kad organizacija ar jos klientų aptarnavimo centras gauna mažiau klientų užklausų, leidžiančių organizacijai sutelkti dėmesį ir išteklius savo pagrindinei verslo veiklai. Įmonėms, kurioms duomenys reikalingi pakartotiniam naudojimui, atvirai prieinami duomenys taip pat gali didinti efektyvumą, nes organizacijos gali per trumpesnę laiką atrasti reikiamus duomenis. Analizuojant atvirai prieinamų duomenų naudą buvo išskirti trys pagrindiniai vertinimo rodikliai. Kaip atvirai prieinami duomenys gali išgelbėti gyvybes? Kaip atvirai prieinami duomenys gali sutaupyti laiko? Kokia nauda aplinkai gali suteikti atvirai prieinami duomenys? Buvo įvertinta, kad atvirai prieinami duomenys gali išgelbėti 1425 gyvybes visoje Europoje kasmet, tai yra 5,5% visų mirtinų auto avarių, pasinaudojus atvirai prieinamais duomenimis potencialiai galima sutaupyti 629 mln. valandų praleistų eismo spūstyse. Žvelgiant iš aplinkos perspektyvos, buvo įvertinta, kad atvirai prieinami duomenys gali teikti naudą tiek aplinkai tiek namų ūkiams. Pateikiant komunalinių paslaugų suvartojimo duomenis, lyginant juos su panašių namų ūkių suvartojimu. Manoma, kad atvirai prieinamų duomenų panaudojimas leistų sutaupyti iki 16 % šiuo metu suvartojamos energijos.

Augant atvirai prieinamų duomenų rinkai kartu auga ir kompetentingų darbuotojų gebančių dirbti su atvirai prieinamais duomenimis. Remiantis Open data portal ataskaitos prognoze manoma, kad atvirai prieinami duomenys turėtų teikti naudą kuriant naujas darbo vietas. Įvertinta, kad nuo 2018 metų iki 2020 metų visoje Europos sąjungoje turėtų padidėti su atvirai prieinamais duomenimis dirbančių žmonių poreikis, iki 2020 turėtų būti dar apie 13 000 papildomų laisvų darbo vietų atvirai prieinamų duomenų specialistams.

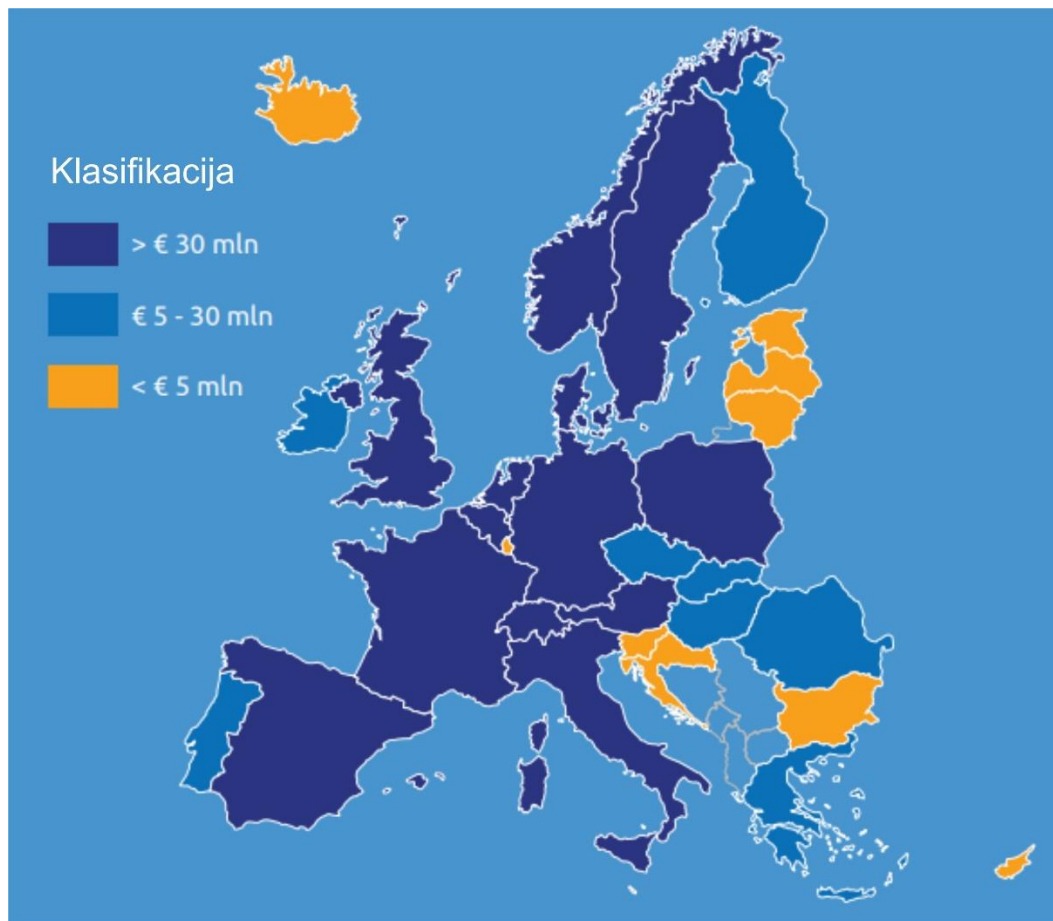


9 pav. Prognozuojamas darbo vietų skaičius darbui su atvirai prieinamais duomenimis Europos sąjungoje

(D. W. Carrara ir kt., 2015)

2017 metais buvo apskaičiuota kiek žmonių dirba su duomenimis Europos sąjungoje, tai tokie darbuotojai, kurie renka, saugo, tvarko ir analizuoja duomenis kaip savo pagrindinę arba ne pagrindinę užduotį darbe. Remiantis Europos komisijos 2017 metų duomenų rinkos ataskaita tokių darbuotojų visoje Europos sąjungoje yra apie 6,1 mln. ir manoma, kad toliau skaičius didės apie 3 % kasmet. Nepaisant įsitikinimo, kad darbas su duomenimis yra labiausiai siejamas su informacijos, komunikacijos ir technologijų sektoriumi, tačiau pagal pasiskirstymą šiame sektoriuje dirba tik 11 % visų Europos sąjungoje su duomenimis dirbančių darbuotojų, tuo tarpu didmeninėje ir mažmeninėje prekyboje dirba net 18 %, o gamyboje 12 % visų darbuotojų dirbančių su duomenimis. Šių darbo vietų pasiskirstymas rodo, kad ekonomika tampa vis labiau orientuota į duomenis, įskaitant tradicinius sektorius.

Kalbant apie atvirai prieinamų duomenų kuriamą vertę būtina paminėti, kad tai labai priklauso nuo sektoriaus, kuriame tikimasi naudoti. Tikima, kad atvirai prieinami duomenys turi didelį pakartotinio naudojimo potencialą naujose paslaugose ir kuriamose aplikacijose, taip pat sprendžiant visuomenės problemas, iššūkius, skatinant piliečių dalyvavimą politiniame ir socialiniame gyvenime, didinant skaidrumą ir atskaitomybę ir gerinant duomenų dalijimąsi tarp viešojo administravimo institucijų. Viską įvertinus, manoma, kad viešasis sektorius iki 2020 metų Europos sąjungos mastu turėtų gauti naudą už 22 mlrd. €. Žemiau pateiktame paveikslėlyje – valstybės klasifikuojamas pagal 2020 metų prognozę, kiek viešai prieinamų duomenų paskelbimas sumažins viešojo sektoriaus patiriamas išlaidas. Be to, viešosios valdžios institucijos skelbia atviruosius duomenis, kurie padeda skatinti naujoves kuriant naujas paslaugas ir verslo galimybes. Todėl nenuostabu, kad daugiausiai sutaupančios šalys taip pat pasirodo reitinguose, kuriose yra daugiausiai inovacijas kuriančių įmonių, arba tarp aukščiausiai vertinamų šalių Europos Komisijos ODINE programoje.



10 pav. 2020 metų prognozuojamas viešojo sektoriaus išlaidų sumažėjimas įtakotas atvirai prieinamų duomenų paskelbimo

(D. W. Carrara ir kt., 2015)

Atvirai prieinamų duomenų brandos lygmuo	Šalys narės	ODINE programos finansuotų įmonių skaičius
Diktuojantys madas	Austrija, Bulgarija, Suomija, Prancūzija, Airija, Nyderlandai, Ispanija, Didžioji Britanija	Didžioji Britanija (17) Nyderlandai (4) Ispanija (4) Austrija (3) Prancūzija (2) Airija (2) Suomija (1)
Pažengę	Estija, Graikija, Slovakija, Rumunija, Norvegija, Kroatija, Slovėnija, Liuksemburgas	Graikija (2) Slovėnija, (2) Slovakija (1) Estija (1) Rumunija (1)
Besivejantis	Belgija, Čekija, Kipras, Danija, Vokietija, Vengrija, Italija, Lietuva, Lenkija, Portugalija, Švedija, Šveicarija	Vokietija (10) Belgija (2) Italija (1) Čekija (1)
Pradedantieji	Lichtenšteinas, Latvija, Malta	Latvija (1)

2 lent. ODINE programos šalių reitingas iš ODINE 2017 m. ataskaitos

Apibendrinant atvirai prieinamų duomenų naudą matoma, kad atvirų duomenų teikiama nauda yra įvairi, nuo geresnio viešojo administravimo efektyvumo, ekonomikos augimo privačiame sektoriuje iki platesnės socialinės gerovės. Open data portal pateikia atvirų duomenų teikiamos naudos klasifikavimą. Nauda skirstoma į tris grupes, kiekviena grupė turi po dvi funkcijas.



11 pav. Atvirų duomenų teikiama nauda

(European data portal., 2017)

## **2. ATVIRAI PRIEINAMŲ DUOMENŲ PANAUDOJIMO ATVEJŲ TYRIMAS**

Šioje darbo dalyje pateikiama tyrimo metodika, kuria vadovaujantis buvo atliekamas tyrimas panaudojant surinktus duomenis apie atvirų duomenų rūšių panaudojimą pasaulyje ir Lietuvoje siekiant rasti asociacijų tarp naudojamų duomenų rūšių ir atvaizduoti poreikį naujų duomenų atvėrimui.

### **2.1 Tyrimo metodai ir darbo eiga**

Analitinė šio darbo dalis padalinta į du etapus, pirmame etape surenkami, suklasifikuojami ir pateikiami Lietuvos valstybinių įstaigų ir Lietuvos didmiesčių savivaldybių skelbiami atviri duomenų rinkiniai ir išanalizuojami pavyzdžiai kaip Lietuvoje ir pasaulyje veikiančios įmonės panaudoja viešai prieinamus duomenis, taip sukurdamos vartotojui patrauklų produktus, kurių pagalba klientams padedama priimti sprendimus, taip pat generuojamos pajamos ir kuriamos naujos darbo vietos. Antroje dalyje naudojantis duomenų gavybos metodika analizuojami atvirų duomenų panaudojimo atvejai ieškant asociacijų tarp panaudojamų duomenų rinkinių rūšių. Tyrimo metu buvo naudojami šie metodai: duomenų rinkimas, duomenų rinkinių rūšių klasifikavimas ir duomenų asociacijų paieška „Carma“ metodu.

1. Duomenų rinkinių rūšių klasifikavimas. Šioje tyrimo dalyje analizuojamos atvirų duomenų rūšys, apžvelgiama kitų šalių atvirų duomenų praktika. Taip pat surinkti ir suklasifikuoti Lietuvoje atvirai prieinami duomenų rinkiniai, bei nurodyti jų šaltiniai.
2. Duomenų rinkimas. Atliekant antrą tyrimo dalį analizuojami atvirų duomenų panaudojimo atvejai populiariausiuose ir sėkmingiausiuose atvirais duomenimis paremtuose projektuose Lietuvoje ir pasaulyje, surenkama informacija apie atvirais duomenimis paremtuose projektuose naudotas atvirų duomenų rinkinių rūšis.
3. Duomenų paruošimas. Surinkami duomenys apie atvirų duomenų rūšių panaudojimą sutvarkomi ir paruošiami apdorojimui SPSS Modeler programine įranga.
4. Asociacijų paieška. Naudojantis SPSS Modeler atliekamas atvirais duomenimis paremtų projektų, naudotų atvirų duomenų rūšių asociacijų tyrimas „Carma“ metodu, analizuojami ir aprašomi rezultatai.

### 3. ATVIRAI PRIEINAMŲ DUOMENŲ RINKINIŲ KLASIFIKAVIMAS

#### 3.1 Atvirai prieinamų duomenų rūšys

Pasaulyje nuolat susiduriama su daugybe iššūkių, tokių kaip lyčių nelygybė, valstybinių organizacijų skaidrumo trūkumas ar klimato kaita. Buvo pastebėta, kad siekiant valdyti ar mažinti įvardintas problemas duomenimis paremti sprendimai tiesiog būtini, o atviri duomenys yra raktas į tuos sprendimus. Todėl 2013 metais įkurta ne pelno siekianti organizacija „Open data watch“ siekianti pakeisti duomenis renkančių ir tvarkančių organizacijų įpročius. „Open data watch“ daug dėmesio skiria viso pasaulio, ypač blogiau išsivysčiusių, šalių nacionalinių statistikos tarnybų veiklos tobulinimui ir revoliucijos duomenų atskleidime skatinimu. Naudodamiesi duomenų parama ir stebėseną, „Open data watch“ daro įtaką ir padeda nacionalinėms statistikos organizacijoms ir kitoms organizacijoms įgyvendinti savo nacionalinių planų ir tvaraus vystymosi tikslus. Šios organizacijos tikslas - padaryti ilgalaikius pokyčius, siekiant pagerinti oficialios statistikos aprėptį ir atvirumą. 2015 metais „Open data watch“ siekdama skatinti duomenų atvirumą įkūrė ne pelno siekiančią organizaciją „ODIN“ (Open data inventory). „ODIN“ tai kasmetinis pasaulio šalių atvirų duomenų vertinimas ir šalių reitingavimas pagal duomenų atvirumą ir kokybę. „ODIN“ vertina šalių duomenis suklasifikuotus į 3 grupes, tai socialiniai duomenys, ekonominiai duomenys ir aplinkos duomenys, kurių kiekviena grupė turi 5-9 duomenų rūšis, kurios taip pat vertinamos atskirai ir įvertinus visų duomenų atvirumą ir duomenų gausą atliekamas bendras šalies vertinimas. 2018 metais „ODIN“ sureitingavo net 178 šalis, tarp kurių Lietuva yra 14 vietoje.

Atvirai prieinami duomenys yra labai skirtingai rūšiuojami, „Open knowledge international“ atvirus duomenis mato suskirstytus į 5 rūšis – kultūros, moksliniai, finansiniai, statistiniai, orų ir aplinkos duomenų rinkiniai.

2013 metais Laura James apibrėžė 8 duomenų rinkinių rūšis – geografiniai, kultūros, moksliniai, finansiniai, statistiniai, metrologiniai, aplinkos ir transporto.

Taip pat šalys savo atvirų duomenų dalinimosi platformose labai skirtingai rūšiuoja atvirai prieinamus duomenys. Palyginus kelias labiausiai išsiskiriančias šalis kaip JAV (vieną iš pirmaujančių ekonomikų pasaulyje), Belgiją (ES šalis turinti žemiausią ODIN reitingą), Singapūrą (ODIN lyderė), Lietuvą, Latviją ir Somalį (blogiausiai vertinama, duomenis atvėrusi šalis pasaulyje) pastebime aiškius skirtumus kaip šalys mato savo atvirus duomenis. JAV ir Belgija duomenis suskirstė į 14 rūšių, Lietuva į 13, o Singapūras tik į 9 rūšis. Iš kurių pagrindinės ir bendros visoms šalims yra finansai, švietimas, transportas, aplinka ir sveikata.

Taip pat pastebima, kad nepaisant labai smulkaus rūšiavimo ar duomenų gausos Belgija yra tik 73 vietoje „ODIN“ reitinge, JAV turėdama išpūdingą duomenų rinkinių sąrašą, viršijantį 200 000 duomenų rinkinių, yra tik 13 vietoje, tuo tarpu Singapūras duomenų rinkinius surūšiuojo į 9 rūšis ir paviešindamas tik 1690 duomenų rinkinių yra „ODIN“ reitingo lyderis, dėl savo duomenų kokybės ir atvirumo, pateikimo laiku ir aukšto panaudojamumo.

	ODIN reitingas (vieta 2018 m.)	Duomenų rūšys	Duomenų rinkinių kiekis
Singapūras	1	Ekonomika, švietimas, aplinka, finansai, sveikata, infrastruktūra, visuomenė, technologijos, transportas (9)	1690
JAV	13	Žemės ūkis, klimatas, vartojimas, ekosistema, švietimas, energija, finansai, sveikata, vietos valdžia, gamyba, uostai, vandenynai, viešoji sauga, moksliniai tyrimai (14)	Daugiau nei 214 000
Lietuva	14	Aplinka, ekonomika ir finansai, kultūra, poilsis ir religija, socialinė apsauga, sveikatos apsauga, švietimas ir mokslas, tarptautiniai santykiai, teisėsauga ir teisėtvara, transportas ir ryšiai, ūkinė veikla ir verslas, valstybės valdymas ir viešasis administravimas, žemės ūkis (13)	745
Latvija	34	Regionai ir savivaldybės, gyventojų ir socialiniai, ekonomika ir verslas, valstybės valdymas, žemės ūkis, sveikatos apsauga, aplinka, kultūra, teisingumas ir vidaus reikalai, transportas (10)	334
Belgija	73	Žemės ūkis ir žuvininkystė, kultūra ir sportas, ekonomika ir finansai, švietimas, energija, aplinka, sveikata, tarptautiniai duomenys, teismai, populiacija, viešasis sektorius, regionai, mokslas ir technologijos, transportas. (14)	Daugiau nei 10 000
Somalis	177	Nerūšiuoja	11

3 lent. Atvirų duomenų rinkinių rūšiavimas. Sudaryta darbo autoriaus remiantis vyriausybinėmis atvirų duomenų internetinėmis platformomis 2018-04-16



## 3.2 Atvirai prienami duomenys ir jų rūšys Lietuvoje

Nuo tos dienos kai 2011 metų gruodį Europos komisijos pranešimas buvo pradėtas fraze „Atvirai duomenys tai variklis skatinantis inovacijas, augimą ir vyriausybės skaidrumą“ Lietuva atvirų duomenų srityje padarė didelį progresą. Šiandien Lietuva, pasaulio mastu, yra gana gerai vertinama „ODIN“ reitinge ir vidutiniokė tarp išsivysčiusių šalių, Lietuvoje nuolat rengiamos atvirų duomenų skatinimo iniciatyvos, kuriami atskleistų duomenų rinkinių plėtros planai ir keliami tikslai, todėl galima manyti, kad Lietuvos atvirų duomenų kiekis didės ir kokybė nuolat gerės.

Šiuo metu Lietuvoje yra sukurta atvirų duomenų platforma, kurios informacijos sąrašo valdytojas ir tvarkytojas yra Informacinės visuomenės plėtros komitetas prie Susisiekimo ministerijos. Pagrindiniame Lietuvos atvirų duomenų rinkmenų šaltinyje <http://www.opendata.gov.lt/> 2019 metų balandžio mėnesį Lietuvos informacinių rinkmenų sąrašė yra surinkta 745 duomenų rinkmenos, kurios yra skirstomos į 13 veiklos sričių ir duomenys, priklausomai nuo duomenų rinkinio, yra prieinami daugiau nei 10 formatų, tarp kurių CSV, XLS ar XML formatais.

	<b>Duomenų rinkinių skaičius</b>	<b>Atvertų duomenų rinkinių tipai</b>
Aplinka	55	Nekilojamo turto, vandens telkinių registrai, oro taršos, aplinkos duomenys, atliekų tvarkymo duomenys, miškų duomenų rinkiniai ir kt.
Ekonomika ir finansai	94	Įvairūs draudimo duomenų rinkiniai, savivaldybių veiklos ataskaitos, muitinės įstaigų duomenys ir kt.
Kultūra	20	Kultūros vertybių registras, Lietuvos periodinės spaudos straipsnių bazė, nacionalinės bibliografijos duomenų bankas ir kt.
Poilsis ir religija	2	Lietuvos sporto statistikos leidiniai, tarifai gamintojams, naudojantiems atsinaujinančius išteklius
Socialinė apsauga	54	Laisvų darbo vietų duomenys, gyventojų užimtumo tyrimo duomenys, namų ūkių biudžeto tyrimų duomenys ir kt.
Sveikatos apsauga	111	Gyventojų sveikatos tyrimo duomenys, statistiniai mirties priežasčių duomenys, sergamumo pagal regionus duomenys ir kt.

Švietimas ir mokslas	15	Studijų programų vertinimų duomenys, vertintų studijų krypčių apžvalgos ar KGB veiklos dokumentai ir kt.
Tarptautiniai santykiai	3	Nacionalinio CERT-LT užregistruotų incidentų statistika, Užsienio reikalų ministerijos Konsulinio departamento konsulinės informacijos keliaujantiems ir gyvenantiems užsienyje tinklapis „keliauk.urm.lt“ ir dvi nuorodos į interneto svetaines, susijusias su pagalba keliaujant ir vizomis.
Teisėsauga ir teisėtvara	31	Ginklų , sutarčių, įgaliojimų ar testamentų registrai ir kt.
Transportas ir ryšiai	89	Laivų registras, viršutinės kelio dangos duomenys, įvairių, tokių kaip gyventojų pasitikėjimo interneto saugumu, gyventojų naudojimosi internetu tyrimų duomenys ir kt.
Ūkinė veikla ir verslas	90	Auditorių sąrašas, licencijuotų ginklų gamintojų sąrašas, juridinių asmenų registras, fizinių asmenų, kuriems draudžiama eiti vadovo pareigas, sąrašas ir kt.
Valstybės valdymas ir viešasis administravimas	151	Elektros energijos gamintojų sąrašas, įvairių komunalinių paslaugų ir kuro tiekėjų sąrašai ir kainos, bankroto ir restruktūrizavimo procesų duomenys ir kt.
Žemės ūkis	51	Ūkininkų ūkių registras, įvairių ūkio šakų paramos ir veiklos statistiniai duomenys ir kt.

4 lent. Lietuvos atvirų duomenų rinkmenų sąrašas sudarytas darbo autoriaus, remiantis

[www.opendata.gov.lt](http://www.opendata.gov.lt) 2018-04-18

Informacinės visuomenės plėtros komiteto valdoma Lietuvos atvirų duomenų platforma yra pagrindinis ir didžiausias atvirų duomenų šaltinis surinktas vienoje vietoje, tačiau tai nėra vienintelis priėjimas prie atvirų duomenų rinkinių. Savo ruožtu duomenis viešina ir įvairius valstybinės institucijos, tokios kaip „Sodra“, „Regitra“, daugelio miestų savivaldybės, Lietuvos valstybės vyriausioji rinkimų komisija, Lietuvos automobilių kelių direkcija ir kt.

Duomenimis dalinasi ir didžiausi duomenų tiekėjai, kaip Vilniaus savivaldybė, kuri pateikia daug duomenų įvairiomis temomis, nuo transporto spūsčių ar medžių būklės iki mažų duomenų valdytojų kaip VĮ Vidaus vandenų kelio direkcijos, kuri pateikia duomenis apie vidaus vandenų gylius.

Klasifikuojant duomenų rinkinius prieinamus per pačių įstaigų internetines svetaines Lietuvoje, remiantis Lauros James (2013) atvirų duomenų rinkinių klasifikavimo modeliu galime suskirstyti visus atvirai prieinamus duomenis į 8 duomenų rinkinių rūšis.

Duomenų valdytojas	Viešinami duomenų rinkiniai	Nuoroda į atvirus duomenis
<b>Geografiniai duomenys</b>		
Kauno savivaldybė	Kauno miesto ribos, laisvalaikio zonos, parkavimo zonų žemėlapis ir kt.	<a href="https://github.com/Kaunas">https://github.com/Kaunas</a>
Nacionalinė žemės tarnyba	Skirtingų mastelių Lietuvos teritorijos rastriniai žemėlapiai	<a href="http://www.nzt.lt/go.php/lit/Atviri-duomenys">http://www.nzt.lt/go.php/lit/Atviri-duomenys</a>
Vidaus vandenų kelių direkcija	Vidaus vandenų gilių duomenys	<a href="http://vvkd.lt/atviri-duomenys/">http://vvkd.lt/atviri-duomenys/</a>
Informatikos ir ryšių departamentas	Lietuvos nusikaltimų žemėlapis	<a href="https://www.ird.lt/nvzrgis/map/">https://www.ird.lt/nvzrgis/map/</a>
VĮ „Sodra“	Vidutinių darbo užmokesčių, socialinių išmokų ir kt. pasiskirstymo pagal savivaldybes žemėlapiai	<a href="http://atvira.sodra.lt/lt-eur/">http://atvira.sodra.lt/lt-eur/</a>
GeoData: Lietuvos geoduomenys	Įvairūs duomenys atvaizduoti geografiškai	<a href="http://www.geodata.lt/">http://www.geodata.lt/</a>
<b>Finansiniai duomenys</b>		
Viešųjų pirkimų tarnyba	Duomenys apie viešuosius pirkimus, viešųjų pirkimų įmones laimėtojas ir kt.	<a href="http://www.freedata.lt/vpt/search">http://www.freedata.lt/vpt/search</a>
VĮ „Sodra“	VSDF biudžeto pajamų ir išlaidų struktūra, Fondo administravimo įstaigų veiklos sąnaudų duomenys ir kt.	<a href="http://atvira.sodra.lt/lt-eur/">http://atvira.sodra.lt/lt-eur/</a>
VĮ „Valdymo koordinavimo centras“	Valstybės valdomų finansiniai duomenys ir kt.	<a href="https://vkc.sipa.lt/apie-imonas/atviri-duomenys/">https://vkc.sipa.lt/apie-imonas/atviri-duomenys/</a>
<b>Statistiniai duomenys</b>		
VĮ „Turto bankas“	Valstybės nekilnojamo turto aukcionų duomenys, duomenys apie parduotą privatizuotą NT	<a href="https://www.turtas.lt/lt/administracine-informacija/atviri-duomenys/">https://www.turtas.lt/lt/administracine-informacija/atviri-duomenys/</a>

Ryšių reguliavimo tarnyba	Belaidžio interneto prieigos paslaugų, teikiamų įvairiuose Lietuvos regionuose, kokybės matavimo duomenys	<a href="https://www.rrt.lt/istekliai/atviri-duomenys/">https://www.rrt.lt/istekliai/atviri-duomenys/</a>
Audito, apskaitos, turto vertinimo ir nemokumo valdymo tarnyba	Tyčinių bankrotų sąrašas, fizinių asmenų bankroto duomenis ir kt.	<a href="http://www.avnt.lt/veiklos-sritys/nemokumas-2/">http://www.avnt.lt/veiklos-sritys/nemokumas-2/</a>
Valstybės tarnybos departamentas	Valstybės tarnautojų duomenys suskirstyti pagal savivaldybes, kategorijas, pareigas ir kt.	<a href="http://statistika.vtd.lt/index.html">http://statistika.vtd.lt/index.html</a>
Lietuvos statistikos departamentas	Gyventojų užimtumo, pajamų, sveikatos ir pan. tyrimų atviri duomenys	<a href="https://osp.stat.gov.lt/viesos-duomeni-rinkmenos">https://osp.stat.gov.lt/viesos-duomeni-rinkmenos</a>
Lietuvos respublikos seimas	Įvairūs duomenys apie seimo narius, posėdžius, komandiruotes ir kt.	<a href="https://www.lrs.lt/sip/portal.show?p_r=15818&amp;p_k=1">https://www.lrs.lt/sip/portal.show?p_r=15818&amp;p_k=1</a>
Valstybinė mokesčių inspekcija	Lietuvos PVM mokėtojų duomenys, kasos aparatų registras ir kt.	<a href="https://www.vmi.lt/cms/atviri-duomenys">https://www.vmi.lt/cms/atviri-duomenys</a>
Lietuvos policija	Eismo įvykių duomenys	<a href="https://www.epolicija.lt/atviri-duomenys">https://www.epolicija.lt/atviri-duomenys</a>
VĮ „Sodra“	Draudėjų, išmokų, pensijų, nedarbingumo pažymėjimų ir kt. duomenys	<a href="http://atvira.sodra.lt/lt-eur/">http://atvira.sodra.lt/lt-eur/</a>
Higienos institutas	Sveikatos įstaigų metinių ataskaitų suvestinės, Lietuvos sveikatos statistika, ligų, mirčių ir kt. duomenys	<a href="http://www.hi.lt/atviri-duomenys.html">http://www.hi.lt/atviri-duomenys.html</a>
Lietuvos Respublikos vyriausioji rinkimų komisija	Rinkimų balsavimų, reitingavimų, balsavusių asmenų ir kt. duomenys	<a href="https://www.vrk.lt/atviri-duomenys">https://www.vrk.lt/atviri-duomenys</a>
Lietuvos Respublikos konkurencijos taryba	Konkurencijos tarybos paskirtų baudų, sumokėtų baudų, diskvalifikavimų iš viešųjų pirkimų ir kt. duomenų rinkiniai	<a href="http://kt.gov.lt/lt/atviri-duomenys">http://kt.gov.lt/lt/atviri-duomenys</a>
VĮ „Valdymo koordinavimo centras“	Valstybės valdomų įmonių sąrašai, valdybos narių sąrašai ir kt.	<a href="https://vkc.sipa.lt/apie-imonas/atviri-duomenys/">https://vkc.sipa.lt/apie-imonas/atviri-duomenys/</a>

<b>Metrologiniai duomenys</b>		
Lietuvos metrologijos inspekcija	Įstaigų, paskirtų atlikti įvairių įrenginių patikras sąrašai, pažymų, įstaigų vadovų ir kt. sąrašai	<a href="http://metrinsp.lrv.lt/lt/atviri-duomenys">http://metrinsp.lrv.lt/lt/atviri-duomenys</a>
<b>Kultūros duomenys</b>		
GeoData: Lietuvos geoduomenys	Bažnyčių, muziejų ir piliakalnių žemėlapiai	<a href="http://www.geodata.lt/category/kultura/">http://www.geodata.lt/category/kultura/</a>
<b>Aplinkos duomenys</b>		
Vidaus vandenų kelių direkcija	Vidaus vandenų gilių duomenys	<a href="http://vvkd.lt/atviri-duomenys/">http://vvkd.lt/atviri-duomenys/</a>
Vilniaus miesto savivaldybė	Vilniaus šalinamų medžių sąrašas, medžių būklės mieste sąrašas, Vilniaus miesto tyliosios gamtos, viešosios zonos	<a href="https://data-vplanas.opendata.arcgis.com/search?collection=Dataset">https://data-vplanas.opendata.arcgis.com/search?collection=Dataset</a>
GeoData: Lietuvos geoduomenys	Miškų kadastro duomenys, saugomų teritorijų žemėlapis, potvynių grėsmės ir rizikos žemėlapis	<a href="http://www.geodata.lt/category/aplinka/">http://www.geodata.lt/category/aplinka/</a>
Aplinkos apsaugos agentūra	Kasdieniai oro kokybės duomenys, oro taršos ataskaitos	<a href="http://oras.gamta.lt/cms/index">http://oras.gamta.lt/cms/index</a>
<b>Transporto duomenys</b>		
Lietuvos policija	Eismo įvykiu duomenys 2013-2017	<a href="https://www.epolicija.lt/atviri-duomenys">https://www.epolicija.lt/atviri-duomenys</a>
Lietuvos automobilių kelių direkcija	Eismo intensyvumo, kelio dangos ir kelio darbų interaktyvus žemėlapis	<a href="https://eismoinfo.lt/#/">https://eismoinfo.lt/#/</a>
Vilniaus savivaldybė	Vilniaus miesto viešojo transporto keleivių srautų tyrimas	<a href="http://sisp.maps.arcgis.com/apps/MapJournal/index.html?appid=5f7aaa93111d4e82af2c70409e586897#">http://sisp.maps.arcgis.com/apps/MapJournal/index.html?appid=5f7aaa93111d4e82af2c70409e586897#</a>
Vilniaus miesto savivaldybė	Vilniaus miesto viešojo transporto stotelės, Gatvių ir privažiavimų ašinės linijos, dviračių trasų sąrašai ir kt.	<a href="https://data-vplanas.opendata.arcgis.com/search?q=transportas">https://data-vplanas.opendata.arcgis.com/search?q=transportas</a>
Vilniaus miesto savivaldybė	Vilniaus grūsčių duomenys	<a href="https://atviras.vilnius.lt/eismo-situacijos">https://atviras.vilnius.lt/eismo-situacijos</a>

Lietuvos automobilių kelių direkcija	Greičių matuoklių, greitį mažinančių kalnelių, atitvarų, avaringų ruožų ir kt. sąrašai	<a href="http://lakd-lakd.opendata.arcgis.com/datasets">http://lakd-lakd.opendata.arcgis.com/datasets</a>
Lietuvos automobilių kelių direkcija	Vidutinio metinio paros eismo intensyvumo valstybinės reikšmės keliuose duomenys ir Valstybinės reikšmės kelių viršutinės kelio dangos duomenys	<a href="http://lakd.lrv.lt/lt/atviri-duomenys">http://lakd.lrv.lt/lt/atviri-duomenys</a>
VĮ „Regitra“	Visų šiuo metu registruotų automobilių sąrašas, daugybė statistinių duomenų pagal markę, metus ir kt.	<a href="https://www.regitra.lt/lt/atviri-duomenys/">https://www.regitra.lt/lt/atviri-duomenys/</a>
SĮ „Susisiekimo paslaugos“	Vilniaus keleivių srautų tyrimų rezultatai	<a href="http://www.vilniustransport.lt/lt/pages/view/?id=194">http://www.vilniustransport.lt/lt/pages/view/?id=194</a>
Klaipėdos miesto savivaldybė	Įlipusių keleivių skaičius stotelėse pagal el. bilietų žymėjimą ir viso viešojo transporto judėjimas žemėlapyje realiu laiku	<a href="http://opendata.klaipeda.lt/klaipedos-transporto-duomenys/">http://opendata.klaipeda.lt/klaipedos-transporto-duomenys/</a>

5 lent. Lietuvos atvirų duomenų, kurių nėra [www.opendata.gov.lt](http://www.opendata.gov.lt), rinkinių sąrašas sudarytas darbo autoriaus 2018-04-23

Atvirų duomenų iniciatyva Lietuvoje stipriai išsibėgėja. 2019 metais Lietuvos miestų savivaldybės, valstybinės įmonės ir įstaigos yra pavišinusios daugiau nei 1000 duomenų rinkinių, tačiau to neužtenka. 2016 metais Lietuvoje vykusioje tarptautinėje mokslinėje konferencijose pranešimą skaitęs Nyderlandų profesorius dr. Michiel de Vries (2016) išsakė savo nuomonę, kad vien duomenų paskelbimas ir tūkstančio Excel lentelių su duomenimis pavišėjimas neduoda jokios naudos, tai tėra formalumas. Duomenys turi būti pateikti su kontekstu, su ekspertų komentarais ir išaiškinimais. Informacija turi būti ne tik sukelta ir pavišinta, bet turi būti lengvai skaitoma ir suprantama, kad žmonės galėtų duomenimis pasinaudoti, juos išanalizavę gauti tam tikrą informaciją padėsiančią priimti sprendimus. Galima pagirti gerus pavyzdžius, kaip VĮ „Regitra“, kuri pateikia duomenų rinkinius ir papildomą Excel failą su visais trumpinių paaiškinimais arba Lietuvos statistikos departamento viešinami duomenų rinkiniai yra prieinami (ZIP) formatu, pastarieji komplektuojami metiniais rinkiniais, kuriuos sudaro tyrimo rezultatai ketvirčiais ir papildomas informacinis failas su visų trumpinių reikšmėmis ir galimais kintamaisiais, taip sudaromos geros sąlygos duomenis analizuojančiam asmeniui lengvai skaityti ir apdoroti atvirai prieinamus duomenis. Tačiau yra nemažai ir blogų bandymų prisijungti prie atvirų duomenų iniciatyvos.

Pavyzdžiui Vyriausioji tarnybinės etikos komisija savo internetinės svetainės atvirų duomenų skiltyje yra pavišiusi 2012-2017 metų veiklos ataskaitas 2019 metais, dalis kitų organizacijų turi pavišiusę vieną ar labai mažai duomenų rinkinių, dažnu atveju jie būna nebeaktualūs, kadangi yra mažiausiai dviejų metų senumo. Nemažiau svarbu paminėti, kad dalis atvirai prieinamų duomenų rinkinių esančių didžiausiame Lietuvos atvirų duomenų portale <http://opendata.gov.lt> visgi nėra atvirai prieinami. Internetinėje svetainėje nurodomas pasirinktos rinkmenos tvarkytojas ir kontaktiniai asmenys, tačiau nei skambučio, ne parašyto prašymo pagalba neapdoroti atviri duomenys nėra prieinami. 2016 metais Prof. dr. Tomo Kirilavičiaus paminėta problema vis dar aktuali. Profesorius teigė, kad atviri duomenys visų pirma garantuoja viešojo sektoriaus skaidrumą ir atskaitomybę, tačiau vien duomenų atvėrimas tokio rezultato neduoda, problema, kad įstaigos atverdamos duomenis ne retai juos stipriai agreguoja, taip paversdamos viešinamus duomenis praktiškai netinkamais panaudoti. Lietuvoje galima rasti ne vieną stipriai agreguotą atvirų duomenų rinkinį. Vienas iš atvejų - atviro Vilniaus savivaldybės internetinėje svetainėje galima rasti miesto savivaldybės pateiktus Vilniaus spūsčių duomenis, tačiau duomenys yra tik apie 10 labiausiai užsikišusių sankryžų ir duomenys apie užfiksuotas eismo situacijas per paskutines 30 dienų, duomenų atsisiųsti jokių formatu galimybės nėra.

Informacinės visuomenės plėtros komitetas 2015 metais pateikė ataskaitą pavadinimu: „Galimybių tyrimo, skirto apibrėžti atvirų duomenų iniciatyvos įgyvendinimo Lietuvoje architektūros modelį, ataskaita“, joje naudojantis atvirų duomenų 5 žvaigždučių brandos lygio schema, kurią 2010 metais pristatė Tim Berners Lee, ji aptarta pirmoje darbo dalyje. Informacinės visuomenės plėtros komiteto pasiūlytas modelis sieks viešinamus atvirų duomenų kokybę pakelti iki 5 žvaigždučių brandos lygio, tačiau 2015 metų ataskaitos išanalizavus atvirų duomenų rinkinių kokybę buvo nustatyta, kad tik 4 % duomenų rinkinių pasiekė trečia brandos lygį, t.y. duomenys skelbiami struktūrizuotų duomenų forma ir atvirais formatais, tokiais kaip CSV, 23 % duomenų rinkinių yra antro brandos lygmens, t.y. duomenys skelbiami struktūrizuota duomenų forma, bet nepatogiais formatais, tokiais kaip Excel ir net 73 % duomenų rinkinių yra žemiausio, pirmo brandos lygmens, t.y. duomenys skelbiami nestruktūrizuotais, neredaguojamais formatais, kartais duomenys būna tiesiog skanuoti. Nemaža dalis duomenų institucijų puslapiuose yra skelbiama HTML formatu be galimybės juos atsisiųsti, taip sumažinama galimybė duomenis toliau panaudoti. Ataskaitoje taip pat išskiriamas duomenų rinkinių formatų integralumo trūkumas. Didelė dalis giminingų duomenų rinkinių yra pateikiami už skirtingus laikotarpius arba tos pačios rūšies duomenys už skirtingus laikotarpius yra pateikiami skirtingais formatais, pvz. duomenų rinkinys už 2014 metus pateikiamas XML formatu, o už 2015 metus – PDF formatu. Toks duomenų pateikimas taip pat trukdo duomenis patogiai panaudoti, taip mažinama nauda verslui ir visuomenei.

### 3.3 Duomenų panaudojimo atvejų pasaulyje klasifikavimas

Atsirandant vis daugiau galimybių panaudoti atvirus duomenis ir daugėjant sričių ir duomenų rinkinių, atvirų duomenų panaudojimas tampa vis dažnesnis tiek Lietuvoje, tiek pasaulyje. Siekiant išanalizuoti atvirų duomenų panaudojimą būtina pažvelgti į panaudos atvejus giliau. Žemiau esančiose lentelėse pateikti atvirai prieinamų duomenų panaudojimo atvejai, jų tikslai, šalis bei projektui naudojamų duomenų rūšys.

<b>Žemės ūkis</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
M-Farm	Padėti ūkininkams priimti pagrįstus sprendimus dėl to, kaip gaminti prekes ir kur parduoti už didžiausią pelną	Kenija	Geografiniai, moksliniai ir švietimo, statistiniai, metrologiniai ir aplinkos
Mandi Trades	Didinti žemės ūkio rinkos efektyvumą ir mažina smulkiųjų ūkininkų informacijos asimetriją	Indija	Statistiniai
Rainbowagri	Padėti ūkininkams tiksliau įvertinti užauginto derliaus kainas	Indija	Geografiniai, statistiniai, aplinkos ir metrologiniai
Rural E-Market	Teikti informaciją ūkininkams ir palengvinti vartotojų ir gamintojų mainus	Madagaskaras	Geografiniai ir aplinkos

6 lent. Atvirų duomenų panaudojimas žemės ūkio srityje. Sudaryta darbo autoriaus



<b>Kultūra ir turizmas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Cultural routes	Sumažinti laiką ieškant turistinės informacijos	Rusija	Geografiniai, kultūros, aplinkos ir transporto
Michael Culture Association	Apjungti ir skatinti skaitmeninį kultūrinį turinį iš muziejų ir bibliotekų visoje Europoje	Prancūzija	Kultūros

7 lent. Atvirų duomenų panaudojimas kultūros ir turizmo srityje. Sudaryta darbo autoriaus

<b>Verslas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
District Metrics	Pagerinti rinkos tyrimus, remiantis pajamomis ir verslo veiklą, naudojantis viešaisiais duomenimis	Indija	Finansiniai ir statistiniai
New Zealand Institute of Economic Research (NZIER)	Palaikyti duomenų bazę, kuri suskirsto nacionalinę ekonominę analizę į patogius duomenų rinkinius, lengvai skaitomus vartotojams	Naujoji Zelandija	Finansiniai ir statistiniai
Rekvizitai	Teikti informacija apie įmones registruotas Lietuvoje	Lietuva	Finansiniai ir statistiniai
SumoBrain	Padėti lengviau rasti patentus, perduoti duomenis ir atnaujinti įspėjimus apie patentų duomenis	Didžioji Britanija	Statistiniai
Young Innovations	Duomenų pagalba skatinti šalies verslo plėtrą	Nepalas	Moksliniai ir švietimo, statistiniai ir finansiniai

8 lent. Atvirų duomenų panaudojimas verslo srityje. Sudaryta darbo autoriaus

<b>Švietimas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Afla	Didinti švietimo išlaidų skaidrumą	Moldova	Geografiniai, moksliniai ir švietimo, finansiniai ir statistiniai
Eduweb	Teikti paprastą interaktyvų mechanizmą, skirtą mokykloms identifikuoti, rasti ir įvertinti	Kenija	Geografiniai, moksliniai ir švietimo
Enova	Sudaryti sąlygas papildomam švietimui naudojantis atvirais duomenimis	Meksika	Geografiniai, moksliniai ir švietimo, finansiniai, statistiniai ir transporto
Map Kibera	Teikti informaciją apie švietimo įstaigas Kenijoje	Kenija	Statistiniai, moksliniai ir švietimo
Nigeria Open Data Access	Didinti visuomenės prieinamumą prie atviros informacijos	Nigerija	Geografiniai, moksliniai ir švietimo, finansiniai, statistiniai, aplinkos ir transporto
QEdu	Teikti patikimus duomenis švietimo vadovams, kad jie galėtų planuoti ir tobulinti švietimo sistemas	Brazilija	Statistiniai, moksliniai ir švietimo
TuvaLabs	Skatinti atvirų duomenų naudojimą ir mokyti jais pasinaudoti	JAV	Geografiniai, moksliniai ir švietimo, statistiniai ir aplinkos
Wuni Zaligu Development Association	Moksliniai tyrimai ir propagavimas švietimo ir aplinkosaugos klausimais	Gana	Moksliniai ir švietimo, statistiniai ir aplinkos

9 lent. Atvirų duomenų panaudojimas švietimo srityje. Sudaryta darbo autoriaus

<b>Energija ir klimatas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Green patrol	Viešinti aplinkosaugos pažeidimų duomenis, taip siekiant švaresnės šalies	Rusija	Geografiniai, statistiniai ir aplinkos
Haezoom	Siūlo bendrą standartą saulės kolektorių pramonei ir mažina asimetrinę informaciją tarp mažų saulės kolektorių tarpininkų ir vartotojų	Pietų Korėja	Moksliniai ir švietimo, statistiniai ir aplinkos
Meteo	Teikti informaciją apie orus šalyje	Lietuva	Geografiniai, moksliniai ir švietimo, statistiniai, metrologiniai ir aplinkos
Open oil	Sumažinti naftos pramonės neaiškumą ir taip didinti naftos bendrovių atskaitomybę	Vokietija	Finansiniai ir aplinkos
Päästöt	Interaktyviame žemėlapyje atvaizduoti skleidžiamos taršos informaciją	Suomija	Geografiniai ir aplinkos
ReConnect Energy	Mažinti energijos vartojimą ir skatinti atsinaujinančią energiją	Indija	Geografiniai, statistiniai, metrologiniai ir aplinkos
The Energy and Resources Institute (TERI)	Didinti informuotumą ir teikti politikos rekomendacijas aplinkos politikos klausimais	Indija	Geografiniai ir aplinkos
Upande	Duomenų pagrindu atliekami tyrimai, vizualizacijos kiti ir techniniai sprendimai siekiant gerinti vandens kokybę	Kenija	Geografiniai, metrologiniai ir aplinkos

10 lent. Atvirų duomenų panaudojimas energijos ir klimato srityje. Sudaryta darbo autoriaus

<b>Finansai ir draudimas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Credit reform	Teikti informaciją apie privačių ir juridinių asmenų mokumą ir finansinę būklę ir rizikingumą	Lietuva	Finansiniai ir statistiniai
Creditinfo	Teikti informaciją apie privačių ir juridinių asmenų mokumą ir finansinę būklę ir rizikingumą	Lietuva	Finansiniai ir statistiniai
IAG Ltd	Nustatyti sritis, kurios yra labiausiai pažeidžiamos stichinių nelaimių atveju.	Australija	Geografiniai, statistiniai ir metrologiniai
Thinknum	Pateikia duomenis ir įžvalgas investuotojams ir įmonėms	JAV	Geografiniai, statistiniai, finansiniai, metrologiniai ir aplinkos

11 lent. Atvirų duomenų panaudojimas finansų ir draudimo srityje. Sudaryta darbo autoriaus

<b>Valstybės valdymas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Budget stories	Teikti politinių sluoksnių analizę ir priemones, skirtas Moldovos biudžetui ir politikai	Moldova	Finansiniai, statistiniai ir aplinkos
Centre for Governance and Public Accountability (CGPA)	Informuoti piliečius apie svarbiausius valdymo klausimus ir stiprina valstybės reagavimą į piliečių skundus	Pakistanas	Moksliniai ir švietimo, finansiniai ir statistiniai
Code for Ghana	Siekti skaidrumo, atvirų duomenų ir piliečių dalyvavimo dėka vietos valdžios sistemos gerinimo..	Gana	Moksliniai ir švietimo, finansiniai ir statistiniai

De Olho nas Metas	Padėti pilietinei visuomenei stebėti savivaldybių projektus savo bendruomenėje	Brazilija	Geografiniai, moksliniai ir švietimo, finansiniai ir statistiniai
Economics Association of Zambia	Teikti rekomendacijas ir propaguoti ekonominę politiką, kuria siekiama sumažinti skurdą Zambijoje	Zambija	Moksliniai ir švietimo, finansiniai ir statistiniai
Janaagraha Centre for Citizenship and Democracy	Padėti vietos valdžios institucijoms atverti savo duomenis, o piliečiams geriau suprasti savivaldybių biudžetus	Indija	Finansiniai ir statistiniai
Democratic Elections	Skatinti rinkimų skaidrumą ir reformas Libane	Libanas	Geografiniai ir statistiniai
Odekro	Skatinti vyriausybės atskaitomybę ir piliečių dalyvavimą, suteikiant galimybę visuomenei susipažinti su parlamentiniais procesais	Gana	Statistiniai
Open budget	Skatinti gyventojus suprasti biudžeto informaciją per vizualizaciją, skaidrus valstybės valdymas	Ukraina	Finansiniai
Orodata	Skatinti gyventojų įsitraukimą ir valstybės problemų supratimą	Nigerija	Finansiniai ir statistiniai
Pulse Lab Jakarta	Analizuoti ir teikti informaciją apie nacionalinius vystymosi prioritetus	Indonezija	Finansiniai, statistiniai, aplinkos ir transporto
Rinkėjo puslapis	Skleisti su rinkimais susijusią informaciją	Lietuva	Geografiniai ir statistiniai

12 lent. Atvirų duomenų panaudojimas valstybės valdymo srityje. Sudaryta darbo autoriaus

<b>Sveikata</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Aidin	Sumažinti asimetrinę informaciją tarp pacientų ir poūminės priežiūros teikėjų	JAV	Statistiniai, moksliniai ir švietimo
Data Uruguay	Interaktyvios vizualizacijos pagalba padidinti piliečių prieigą prie atvirų vyriausybės sveikatos duomenų	Urugvajus	Statistiniai, moksliniai ir švietimo
MapMyGenome	Siūlo individualizuotus sveikatos sprendimus, pagrįstus genetiniais tyrimais	Indija	Statistiniai, moksliniai ir švietimo
Medicat	Padėti pacientams priimti teisingesnius sprendimus dėl medikamentų vartojimo	Prancūzija	Statistiniai, moksliniai ir švietimo
Yeleman	Duomenų pagalba nustatyti ir sumažinti maliarijos protrūkius	Malis	Statistiniai ir geografiniai

13 lent. Atvirų duomenų panaudojimas sveikatos srityje. Sudaryta darbo autoriaus

<b>Statybos ir nekilnojamas turtas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Citynow	Informuoti apie vykdomus ir planuojamus statybų projektus	Lietuva	Geografiniai ir statistiniai
City scale	Suteikti miesto piliečiams atitinkamą statistiką apie savo miestą.	Ukraina	Geografiniai, kultūros, moksliniai ir švietimo, finansiniai, statistiniai, metrologiniai, aplinkos ir transporto
Housing social	Pagerinti Indijos nekilnojamojo turto rinkos efektyvumą, suteikiant daugiau informacijos	Indija	Geografiniai, finansiniai, statistiniai ir transporto

Kur gyvenu?	Teikti informaciją apie NT objektą ir kas yra netoli jo	Lietuva	Geografiniai, moksliniai ir švietimo, finansiniai, statistiniai, aplinkos ir transporto
Navut	Sumažinti išlaidas ieškant NT	Kanada	Geografiniai, statistiniai, mokslo ir švietimo
Otthonterkep	Palengvinti NT įsigijimą ir gerinti informacijos apie NT prieinamumą	Vengrija	Geografiniai ir statistiniai

14 lent. Atvirų duomenų panaudojimas statybų ir NT srityje. Sudaryta darbo autoriaus

<b>Geografinis atvaizdavimas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Cemety	Padėti lengviau rasti kapavietes	Lietuva	Geografiniai
Geospheres	Padidinti Maroko geografinių duomenų naudojimą ir prieigą prie jų	Marokas	Geografiniai ir statistiniai
Kiemo broliai	Skatinti aktyvų laisvalaikį	Lietuva	Geografiniai
Nesėdėk namuose	Skatinti aktyvų laisvalaikį ir pažinti Lietuvą	Lietuva	Geografiniai ir aplinkos
Nusikalstamų veiklų Lietuvoje žemėlapis	Mažinti nusikalstamumą ir didinti gyventojų informuotumą apie vykdomus nusikaltimus	Lietuva	Geografiniai ir statistiniai

15 lent. Atvirų duomenų panaudojimas juos atvaizduojant žemėlapyje. Sudaryta darbo autoriaus

<b>Žiniasklaida ir ryšiai</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
KataData	Skelbti duomenis, straipsnius, ataskaitas ir vizualizacijas ekonominei ir verslo analizei	Indonezija	Finansiniai ir statistiniai
La Nacion	Tiriamoji žurnalistika ir atvirų duomenų atvėrimas visuomenei	Argentina	Geografiniai, finansiniai, statistiniai, aplinkos ir transporto
Ojoconmipisto	Skatinti pilietišką žurnalistiką ir piliečių aktyvumą	Gvatemala	Finansiniai ir statistiniai
Twaweza	Informuoti visuomenės ir politikos dalyvavimą naudojantis žiniasklaida	Tanzanija	Moksliniai ir švietimo, finansiniai ir statistiniai

16 lent. Atvirų duomenų panaudojimas žiniasklaidos ir ryšių srityje. Sudaryta darbo autoriaus

<b>Transportas</b>			
<b>Pavadinimas</b>	<b>Tikslas</b>	<b>Šalis</b>	<b>Naudojamų duomenų rūšys</b>
Carvertical	Teikti patikimą informaciją apie naudotą automobilį, jo istoriją	Lietuva	Statistiniai ir transporto
Eismoinfo	Padėti rasti geriausią ir greičiausią maršrutą ir teikti informaciją apie kelių būklę	Lietuva	Geografiniai, statistiniai, metrologiniai ir transporto
InstaVIN	Teikti vartotojams automobilio istoriją	JAV	Statistiniai ir transporto
RailYatri	Supaprastinti keliones traukiniais	Indija	Geografiniai ir transporto
Šviesoforai	Padėti rasti geriausią ir greičiausią maršrutą Lietuvos keliuose	Lietuva	Geografiniai, statistiniai ir transporto



Trafi	Padėti surasti patogią viešąją transporto priemonę ir maršrutą	Lietuva	Geografiniai, statistiniai ir transporto
Vilniaus viešojo transporto judėjimas realiu laiku	Padėti geriau planuoti kelionę viešuoju transportu	Lietuva	Geografiniai ir transporto

17 lent. Atvirų duomenų panaudojimas transporto srityje. Sudaryta darbo autorius

### 3.4 Lietuvos atvirų duomenų panaudojimo atvejai

Internete galima aptikti gausybę duomenų įvairiausiomis temomis, tačiau duomenys yra beverčiai kol jie nėra panaudojami praktiškai. Įvairios valstybinės įstaigos, miestų savivaldybės ar privatus verslai kaupia daug socialinės informacijos, duomenų apie verslo aplinką ar kitų statistinių duomenų, kurių atvėrimas ir panaudojimas pasitarnautų visuomenės gerovei ir verslo skatinimui. Daugėjant gerų ir tvarkingų atvirų duomenų rinkinių skaičiui atsiranda vis daugiau projektų, kurie surenka atvires duomenis iš skirtingų šaltinių, juos apdoroja, analizuoja ir pateikia vienoje vartotoji patogioje sistemoje. Taip žmonės turi galimybę naudotis lengvai prieinama, susisteminta informacija, kuri leidžia lengviau, greičiau ir tiksliau priimti sprendimus.

Atviri duomenys atlieka svarbų vaidmenį verslo, technologijų ir startuolių atsiradimui ir plėtrai, tačiau Lietuvoje dar vis yra didelis institucijų atvirų duomenų brandos lygių skirtumas, dalis duomenų būtinų startuoliui veikti dažnai dar nėra pateikti ir būtina raginti juos kuo greičiau pavišinti, neretai tenka individualiai derėtis dėl duomenų pateikimo konkrečiam projektui, pasirašant sutartis dėl duomenų pateikimo ir panaudojimo. Atvirų duomenų panaudojimo privalumus galime pastebėti sekant užsienio šalių pavyzdžius. Ispanijos vyriausybės pateiktoje sektoriaus, kuriame įmonės panaudoja atvires duomenis komerciniais tikslais 2014 metų ataskaitoje matyti, kad atviri duomenys padėjo sukurti net 494 verslus, kurių daugiau nei pusė jau veikia ilgiau nei 5 metus bendrai sukurdami iki 4700 darbo vietų. Ispanijos įmonės naudojančios atvires duomenis komerciniais tikslais 2014 metais bendrai generavo iki 500 mln. eurų apyvartos. Statistinių duomenų apie Lietuvoje sėkmingai veikiančių verslų paremtų atvirų duomenų panaudojimu nėra, tačiau sėkmingų pavyzdžių šalyje turime mažai, kadangi atvirų duomenų potencialas nėra iki galo išnaudojamas.

#### „Kur gyvenu“?

Tarp geriausiai žinomų atvirais duomenimis besinaudojančių verslų yra UAB „Codein“, kuriantis projektą „Kur gyvenu?“. Projektas 2013 metais gavo fondo „Practica capital“ 200 000 € investiciją, veiklą pradėjęs Vilniuje ilgainiui, gavęs finansavimą išplito ir kitų šalių

didžiuosiuose miestuose – San Franciske, Berlyne, Londone Čikagoje ir Niujorke, pavadinimu „Place I live“. Pagrindinis projekto tikslas padėti nekilnojamo turto pirkėjams, pardavėjams, nuomininkams ar nuomotojams, brokeriams ar NT valdytojams lengviau priimti teisingus sprendimus susijusius su NT. Projektas vienoje vietoje sistemina ir patogiai pateikia aktualiausias duomenis apie NT ir gyvenimo kokybę, apžvelgdamas tokią informaciją gautą iš atvirų duomenų kaip: nurodyto konkretaus nekilnojamo turto apytikslę rinkos vertę, šildymo vidutinį suvartojimą, oro taršos, triukšmo bei 500 metrų spindulių įvykdytų nusikaltimų duomenis, atstumą iki populiariausių ir svarbiausių miesto objektų ir centro, demografinę rajono situaciją ar net aplink esančių vidurinių mokyklų reitingą ir kitus duomenis.

Internetinio projekto platforma yra gana paprasta, tituliname puslapyje pasitinka didelis paieškos laukelis, kurį prašoma užpildyti norimu adresu arba pasirinkus norimą iš siūlomų miestų atsidaro langas su nedideliu, interaktyviu miesto dalies žemėlapiu, laukeliu tiksliam adresui įvesti ir miesto duomenimis, tokiais kaip: vidutinė miesto NT kvadratinio metro kaina, butų ir daugiabučių mieste skaičius, statistiniai nusikaltimų duomenys mieste ir kt. Užpildžius adreso laukelį yra pateikiama ataskaita su konkrečiu NT susijusia informacija ir atsiliepimais apie statinį.

Žemiau pateikiamas duomenų sąrašas, kuris būtinas ir šiuo metu yra naudojamas projektui „Kur gyvenu?“ veiklai vykdyti.

Projektas „Kur gyvenu?“	
Duomenys	Šaltinis
Statybos metai, pastato tipas	VĮ „Registru centras“
Sezono vidutinė šildymo kaina	UAB „Vilniaus energija“
Preliminari vidutinė rinkos kaina	VĮ „Registru centras“
Oro taršos duomenys	Aplinkos apsaugos agentūra
Triukšmo duomenys	Miesto savivaldybė
Atstumas iki svarbių objektų ar mokyklų	Google maps
Nusikaltimų duomenys	Lietuvos Respublikos vidaus reikalų ministerija
Rajono gyventojų demografiniai duomenys	Lietuvos statistikos departamentas
Artimiausių vidurinių mokyklų reitingai	Lietuvos aukštųjų mokyklų asociacija

18 lent. Projekto „Kur gyvenu?“ naudojami atviri duomenys. Sudaryta darbo autoriaus

Svarbu paminėti, kad projektas Lietuvoje atrodo apleistas, duomenys apie vykdomus projektus yra pasenę (projektai jau įvykdyti), naujausių pastatytų gyvenamųjų namų nėra, rinkos kainos, nusikalstamumo duomenys ir kiti rodikliai yra kelių metų senumo, todėl gali neatspindėti

realios šiuo metu esančios situacijos, tačiau kitose šalyse vykdomas projektas „Place I live“ plėtojamas toliau, duomenys yra naujausi ir aktualūs.

### **„Creditinfo“**

Neabejotinai visų bankų, kredito įstaigų ar dalies sąmoningų verslų vienas pagrindinių įrankių išduodant paskolas, pradėdant bendradarbiauti ar net parduodant prekes yra „Creditinfo“ platforma. „Creditinfo“ tai brandus ir labai svarbų vaidmenį Lietuvoje atliekantis atvirais duomenimis paremtas verslas. Daugiau nei 15 metų veiklą vykdomantis UAB „Creditinfo Lietuva“ 2019 balandžio 30 dieną yra sukūrus 47 darbo vietas ir generuoja daugiau nei 3 mln. € apyvartos kasmet. UAB „Creditinfo Lietuva“ 2003 metais įkūrė kredito biurą, taip tapdama agentūra, kuri renka ir saugo privačių ir juridinių asmenų duomenis, kuriais remiantis nustatomas jų mokumas. Kredito biuras, tai informacinė sistema, kurios pagrindinė užduotis apjungti šalies komercinius bankus, kredito unijas, lizingo, vartojimo kredito, taip pat telekomunikacijų ir kitos įmones su tikslu gauti ir teikti informaciją apie privačių ir juridinių asmenų mokumą. Kredito biuras atlieka svarbų vaidmenį visoje šalies finansinėje sistemoje, kadangi finansų įstaigoms padeda priimti teisingus ir pagrįstus kreditavimo sprendimus, skatina atsakingą skolinimąsi, taip pat sudaro sąlygas tvariam paskolų portfelio augimui. Pasaulio bankas ne kartą paminėjo, kad išsami ir lengvai prieinama informacija didina kredito prieinamumą. Pasaulio bankas kasmet vertina pasaulio šalių kredito biurus ir 2017 metų „Doing business 2017“ ataskaitoje informacijos detalumas įvertintas aukščiausiu įvertinimu - 8.0 balu. Lietuvos kredito biuras yra sukaupęs informacijos apie 84 % visų aktyvių šalies gyventojų. Lietuva papildomai įvertinta už tai, kad finansų įstaigos turi galimybę ataskaitoje matyti kredito biuro apskaičiuotą vertinamo privataus ar juridinio asmens kredito reitingą.

Visa „Creditinfo“ veikla paremta dviejų rūšių gaunamais duomenimis, tai mokami arba nemokami atviri duomenys iš valstybės registru ir kreditų biuro dalyvių valdoma informacija, kuria, sutartomis sąlygomis, kredito biuras turi teisę naudotis ir dalintis. „Creditinfo“ vadovo Andriaus Bogdanovičiaus teigimu, jo valdoma įmonė kasmet tiesiogiai arba per tarpininkus iš valstybės registru perka duomenų už maždaug 600 000 €.



12 pav. „Creditinfo“ duomenų gavimo ir dalijimosi grafinis atvaizdavimas

(UAB „Creditinfo Lietuva“, 2019)

Tik prisijungus prie internetinės svetainės visa platforma yra dalinama į du vartotojų segmentus, tai paslaugos gyventojams ir paslaugos verslui. Pasirinkus paslaugas verslui svetainė nukreipia į skiltį su informacija apie teikiamas paslaugas, tačiau paslaugos teikiamos tik registruotiems kreditų biuro klientams. Privačių asmenų paslaugų skiltyje identifikavus savo tapatybę „Creditinfo“ suteikia galimybę gauti prisijungusio asmens kredito istorijos ataskaitą, užsisakyti paslaugą „nenoriu kredito“, kuri mažiausiai 90 dienų apribos galimybę skolintis, „tapatybės apsaugos“ paslaugą, kuri kredito istorijoje sukuria 100 mln € skolos įrašą, tačiau gali būti išjungta iškart bet kuriuo metu nepalikdama jokių pėdsakų kredito istorijoje arba susipažinti su duomenimis apie save, kartu ir sužinoti kredito reitingą.

Žemiau pateikiamas duomenų sąrašas, kuris būtinas ir šiuo metu yra naudojamas kreditų biuro „Creditinfo“ veiklai vykdyti.

<b>Kreditų biuras „Creditinfo“</b>	
<b>Duomenys</b>	<b>Šaltinis</b>
Verslo ir privačių asmenų mokėjimo istorija	Kreditų biuro dalyviai
Vėlavimo atsiskaityti ir bankroto reitingus	Kreditų biuro dalyviai
Informacija apie sąsają su kitais verslo subjektais	Registrų centras, juridinių asmenų registras
Finansinės ataskaitos	Registrų centras, juridinių asmenų registras
Informacija apie darbuotojų ir vadovų kaitą	Registrų centras, juridinių asmenų registras
Informacija apie juridinių asmenų transporto priemones	VĮ „Regitra“
Turto areštų informacija	Registrų centras, turto arešto aktų registras
Teisminės bylos	Lietuvos teismai, teisiniu bylu registras
Antstolių informacija	Registrų centras, antstolių informacinė sistema
Privačių asmenų finansiniai įsipareigojimai	Kreditų biuro dalyviai
Aktyvių sudarytų kredito sutarčių sąrašas	Kreditų biuro dalyviai
Pradelstų mokėjimų sąrašas	Kreditų biuro dalyviai
Asmenine kredito istorija besidomėjusių įmonių sąrašas	Kreditų biuro dalyviai

19 lent. Kreditų biuro „Creditinfo“ naudojami duomenys. Sudaryta darbo autoriaus

### **„Credit reform“**

Dar viena atvirais duomenimis paremta organizacija, kuri yra padeda mažinti riziką ir yra lyderė sprendžiant prekinio kreditavimo problemas Lietuvoje. UAB „Credit reform Lietuva“ yra didžiausio Europoje kreditinės informacijos ir skolų išieškojimo susivienijimo atstovas Lietuvoje. Šiuo metu įmonė aptarnauja didžiausius Lietuvoje veikiančius lietuviško ir užsienio kapitalo verslus, draudimo ar lizingo kompanijas ir bankus, kasmet ruošdama tūkstančius ataskaitų apie Lietuvos įmones, taip pat vykdo privačių ir juridinių asmenų skolų išieškojimą. „Credit reform“ duomenų bazėje vartotojai pateikia virš 25 000 tiesioginių paieškų kasdien. Užsienio įmonės siekiančios įsitvirtinti Lietuvoje vis dažniau naudojasi „Credit reform“ siūloma masine rinkodaros atranka pagal užsakovo nurodytus kriterijus, taip sudarant tinkamų pirkėjų ir tiekėjų sąrašus su išsamia kontaktine informacija. Remiantis „Credit reform“ pateiktų veiklos aprašu, pagrindinės teikiamos veiklos sritys: - nemokumo bei skolų prevencija - klientų mokumo ir kreditavimo rizikų įvertinimas internetu: [www.cr.lt](http://www.cr.lt) informacija: skolininkų kontaktai, kreditingumas, skolos, debitoriai, teismai, negatyvi vieša, balansai, pelnas, apyvartos, darbuotojų skaičius, akcininkai, mobilūs telefonai, PVM statusas,

kreditinė informacija apie įmones, jų kreditavimo limitų ir kreditingumo reitingų nustatymas pavienių ir "masinių" skolų išieškojimas Lietuvoje ir užsienyje ir kt.

„Credit reform“ paslaugų spektras formuojamas iš kelių pagrindinių elementų, prisijungę vartotojai gali užsisakyti dominančios įmonės išsamią kredito ataskaitą, užsisakyti kredito reitingo paslaugą ar pagal nustatytus kriterijus gauti atrinktų klientų ar partnerių sąrašą.

Vienas geriausių įrankių priimant sprendimą ar verta pradėti prekiauti ar turėti kitokių verslo sąsajų su konkrečia įmone yra dominančios įmonės kredito ataskaita, kurioje yra visi įmonės rekvizitai ir kontaktiniai duomenys, įmonės esamas ir buvę vadovai, atstovaujамų prekinių ženklų sąrašas, informacijos iš spaudos santraukos, filialų ir licencijų informacija, darbuotojų kitimo informacija ir finansinės ataskaitos, įmonės dalyvavimo teismuose istorija ir kt.

Pigesnė, galinti tiek papildyti kredito ataskaitą, tiek būti užsakoma atskirai, tačiau mažiau informatyvi yra CR kredito reitingo nustatymo paslauga. Ši paslauga yra subjektyvi įstaigos „Credit reform“, jos skaičiavimo metodika nėra viešai skelbiama, tačiau reitingas skaičiuojamas atsižvelgiant į įvairius nemokumo faktorius, ūkio šakos tendencijas, finansinius duomenis ir kt. Užsisiūkus norimos įmonės ataskaitą, informacija pateikiama trimis detalumo lygiais:

**CR reitingas** - rizikingumo skalė nuo 100 (nepriekaištingas kreditingumas) iki 600 (bankrotas).

**CR rizikos klasė** - labiau apibendrinta rizikingumo klasifikacija nuo 1 iki 10 lengvesniam kreditavimo sprendimui priimti.

**CR limitas** - visiems kreditoriams bendras didžiausias leistinas kredito limitas taikant 30 d. atidėjimo terminą. (nurodoma rekomenduojama suma eurais).

Teisingai sprendimo paramai pateikiama yra ir istoriniai kredito reitingo kitimo duomenys. Istoriniai duomenys parodo ilgalaikę įmonės veikimo sėkmę, priešingai nei momentinis kreditingumo vertinimas.

Trečioji iš pagrindinių „Credit reform“ teikiamų paslaugų, tai pagal kliento poreikį surinktas ir sudarytas įmonių sąrašas. Klientui nustačius pageidaujamus kriterijus, tokius kaip: įmonės veiklos sritys, regionas, apyvarta, darbuotojų skaičius ir pan. „Credit reform“ naudodamiesi viešai prieinama, pagal sutartis prieinama valstybės registrų ar savo duomenų bazių informacija surenka reikiamą informaciją ir pateikia ją susistemintą klientui su įmonės kontaktinių asmenų informacija pirmam kontaktui užmegzti. Paslauga orientuota į užsienio įmones Lietuvoje ar Lietuviškas įmones siekiančias įsitvirtinti rinkoje ar vykdyti tikslinius pardavimus išskirtinai į jų siūlomą produktą orientuotai auditorijai ir kitais panašiais atvejais.

„Credit reform“ skatindama mažinti įmonių pradelstų įsiskolinimų kiekį ir kartu gerinti prekinio kreditavimo aplinką, kartu su Lietuvos pramonininkų konfederacija sukūrė savo rodiklį – darbdavių skolos indeksą (DIX), kurį peržiūrėti galima „Credit reform“ internetinė svetainėje. DIX rodiklis yra paprastas kelių viešai prieinamų duomenų santykis:

$$DIX = \frac{\text{Bendra įmonių skola Sodrai}}{\text{šalies įmonių skaičiaus, kurios turi bent 1 darbuotoją}}$$

DIX – šiuo metu objektyviausias rodiklis, skirtas matuoti einamajam įmonių mokumui. Rodiklis yra skelbiamas kasdien ir rodiklis parodo ar įmonių sąskaitose tikrai esama pinigų, ar įmonė moki. Kuo mažesnė indekso reikšmė – tuo geriau sumokami ar išieškomi socialinio draudimo mokesčiai, taip pat tai parodo ir geresnę bendrą ekonomikos būklę. „Credit reform“ teigimu: „Faktas, kad įmonė skolinga Sodrai vertingesnis už faktą, kad įmonė skolinga kitai įmonei. Paprastai jei įmonė daugiau kaip mėnesį vėluoja mokėti Sodrai, ji jau būna skolinga ne tik savo tiekėjams ar kreditoriams, bet net ir mokesčių inspekcijai. Rizikos vertinimui įdomiausios yra senesnės nei 30 d. skolos SODRAI.“

Toliau pateikiamas duomenų sąrašas, kuris būtinas ir šiuo metu yra naudojamas UAB „Credit reform Lietuva“ veiklai vykdyti.

<b>Projektas „Credit reform“</b>	
<b>Duomenys</b>	<b>Šaltinis</b>
Įmonės rekvizitai ir kontaktiniai duomenys	Registrų centras
Esamas ir buvę vadovai	Registrų centras, juridinių asmenų registras
Akcininkų sąrašas	Registrų centras, juridinių asmenų registras
Finansinės ataskaitos	Registrų centras, juridinių asmenų registras
Turto areštų informacija	Registrų centras, turto arešto aktų registras
Hipotekos registro informacija	Registrų centras, hipotekos registras
Teisminės bylos	Lietuvos teismai, teisiniu bylu registras
Antstolių informacija	Registrų centras, antstolių informacinė sistema
Darbuotojų kitimo istorija	VĮ „Sodra“
Bendriniai statistiniai ūkio šakos, kuriai priklauso įmonė, rodikliai	Statistikos departamentas + Registrų centras
Valdomų transporto priemonių skaičius	VĮ „Regitra“
Spaudos pranešimų sąrašas	Delfi, Verslo žinios, 15min, Lietuvos rytas ir visi kiti informaciniai kanalai
Informacija apie domėjimąsi konkrečia įmone duomenys	Vidinė „Credit reform“ duomenų bazė
Įmonių valdomas NT	Registrų centras, nekilnojamo turto registras
Skolų Sodrai rodiklis	VĮ „Sodra“
Vidutinio darbo užmokesčio duomenys	VĮ „Sodra“
Juridinių asmenų įsiskolinimų duomenys	VĮ „Sodra“

20 lent. Projekto „Credit reform“ naudojami duomenys. Sudaryta darbo autoriaus



### 3.5 Asociacijų paieškos tyrimas

Ieškant asociacijų tarp pasaulyje ir Lietuvoje sėkmingų, atvirais duomenimis paremtų projektų buvo sudarytas binarinis duomenų rinkinys, kuriame yra informacija apie tiriamus projektus, jų naudotus atvirų duomenų rinkinius ir požymis nurodantis kokiai veiklos sričiai projektas yra priskiriamas.

Sudarius atvirai prieinamų duomenų panaudojimo atvejų pasaulyje duomenų rinkinį, siekiant padaryti išvalgas, naudojome SPSS Modeler programinę įrangą, kurioje sudaromas duomenų rinkinio analizės planas.

Tyrimui tinkami trys iš keturių SPSS Modeler siūlomi asociacijų paieškos metodai:

#### **Apriori**

„Apriori“ iš duomenų sudaro taisyklių rinkinį, sukurdamas taisykles, turinčias didžiausią informacijos turinį. „Apriori“ siūlo penkis skirtingus taisyklių pasirinkimo būdus ir naudoja sudėtingą indeksavimo schemą, kad efektyviai apdorotų didelius duomenų rinkinius. „Apriori“ paprastai yra greitai apmokomas, „Apriori“ reikalauja, kad visi įvesties ir išvesties laukai būtų binariniai (Flag), tačiau užtikrina geresnį našumą, nes jis yra optimizuotas šio tipo duomenims.

Apriori metodo silpnoji savybė – metodas leidžia nustatyti norimą didžiausią priežasčių skaičių taisyklėje, tačiau visada rezultate pateikia tik vieną priežastį. Todėl metodas dažnai nepateikia visų galimų ir įdomių tyrimo rezultatų. Atvirų duomenų panaudojimo tyrimui Apriori metodas nėra tinkamas, nes kuriamiems projektams atvirų duomenų rinkinių aibės yra plačios ir viena asocijuota duomenų rūšis taisyklėje gali teisingai nepadengti visų tiriamų atvejų.

#### **Carma**

„Carma“ yra alternatyvus metodas „Apriori“ metodui, tačiau „Carma“ iš duomenų rinkinio išskiria taisyklių rinkinius nereikalaujantis nurodyti įvesties ar tikslo laukų, dėl to šis asociacijų paieškos būdas yra plačiau pritaikomas. „Carma“ priešingai nei „Apriori“ analizuojant duomenis nustato taisykles tiek pirminiam įvykiui, tiek asocijuotam antriniam įvykiui, priešingai nei „Apriori“. Taip pat „Carma“ leidžia taikyti taisykles su keliais rezultatais.

„Carma“ metodas yra labiausiai tinkamas atvirai prieinamų duomenų rinkinių asociacijų tyrimui, kadangi kurdamas taisykles metodas aptinka asociacijas tarp kelių priežasčių ir kelių pasekmių, remdamasis nuodytu didžiausių visos taisyklės elementu skaičiumi.

## Asociacijų taisyklės

„Asociacijų taisyklė“ yra funkcija, kuri nurodo duomenų rinkinio statistiką ir duomenų rinkinyje esančias įdomiausias asociacijas. Įprastai ši taisyklė sieja konkrečią išvadą, pvz.: konkrečios prekės pirkimas, jei perkama kita ar kelios kitos prekės.

„Asociacijų taisyklės“ metodas taip pat yra siauresnio panaudojimo nei „Carma“, tačiau kartu pateikia papildomą statistinę informaciją apie duomenų rinkinį, kuri gali būti aktuali papildant tyrimą.

Tyrimas buvo atliktas visais trimis metodais, tačiau „Carma“ metodas dėl savo plačiausio spektro ir giliausių įžvalgų galimybių analizuojant atvirai prieinamų duomenų rinkinių panaudojimo duomenų rinkinį buvo tinkamiausias. „Carma“ metodo rezultatai aptariami žemiau.

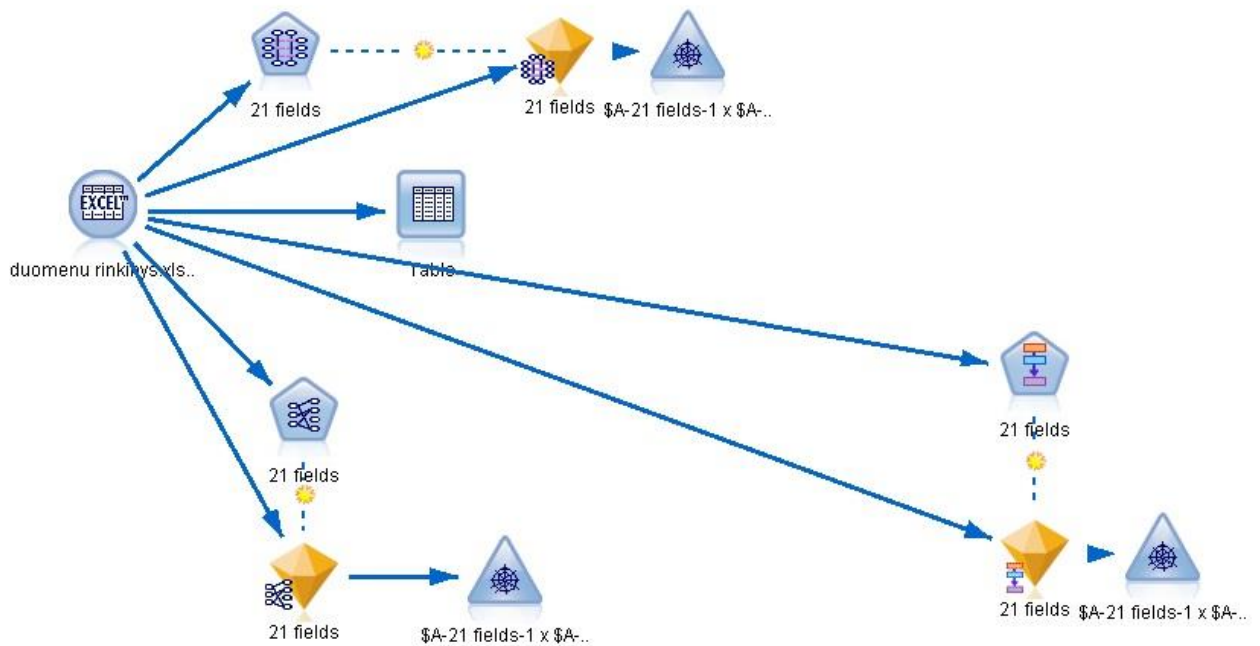
### Tyrimo naudotos sąvokos:

**Paramos rodiklis** (support) – remiantis duotais duomenimis nurodo proporciją, kad pirminis įvykis yra teisingas. Pvz.: jei 50% pateiktų duomenų yra duonos pirkimas, tada taisyklė, kad perkant duoną, bus perkamas ir sūris turės 50% patikimumo paramą.

**Taisyklės paramos rodiklis** (Rule support) – rodiklis nurodo tiesioginį ryšį tarp įvykių. Kokia tikimybė, kad įvykus A įvykių kombinacijai, įvyko ir B įvykis ar įvykių kombinacija. Pvz.: 20 % duomenų turi kombinaciją perkant ir duoną perkamas ir sūris. Tai taisyklės paramos rodiklis bus – 20 %.

**Patikimumas** (Confidence) – Rodo santykį tarp paramos rodiklio ir taisyklės paramos rodiklio. Pvz.: šiuo atveju 20 % duomenų sudaro duonos ir sūrio kombinaciją ir 50 % duomenų pirminis įvykis yra duonos pirkimas. Tokiu atveju patikimumo rodiklis bus -  $\frac{20\%}{50\%} = 40\%$

**Įdiegiamumo rodiklis** (Deployability) – Matas išreikštas procentais nurodantis analizuojamų duomenų dalį, kai įvyko pirmas įvykis, tačiau dar neįvyko asocijuotas įvykis. Šiuo atveju pvz.: visos klientų bazės procentinė dalis, kuri pirko duoną, tačiau dar nepirko sūrio.



13 pav. SPSS Modeler asociacijų paieškos funkcijų žemėlapis

Funkcijų žemėlapyje grafiškai vaizduojama duomenų rinkinio trys asociacijų paieškų tipai. Duomenų rinkinis tinkamas analizuoti „Apriori“, „Carma“ ir Asociacijų taisyklės metodais. Žemiau pateikti tinkamiausio „Carma“ metodo aukščiausius rodiklius turintys rezultatai ir jų analizė.

Ne Lietuvoje veikiančių projektų naudojamų duomenų rinkinių asociacijų tyrimo rezultatai „Carma“ metodu:

Duomenų rinkinių ir jų kombinacijų rūšys	Paramos rodiklis „Support %“
Statistiniai duomenys	81,48 %
Geografiniai duomenys	48,15 %
Finansiniai duomenys	42,59 %
Moksliniai ir švietimo duomenys	40,71 %
Moksliniai, švietimo ir statistiniai duomenys	38,89 %
Finansiniai ir statistiniai duomenys	38,89 %
Aplinkos duomenys	35,19 %

21 lent. Ne Lietuvoje dažniausiai naudotos duomenų rūšys pagal aukščiausią „Paramos rodiklį“

Carma metodo sukurtose taisyklėse galima matyti įdomių duomenų rinkinių kombinacijų:

<b>Duomenų rinkinių ir jų kombinacijų rūšys</b>	<b>Paramos rodiklis „Support %“</b>
Moksliniai, švietimo, finansiniai ir statistiniai duomenys	18,52 %
Geografiniai, moksliniai ir švietimo duomenys	16,67 %
Geografiniai, moksliniai ir švietimo duomenys	14,82 %
Geografiniai, finansiniai ir aplinkos duomenys	7,41 %
Statistiniai, aplinkos ir transporto duomenys	7,41 %

22 lent. Ne Lietuvoje naudotos duomenų rūšys pagal aukščiausią „Paramos rodiklį“

Remiantis „Carma“ pagal aukščiausią paramos rodiklį išrinktais duomenimis, pastebime dažniausiai ne Lietuvoje naudojamas duomenų rinkinių rūšis. Aukščiausius rodiklius turi statistiniai, geografiniai, finansiniai ir kiti duomenų rinkiniai, tačiau matoma ir kiek tolimesnių viena kitai sričių duomenų rinkinių panaudojimą, tokių kaip: moksliniai, švietimo, finansiniai ir statistiniai kartu, geografiniai, moksliniai ir švietimo duomenys kartu ir kitos kombinacijos.

Žemiau pateiktoje lentelėje matome tiesioginį ryšį tarp panaudotų duomenų rūšių. Taisyklės paramos rodiklis parodo kiek atvejų panaudojus vieną duomenų rūšį buvo naudojama konkreti kita duomenų rūšis.

<b>Jeigu naudojami</b>	<b>Kartu naudojami</b>	<b>Taisyklės paramos rodiklis „Rule support %“</b>
Moksliniai ir švietimo duomenys	Statistiniai duomenys	38,89 %
Finansiniai duomenys	Statistiniai duomenys	38,89 %
Geografiniai duomenys	Statistiniai duomenys	35,19 %
Aplinkos duomenys	Geografiniai duomenys	25,92 %
Statistiniai duomenys	Aplinkos duomenys	24,07 %
Moksliniai, švietimo ir finansiniai duomenys	Statistiniai duomenys	18,52 %
Statistiniai ir aplinkos duomenys	Geografiniai duomenys	16,67 %
Moksliniai ir švietimo duomenys	Geografiniai duomenys	16,67 %

23 lent. Ne Lietuvoje naudotos duomenų rūšys pagal aukščiausią „Taisyklės paramos rodiklį“

Atliekant tyrimą pastebėtos nestandartinės, rečiau ne Lietuvoje pasitaikančios, tačiau taip pat ženklų statistinį dažnumą turinčios duomenų rinkinių naudojimo su kitais duomenų rinkiniais asociacijos:

<b>Jeigu naudojami</b>	<b>Kartu naudojami</b>	<b>Taisyklės paramos rodiklis „Rule support %“</b>
Metrologiniai duomenys	Geografiniai duomenys	12,96 %
Geografiniai duomenys	Transporto duomenys	12,96 %
Finansiniai ir transporto duomenys	Statistiniai duomenys	11,11 %
Statistiniai ir metrologiniai duomenys	Geografiniai duomenys	11,11 %
Geografiniai, transporto ir finansiniai duomenys	Statistiniai duomenys	9,26 %

24 lent. Kiti ne Lietuvoje statistiškai reikšmingai dažnai naudotos duomenų rūšys pagal aukščiausią „Taisyklės paramos rodiklį“

Pateiktose lentelėse matoma, kad pasaulyje be Lietuvos aukščiausią taisyklės paramos rodiklį turi moksliniai ir švietimo bei finansiniai duomenys panaudojus vienus iš geografinių, mokslinių ar finansinių duomenų net 38,89 % atvejų kartu naudojami ir statistiniai. Kas ketvirtą kartą panaudojus aplinkos duomenis, naudojami geografiniai duomenys arba panaudojus statistinius duomenis, naudojami aplinkos duomenys. Pasaulio šalyse be Lietuvos pastebimas įdomių daugiau nei dviejų ar trijų atvirų duomenų rūšių naudojimas, tokių kaip: 18,52 % kartų kai panaudojama moksliniai, švietimo ir finansiniai duomenys kartu naudojami ir statistiniai duomenys arba panaudojus finansinius ir transporto duomenis, 11,11 % atvejų kartu naudojami statistiniai duomenys ir kt.

<b>Jeigu naudojami</b>	<b>Kartu nebus naudojami nė viena rūšis iš išvardintų</b>	<b>Įdiegiamumo rodiklis „Deployability %“</b>
Statistiniai duomenys	Aplinkos, geografiniai, moksliniai ir švietimo ar finansiniai duomenys	42,59% - 57,41 %
Geografiniai duomenys	Metrologiniai, transporto, finansiniai ar moksliniai ir švietimo duomenys	31,48 % - 35,19 %
Finansiniai duomenys	Transporto ar aplinkos duomenys	29,63 % -31,48 %
Moksliniai ir švietimo duomenys	Aplinkos duomenys	29,63 %
Finansiniai duomenys	Geografiniai duomenys	27,78 %

25 lent. Ne Lietuvoje naudojamų atvirų duomenų rūšys nebuvo panaudotos su kitomis duomenų rūšimis pagal aukščiausią „Įdiegiamumo rodiklį“

Įdiegiamumo rodiklis parodė, kad pasaulio šalyse be Lietuvos projektai naudoja gana įvairias duomenų rūšių kombinacijas, kadangi išsiskyrė tik statistinių duomenų rūšis, kuriuos panaudojus net 42,59% - 57,41 % atvejų kartu nebuvo naudojami aplinkos, geografiniai, moksliniai ir švietimo ar finansiniai duomenys, tačiau kitos atvirų duomenų rūšys tik iki trečdaliao atvejų nebuvo kartu naudojamos su nuo vienos iki keturių kitų duomenų rūšių. Išvestos taisyklės rodo, kad duomenų rūšių panaudojimas yra gana įvairus ir nėra griežtų taisyklių ar ribų duomenų rūšių naudojimui kartu kuriant naujus produktus sprendimams priimti.

Atlikus užsienio šalių atvirų duomenų rūšių panaudojimo asociacijų tyrimą galima matyti, kad už Lietuvos ribų atvirais duomenimis paremti verslai – projektai naudoja daug įvairių duomenų rinkinių, kartais apjungdami ir rečiau naudojamas duomenų rūšis, taip sukurdami gilesnes ar dar nematytas išvalgas ir pagalbą priimant sprendimus. Dėl to kuriami unikalūs projektai padedantys kokybiškai aplankyti muziejus ar kitas lankytinas vietas be gido, palengvinantys ūkininkavimą, padedantys gelbėti žmonių ir gyvūnų gyvybes ir kt.

Lietuvoje veikiančių projektų naudojamų duomenų rinkinių asociacijų tyrimo rezultatai „Carma“ metodu:

Lietuvoje naudojamų duomenų rinkinių ar jų kombinacijų sąrašas labai siauras ir skurdus. „Carma“ metodo taisyklėse 37-75% paramos rodiklį turi tik geografiniai, statistiniai, geografiniai kartu su statistiniais, transporto arba transporto kartu su geografiniais duomenų rinkiniai ar jų kombinacijos. Tai yra pagrindinės duomenų rūšys, kuriomis remiantis kuriami atvirais duomenimis paremti projektai Lietuvoje, rezultatai parodo labai siaurą duomenų rinkinių panaudojimo aibę ir parodo, kad beveik visi Lietuvoje sukuriami atvirais duomenimis paremti projektai apima vos dvi, tris galimas sritis.

Lietuvoje naudotų duomenų rūšių kombinacijos pagal aukščiausią taisyklės paramos rodiklį, kuris parodo tiesioginį ryšį ir antrinio įvykio tikimybę įvykus pirminiam įvykiui:

<b>Jeigu naudojami</b>	<b>Kartu naudojami</b>	<b>Taisyklės paramos rodiklis „Rule support %“</b>
Geografiniai duomenys	Statistiniai duomenys	50,00 %
Transporto duomenys	Geografiniai duomenys	31,25 %
Transporto duomenys	Statistiniai duomenys	31,25 %
Finansiniai duomenys	Statistiniai duomenys	25,00 %
Geografiniai ir transporto duomenys	Statistiniai duomenys	25,00 %
Aplinkos duomenys	Geografiniai duomenys	18,75 %
Moksliniai ir švietimo duomenys	Geografiniai duomenys	12,50 %
Metrologiniai duomenys	Geografiniai duomenys	12,50 %

26 lent. Lietuvoje naudotos duomenų rūšys pagal aukščiausią „Taisyklės paramos rodiklį“

Analizuojant asociacijos taisykles pagal aukščiausią taisyklės paramos rodiklį Lietuvoje matomas siauras ir paprastas duomenų rinkinių rūšių pasirinkimas. Nėra trijų, keturių duomenų rūšių asociacijų, o esamos turi didelį statistinį pasikartojimą vienai baigčiai.

Net pusė atvejų kai naudojami geografiniai atviri duomenys Lietuvoje kartu naudojami ir statistiniai duomenys, beveik vieną trečdalį atvejų panaudojus transporto duomenis kartu naudojami geografiniai, dar beveik vieną trečdalį atvejų panaudojus transporto naudojami statistiniai duomenys. Matoma, kad Lietuvoje vystomi projektai įprastai remiasi vienos srities duomenimis juos apjungdami su geografiniais arba statistiniais duomenimis išvalgoms padaryti, todėl tikėtina, kad retas projektas būna kompleksinis ir sudėtingas, o atlieka tik vieną ar kelias bazines funkcijas.

<b>Jeigu naudojami</b>	<b>Kartu nebus naudojami nė viena rūšis iš išvardintų</b>	<b>Įdiegiamumo rodiklis „Deployability %“</b>
Kultūros	Moksliniai ir švietimo, metrologiniai, aplinkos, finansiniai ar transporto duomenys	62,5% - 87,5%
Geografiniai duomenys	Moksliniai ir švietimo ar metrologiniai duomenys	62,5%
Statistiniai duomenys	Aplinkos, moksliniai ar švietimo ar metrologiniai duomenys	62,5%
Geografiniai duomenys	Aplinkos duomenys	56,25%
Statistiniai duomenys	Aplinkos duomenys	50,00%

27 lent. Lietuvoje naudojamų atvirų duomenų rūšys nebuvo panaudotos su kitomis duomenų rūšimis pagal aukščiausią „Įdiegiamumo rodiklį“

Įdiegiamumo rodiklis rodo kokios atvirų duomenų rūšių panaudojimo kombinacijos, kaip dažnai neįvyko. Lietuvoje galima matyti, kad dažniausiai naudojant kultūros duomenys buvo sunku suderinti daugumą kitų duomenų rūšių, tikėtina, dėl to, kad atviri kultūros duomenys beveik nebuvo naudojami Lietuvoje kuriamuose projektuose, tačiau matoma, kad nebuvo retai buvo sutinkamos kartu tokios atvirų duomenų rūšys kaip geografiniai su moksliniais, geografiniai su metrologiniais, statistiniai su aplinkos ar geografiniai su aplinkos, kurios ,atrodytų, galėti viena kitą stipriai papildyti.

Išanalizavus Lietuvoje naudojamų duomenų rūšių taisykles matoma, kad Lietuvoje atvirais duomenimis paremti projektai remiasi vos keliomis duomenų rūšimis, duomenų rūšių kombinacijos taip pat yra labai standartinės, kurios įprastai sudaromos ir statistinių, finansinių ir / ar geografinių duomenų rinkinių. Lietuvoje sukurti, atvirais duomenimis paremti, projektai dažniausiai yra susiję su viešu transportu, finansiniais duomenimis apie įmones ar privačius asmenis arba žemėlapiuose atvaizduojama kokia nors statistika. Be kelių ryškesnių projektų dauguma yra panaudoję vos dvi ar tris atvirų duomenų rūšis, taip sukurdami labai paprastą bazinį produktą, dažnu atveju nekonkurencingą ir nepatrauklų pritaikymui už Lietuvos ribų.



Lyginant Lietuvoje ir likusiose pasaulio šalyse naudojamų atvirų duomenų rūšių asociacijų tyrimo rezultatus matoma tendencija, kad visame pasaulyje daugiausiai naudojami geografiniai ir statistiniai duomenys, tačiau matomi ir aiškūs skirtumai:

Taisyklės paramos rodikliai parodė, kad Lietuvoje duomenų rūšių derinimas su kitos rūšies duomenų rinkiniais yra labai žemame lygyje, beveik visi projektai panaudoję geografinius duomenis kartu remiasi ir statistiniais duomenimis, panaudojus transporto – naudojami geografiniai arba statistiniai, tačiau kitose pasaulio šalyse tos pačios duomenų rūšys yra plačiau panaudojamos ir sukuriamos įdomesnės duomenų rūšių kombinacijos: Panaudojus mokslinius duomenis net 39% atveju kartu naudojama ir statistiniai duomenys, 25% atvejų geografinius duomenis kartu naudojo su aplinkos duomenimis, kas penktą kartą finansiniai duomenys su moksliniais duomenimis.

Įdiegiamumo rodiklis parodė, kad kultūros duomenys Lietuvoje beveik netinkami naudoti su penkiomis kitomis rūšimis duomenų kaip moksliniai, metrologiniai, aplinkos, finansiniai ir transporto, net iki 82% atvejų kultūros duomenys nebuvo panaudoti kartu su išvardintomis duomenų rūšimis. Tuo tarpu užsienio šalyse tokių aukštų rodiklių neturi ne viena duomenų rūšis.

Tyrimas parodo, kad Lietuvoje ženkliai skurdesnis duomenų rūšių taikymas kuriant atvirais duomenimis paremtus projektus, dauguma projektų remiasi vos keliomis pagrindinėmis duomenų rūšimis. Pasaulinės tendencijos panašios, dažniausiai naudojami statistiniai, finansiniai ir geografiniai duomenys, tačiau priešingai nei Lietuvoje matomas gana dažnas mokslinių ir švietimo, aplinkos ar metrologinių duomenų naudojimas.

Dr. Michiel de Vries (2016) išsakė savo nuomonę, kad vien duomenų paskelbimas ir tūkstančio Excel lentelių su duomenimis paviešinimas neduoda jokios naudos, tai tėra formalumas. Duomenys turi būti pateikti su kontekstu, su ekspertų komentarais ir išaiškinimais, tik taip bus maksimaliai skatinamas atvirų duomenų panaudojimas.

Vienos didžiausios pasaulyje programinės įrangos kūrėjos SAP valdybos nario, Juergeno Muellero teigimu<sup>11</sup> (2019) duomenų vertė nustatoma naudojantis tokia formule:

**Duomenų kiekis \* Duomenų kokybė \* Duomenų panaudojamumas = Duomenų vertė**

Remiantis Juergeno Muellero teigimu galime manyti, kad Lietuvoje atviri duomenys nėra taip sėkmingai naudojami kaip kitose pasaulio šalyse todėl, kad nėra užtikrintos visos trys duomenų vertės sudedamosios dalys. Mažas duomenų kiekis, prasta kokybė arba netinkamas ar nepatogus duomenų panaudojamumas, net esant žemam tik vienam iš rodiklių bendra duomenų vertė tampa žema.

---

<sup>11</sup> <https://www.datanami.com/2019/10/09/hana-in-the-cloud-by-christmas-sap-says/> (peržiūrėta: 2019-12-18)

## IŠVADOS IR REKOMENDACIJOS

1. Atvirai prieinamų duomenų atvėrimo judėjimo tikslas yra skatinti dalintis jau kartą panaudotais ir paruoštais kompiuteriniam apdorojimui duomenų rinkiniais, kad verslas, privatūs asmenys ar visuomenė gautų naudą. Naudojimo mastai sparčiai auga ir spėjama augti, skaičiuojama, kad atvirai prieinami duomenys gali padėti kasmet pasaulyje sugeneruoti ir padėti sutaupyti apie 3 trilijonus JAV dolerių.
2. Atvirų duomenų panaudojimas teikia tiesioginę naudą. Antrinis duomenų panaudojimas lengvina priimamus sprendimus, gerina viešojo sektoriaus efektyvumą, didina skaidrumą bei atskaitomybę, teikia ekonominę naudą, kadangi kuriami inovatyvūs verslai - kuriamos naujos, gerai apmokamos, darbo vietos. Manoma, kad 2020 metais Europoje bus 13 000 laisvų darbo vietų darbui su atvirais duomenimis.
3. Lietuva pasaulio mastu yra vertinama kaip atvirų duomenų kultūros vidutiniokė. Pirmoje vietoje esantis Singapūras yra atvėręs tik ~1600 duomenų rinkinių daugiau (100 kartų mažiau nei 13 vietoje esanti JAV), todėl galima teigti, kad duomenų kokybė yra ne mažiau svarbu nei duomenų kiekis. Mažas Lietuvoje atvertų duomenų rinkinių skaičius nėra vienintelė problema, ne mažesnę dėmesį reikėtų skirti duomenų kokybei, taip būtų galima gerinti Lietuvos įvaizdį atvirų duomenų reitinge.
4. Lyginant Lietuvoje ir likusiame pasaulyje atvirų duomenų panaudojimo atvejus matyti, kad Lietuvoje dažniausiai naudojami tik kelių rūšių atviri duomenys, statistiniai, geografiniai ir transporto, o likusiame pasaulyje naudojamų atvirų duomenų rinkinių kombinacijos yra įvairesnės. Galima manyti, kad Lietuvoje trūksta atvirų duomenų rinkinių retai naudojamose srityse arba jų kokybė yra netinkama antriniam naudojimui.
5. Asociacijų tyrimą papildant surinkta informacija apie Lietuvos atvertų duomenų rinkinius galima teigti, kad norint praplėsti naudojamų duomenų rūšių aibę ir padėti kurti įvairesnius projektus būtina atverti daugiau kultūros, švietimo ir mokslo ir metrologinių duomenų rinkinių, šiose srityse atverta nuo vos kelių iki 20 duomenų rinkinių. Aplinkos rūšies atverti 55 duomenų rinkiniai – duomenų rinkinius kiekis nėra išskirtinai mažas, todėl reikėtų atkreipti dėmesį į duomenų vertę, tikėtina, kad dėl žemos duomenų rinkinių vertės aplinkos duomenų rinkiniai Lietuvoje beveik nėra naudojami.
6. Remiantis pasaulyje panaudotų duomenų rūšių atvejų asociacijų tyrimu galima būtų papildyti Informacinės visuomenės plėtros komiteto vykdomo duomenų atvėrimo poreikio nustatymo metodiką. Asociacijų tyrimas gali padėti nustatyti duomenų rūšių atvėrimo poreikius, taip

pagilinant išvalgas kai asociacijų tyrimas taikomas su standartiniais, jau nustatytais metodais kaip duomenų atvėrimo paklausimų analizė, apklausų organizavimas, pastabų analizavimas ir kt.

## LITERATŪRA

1. 2016-2018 Lithuania's action plan for participation in international initiative of open government partnership, (2016),  
URL:[https://www.opengovpartnership.org/sites/default/files/AVP\\_planas\\_2016-2018\\_en%20\(1\).pdf](https://www.opengovpartnership.org/sites/default/files/AVP_planas_2016-2018_en%20(1).pdf) [žiūrėta 2018-12-20].
2. Analytical Report 9: The Economic Benefits of Open Data (2017)  
URL:  
[https://www.europeandataportal.eu/sites/default/files/analytical\\_report\\_n9\\_economic\\_benefits\\_of\\_open\\_data.pdf](https://www.europeandataportal.eu/sites/default/files/analytical_report_n9_economic_benefits_of_open_data.pdf) [žiūrėta 2020-01-04].
3. A World Bank Group Flagship Report, „Doing Business 2017“, (2017)  
URL:<http://www.doingbusiness.org/content/dam/doingBusiness/media/Annual-Reports/English/DB17-Report.pdf> [žiūrėta 2019-05-01].
4. Ali M. Al-Khouri, Open Data: A Paradigm Shift in the Heart of Government, (2014),  
URL:<http://www.macrothink.org/journal/index.php/jpag/article/view/6407/52> [žiūrėta 2018-12-19].
5. ASEDIE, Asedie Infomediary Sector Report (2017)  
URL:<http://www.asedie.es/assets/asedie-infomediary-sector-report-2017-.pdf> [žiūrėta 2019-01-01].
6. Belgijos atvirų duomenų portalas  
URL: <https://data.gov.be/en> [žiūrėta 2019-04-16].
7. D W. Carrara ir kt., Creating Value through Open Data, 2015,  
URL:[http://www.europeandataportal.eu/sites/default/files/edp\\_creating\\_value\\_through\\_open\\_data\\_0.pdf](http://www.europeandataportal.eu/sites/default/files/edp_creating_value_through_open_data_0.pdf) [žiūrėta 2018-12-18].
8. European Commission, Introduction to Linked Data, (2014)  
URL:[https://www.europeandataportal.eu/sites/default/files/d2.1.2\\_training\\_module\\_1.2\\_introduction\\_to\\_linked\\_data\\_en\\_edp.pdf](https://www.europeandataportal.eu/sites/default/files/d2.1.2_training_module_1.2_introduction_to_linked_data_en_edp.pdf) [žiūrėta 2018-12-29].
9. European Commission, Įvadas į susietuosius duomenis, (2013)  
URL:[https://www.europeandataportal.eu/sites/default/files/d2.1.2\\_training\\_module\\_1.2\\_introduction\\_to\\_linked\\_data\\_lt\\_edp.pdf](https://www.europeandataportal.eu/sites/default/files/d2.1.2_training_module_1.2_introduction_to_linked_data_lt_edp.pdf) [žiūrėta 2019-01-13].
10. European Commission, The European Data Market Study: Final Report. (2017)  
URL:[https://sites.google.com/a/open-evidence.com/download/repository/SMART20130063\\_Final%20Report\\_030417\\_2.pdf?atredirects=0&d=1](https://sites.google.com/a/open-evidence.com/download/repository/SMART20130063_Final%20Report_030417_2.pdf?atredirects=0&d=1) [žiūrėta 2019-01-01].

11. European commission: Open data An engine for innovation, growth and transparent governance (2011)  
URL:[http://www.europarl.europa.eu/RegData/docs\\_autres\\_institutions/commission\\_euop/enne/com/2011/0882/COM\\_COM\(2011\)0882\\_EN.pdf](http://www.europarl.europa.eu/RegData/docs_autres_institutions/commission_euop/enne/com/2011/0882/COM_COM(2011)0882_EN.pdf) [žiūrėta 2019-04-17].
12. European data portal, Some insights in the most popular Open Data categories, (2016)  
URL:<https://www.europeandataportal.eu/en/highlights/some-insights-most-popular-open-data-categories> [žiūrėta 2018-12-30].
13. Impact assessment of odine programme, (2017)  
URL:[https://opendataincubator.eu/files/2017/07/ODINE\\_Final-report\\_2.0.pdf](https://opendataincubator.eu/files/2017/07/ODINE_Final-report_2.0.pdf) [žiūrėta 2019-01-02].
14. Informacinės visuomenės plėtros komitetas prie Susisiekimo ministerijos: „Galimybių tyrimo, skirto apibrėžti atvirų duomenų iniciatyvos įgyvendinimo Lietuvoje architektūros modelį, ataskaita“, (2015)  
URL:[https://ivpk.lrv.lt/uploads/ivpk/documents/files/IVPK\\_leidiniai/Galimybi%C5%B3%20tyrimo%20ataskaita\\_atviri%20duomenys%202015.pdf](https://ivpk.lrv.lt/uploads/ivpk/documents/files/IVPK_leidiniai/Galimybi%C5%B3%20tyrimo%20ataskaita_atviri%20duomenys%202015.pdf) [žiūrėta 2019-04-28].
15. Informacinės visuomenės plėtros komiteto duomenų atvėrimo koordinavimas  
URL:<https://ivpk.lrv.lt/lt/veiklos-sritys-1/duomenu-atverimo-koordinavimas> [žiūrėta 2019-12-26].
16. J. Manyika ir kt., Open Data: Unlocking Innovation and Performance with Liquid Information, tech. rep. October, McKinsey, (2013)  
URL:<https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information> [žiūrėta 2018-12-18].
17. J.Ovenden, Are companies using their external data?  
URL:<https://channels.theinnovationenterprise.com/articles/are-companies-using-their-external-data> [žiūrėta 2018-12-29].
18. J.Tauberer, Open Government Data Definition: The 8 Principles of Open Government Data  
URL: <https://opengovdata.io/2014/8-principles/> [žiūrėta 2018-12-30];
19. JAV atvirų duomenų portalas  
URL: <https://www.data.gov/> [žiūrėta 2019-04-16].
20. Jorn Berends ir kt., Re-using Open Data: A study of companies transforming Open Data into economic and societal value, (2017)  
URL:[https://www.europeandataportal.eu/sites/default/files/re-using\\_open\\_data.pdf](https://www.europeandataportal.eu/sites/default/files/re-using_open_data.pdf) [žiūrėta 2018-12-18].
21. Juddie Attard ir kt., Strategies and Approaches for Exploiting the value of Open Data, (2016)

- URL:[https://www.researchgate.net/publication/281349915\\_A\\_Systematic\\_Review\\_of\\_Open\\_Government\\_Data\\_Initiatives](https://www.researchgate.net/publication/281349915_A_Systematic_Review_of_Open_Government_Data_Initiatives) [žiūrėta 2018-12-20].
22. Laura James, Defining Open Data, (2013)  
URL: <https://blog.okfn.org/2013/10/03/defining-open-data/> [žiūrėta 2018-12-19].
23. Lietuvos atvirų duomenų portalas  
URL: <http://opendata.gov.lt> [žiūrėta 2019-04-16];
24. LR Valstybės tarnybos departamentas prie Lietuvos Respublikos vidaus reikalų ministerijos: Mokslininkai ir praktikai diskutavo apie atvirąją valdžią, (2016)  
URL:<https://vtd.lrv.lt/lt/naujienos/mokslininkai-ir-praktikai-diskusavo-apie-atviraja-valdzia> [žiūrėta 2019-04-25].
25. M. Janssen, Y. Charalabidis and A. Zuiderwijk, Benefits, Adoption Barriers and Myths of Open Data and Open Government, (2012)  
URL:<https://repository.tudelft.nl/islandora/object/uuid:952c8b0e-c55e-4443-9b41-14c38806a840?collection=research> [žiūrėta 2019-01-02].
26. Michiel S de Vries: Towards Citizen-associations for Participatory Decision-making, (2016)  
URL:<https://drive.google.com/file/d/0ByPi-FecRfyQdIR6X1phbnU1QU0/view> [žiūrėta 2019-04-24].
27. N. Cukier and V. Mayer-Schoenberger, The Rise of Big Data, Foreign Affairs (2013)  
URL: <https://www.foreignaffairs.com/articles/2013-04-03/rise-big-data> [žiūrėta 2019-01-02].
28. N. Huijboom ir kt., Open data: an international comparison of strategies, 2011  
URL:<http://unpan1.un.org/intradoc/groups/public/documents/UN-DPADM/UNPAN046727.pdf> [žiūrėta 2018-12-29].
29. ONTSI, „Characterization Study of the Infomediary Sector, 2014  
URL:[https://www.ontsi.red.es/ontsi/sites/ontsi/files/executive\\_summary\\_public\\_infomediary\\_sector\\_2014.pdf](https://www.ontsi.red.es/ontsi/sites/ontsi/files/executive_summary_public_infomediary_sector_2014.pdf) [žiūrėta 2019-04-28].
30. Open data impact map  
URL: <https://opendataimpactmap.org/index> [žiūrėta 2019-05-26].
31. Open Government Data Definition: The 8 Principles of Open Government Data  
URL: <https://opengovdata.io/2014/8-principles/> [žiūrėta 2020-01-04].
32. Open knowledge international: Open data Handbook  
URL: <http://opendatahandbook.org/> [žiūrėta 2018-12-19].
33. Platforma „Credit reform“  
URL: <https://www.cr.lt/> [žiūrėta 2019-05-02].

34. Platforma „Creditinfo“  
URL: <https://lt.creditinfo.com> [žiūrėta 2019-05-01].
35. Platforma „Kur gyvenu?“  
URL: <https://kurgyvenu.lt> [žiūrėta 2019-04-28].
36. Simon Chignard, A brief history of Open Data, (2013)  
URL: <http://parisinnovationreview.com/articles-en/a-brief-history-of-open-data> [žiūrėta 2018-12-19].
37. Singapūro atvirų duomenų portalas  
URL: <https://data.gov.sg/search> [žiūrėta 2019-04-16].
38. Somalio atvirų duomenų portalas  
URL: <http://www.opendataforsomalia.org/> [žiūrėta 2019-04-16].
39. T. Davenport, L. Prusak, Working Knowledge: How Organizations Manage What They Know, (1998)  
URL: [https://www.researchgate.net/publication/229099904\\_Working\\_Knowledge\\_How\\_Organizations\\_Manage\\_What\\_They\\_Know](https://www.researchgate.net/publication/229099904_Working_Knowledge_How_Organizations_Manage_What_They_Know) [žiūrėta 2019-01-12].
40. T. Berners-Lee, Linked data, (2006)  
URL: <https://www.w3.org/DesignIssues/LinkedData.html> [žiūrėta 2019-01-13];
41. T. Walters, Incorporating External Data Into the Data Warehouse  
URL: <http://www2.sas.com/proceedings/sugi22/DATAWARE/PAPER116.PDF> [žiūrėta 2018-12-29].
42. Tim Berners-Lee: The next Web of open, linked data, (2009)  
URL: [https://www.youtube.com/watch?v=OM6XIICm\\_qo](https://www.youtube.com/watch?v=OM6XIICm_qo) [žiūrėta 2018-12-29].
43. UK Government: Open Data White Paper: Unleashing the potential  
URL: [https://data.gov.uk/sites/default/files/Open\\_data\\_White\\_Paper.pdf](https://data.gov.uk/sites/default/files/Open_data_White_Paper.pdf) [žiūrėta 2018-12-20].
44. Verslo žinios, „Creditinfo“ vadovas: Registrų centro įkainiai yra per dideli“  
URL: <https://www.vz.lt/verslo-aplinka/2018/09/11/creditinfo-vadovas-registru-centro-ikainiai-yra-per-dideli#ixzz5mV9sOVpV> [žiūrėta 2019-05-01].
45. Vyriausioji tarnybinė etikos komisija: atviri duomenys  
URL: <https://www.vtek.lt/index.php/atviri-duomenys> [žiūrėta 2019-04-24].

# **DETERMINING THE NEED FOR DATA OPENING FOR THE LITHUANIAN PUBLIC AND PRIVATE SECTORS**

**Karolis Laucė**

**Paper for Master's degree**

*Strategic information systems management Master's program*

Vilnius University, Faculty of Economics, Economic informatics Department

Supervisor – Asist. dr. **M. Kazimianec**

Vilnius, 2020

## **Summary**

73 pages, 27 charts, 13 pictures, 45 references.

The use of open data is growing at a high rate every year. It is estimated that the use of open data can generate and save about \$ 3 trillion every year worldwide and create about 13,000 new open data job openings by 2020. Currently, the topic of using openly accessible data is particularly relevant.

Master's thesis consists of three parts: literature analysis, methodology and research with research results, conclusions and recommendations.

The first part of the thesis, based on the scientific literature analysis which describes open data, the benefits of open data and programs to promote data opening.

The second part describes the methodology and workflow.

The third part of the thesis analyzes the types of open data in the world, discusses collected and classified data sources and sets in Lithuania, analyzes and describes the open data usage cases in Lithuania, analyzes the open data types in Lithuania and other countries, it is accompanied by a study of associations of use of data types.

The results revealed that only a few types of data sets are used in Lithuania: statistical, geographic and transport data and therefore the diversity of projects developed or under development in Lithuania is very limited, unlike in other countries of the world. Based on the results of the association study, it can be assumed that the use of open data is not ensured in Lithuania - the value of the data is too low and therefore some of the open data is not reused.



According to Juergen Mueller, SAP board member, the value of the data is determined using the following formula:

$$\text{Data Amount} * \text{Data Quality} * \text{Data Usage} = \text{Data Value}$$

It is likely that at least one of the components of the data value is of poor quality, which makes the whole value of the data low. Low value open data is much less common or not used at all.

## PRIEDAI

Consequent	Antecedent	Support %	Confidenc...	Rule Supp...	Lift	Deployability
Moksliniai ir švietimo	Statistiniai	81,481	47,727	38,889	1,171	42,593
Finansiniai	Statistiniai	81,481	47,727	38,889	1,121	42,593
Geografiniai	Statistiniai	81,481	43,182	35,185	0,897	46,296
Aplinkos	Statistiniai	81,481	29,545	24,074	0,84	57,407
Moksliniai ir švietimo Finansiniai	Statistiniai	81,481	22,727	18,519	1,227	62,963
Geografiniai Aplinkos	Statistiniai	81,481	20,455	16,667	0,789	64,815
Statistiniai	Geografiniai	48,148	73,077	35,185	0,897	12,963
Aplinkos	Geografiniai	48,148	53,846	25,926	1,53	22,222
Moksliniai ir švietimo	Geografiniai	48,148	34,615	16,667	0,85	31,481
Statistiniai Aplinkos	Geografiniai	48,148	34,615	16,667	1,438	31,481
Finansiniai	Geografiniai	48,148	30,769	14,815	0,722	33,333
Moksliniai ir švietimo Statistiniai	Geografiniai	48,148	30,769	14,815	0,791	33,333
Finansiniai Statistiniai	Geografiniai	48,148	30,769	14,815	0,791	33,333
Metrologiniai	Geografiniai	48,148	26,923	12,963	2,077	35,185
Transporto	Geografiniai	48,148	26,923	12,963	1,615	35,185
Statistiniai Metrologiniai	Geografiniai	48,148	23,077	11,111	2,077	37,037
Metrologiniai Aplinkos	Geografiniai	48,148	23,077	11,111	2,077	37,037
Statistiniai	Finansiniai	42,593	91,304	38,889	1,121	3,704
Moksliniai ir švietimo	Finansiniai	42,593	43,478	18,519	1,067	24,074

Ne Lietuvoje naudojamų atvirų duomenų rūšys pagal aukščiausią paramos rodiklį

Consequent	Antecedent	Support %	Confidenc...	Rule Supp...	Lift	Deployability
Statistiniai	Geografiniai	75,0	66,667	50,0	0,889	25,0
Geografiniai	Statistiniai	75,0	66,667	50,0	0,889	25,0
Transporto	Geografiniai	75,0	41,667	31,25	1,111	43,75
Transporto	Statistiniai	75,0	41,667	31,25	1,111	43,75
Statistiniai Transporto	Geografiniai	75,0	33,333	25,0	1,067	50,0
Finansiniai	Statistiniai	75,0	33,333	25,0	1,333	50,0
Geografiniai Transporto	Statistiniai	75,0	33,333	25,0	1,067	50,0
Aplinkos	Geografiniai	75,0	25,0	18,75	1,333	56,25
Moksliniai i...	Geografiniai	75,0	16,667	12,5	1,333	62,5
Metrologiniai	Geografiniai	75,0	16,667	12,5	1,333	62,5
Moksliniai i...	Statistiniai	75,0	16,667	12,5	1,333	62,5
Metrologiniai	Statistiniai	75,0	16,667	12,5	1,333	62,5
Aplinkos	Statistiniai	75,0	16,667	12,5	0,889	62,5
Transporto	Geografiniai Statistiniai	50,0	50,0	25,0	1,333	25,0
Geografiniai	Transporto	37,5	83,333	31,25	1,111	6,25
Statistiniai	Transporto	37,5	83,333	31,25	1,111	6,25
Geografiniai Statistiniai	Transporto	37,5	66,667	25,0	1,333	12,5
Metrologiniai	Transporto	37,5	16,667	6,25	1,333	31,25
Statistiniai	Geografiniai Transporto	31,25	80,0	25,0	1,067	6,25
Geografiniai	Statistiniai Transporto	31,25	80,0	25,0	1,067	6,25

Lietuvoje naudojamų atvirų duomenų rūšys pagal aukščiausią paramos rodiklį

Consequent	Antecedent	Support %	Confidenc...	Rule Supp...	Lift	Deployability
Statistiniai	Moksliniai ir švietimo	40,741	95,455	38,889	1,171	1,852
Statistiniai	Finansiniai	42,593	91,304	38,889	1,121	3,704
Moksliniai ir švietimo	Statistiniai	81,481	47,727	38,889	1,171	42,593
Finansiniai	Statistiniai	81,481	47,727	38,889	1,121	42,593
Statistiniai	Geografiniai	48,148	73,077	35,185	0,897	12,963
Geografiniai	Statistiniai	81,481	43,182	35,185	0,897	46,296
Geografiniai	Aplinkos	35,185	73,684	25,926	1,53	9,259
Aplinkos	Geografiniai	48,148	53,846	25,926	1,53	22,222
Statistiniai	Aplinkos	35,185	68,421	24,074	0,84	11,111
Aplinkos	Statistiniai	81,481	29,545	24,074	0,84	57,407
Statistiniai	Moksliniai ir švietimo Finansiniai	18,519	100,0	18,519	1,227	0,0
Finansiniai	Moksliniai ir švietimo Statistiniai	38,889	47,619	18,519	1,118	20,37
Moksliniai ir švietimo	Finansiniai Statistiniai	38,889	47,619	18,519	1,169	20,37
Finansiniai	Moksliniai ir švietimo	40,741	45,455	18,519	1,067	22,222
Finansiniai Statistiniai	Moksliniai ir švietimo	40,741	45,455	18,519	1,169	22,222
Moksliniai ir švietimo	Finansiniai	42,593	43,478	18,519	1,067	24,074
Moksliniai ir švietimo Statistiniai	Finansiniai	42,593	43,478	18,519	1,118	24,074
Moksliniai ir švietimo Finansiniai	Statistiniai	81,481	22,727	18,519	1,227	62,963
Geografiniai	Statistiniai Aplinkos	24,074	69,231	16,667	1,438	7,407

Ne Lietuvoje naudojamų atvirų duomenų rūšys pagal aukščiausią taisyklės paramos rodiklį

Consequent	Antecedent	Support %	Confidenc...	Rule Supp...	Lift	Deployability
Geografiniai	Statistiniai	75,0	66,667	50,0	0,889	25,0
Geografiniai	Transporto	37,5	83,333	31,25	1,111	6,25
Statistiniai	Transporto	37,5	83,333	31,25	1,111	6,25
Transporto	Geografiniai	75,0	41,667	31,25	1,111	43,75
Transporto	Statistiniai	75,0	41,667	31,25	1,111	43,75
Statistiniai	Finansiniai	25,0	100,0	25,0	1,333	0,0
Statistiniai	Geografiniai Transporto	31,25	80,0	25,0	1,067	6,25
Geografiniai	Statistiniai Transporto	31,25	80,0	25,0	1,067	6,25
Geografiniai Statistiniai	Transporto	37,5	66,667	25,0	1,333	12,5
Transporto	Geografiniai Statistiniai	50,0	50,0	25,0	1,333	25,0
Statistiniai Transporto	Geografiniai	75,0	33,333	25,0	1,067	50,0
Finansiniai	Statistiniai	75,0	33,333	25,0	1,333	50,0
Geografiniai Transporto	Statistiniai	75,0	33,333	25,0	1,067	50,0
Geografiniai	Aplinkos	18,75	100,0	18,75	1,333	0,0
Aplinkos	Geografiniai	75,0	25,0	18,75	1,333	56,25
Geografiniai	Moksliniai ir švieti...	12,5	100,0	12,5	1,333	0,0
Statistiniai	Moksliniai ir švieti...	12,5	100,0	12,5	1,333	0,0
Aplinkos	Moksliniai ir švieti...	12,5	100,0	12,5	5,333	0,0
Geografiniai	Metrologiniai	12,5	100,0	12,5	1,333	0,0
Statistiniai	Metrologiniai	12,5	100,0	12,5	1,333	0,0

Lietuvoje naudojamų atvirų duomenų rūšys pagal aukščiausią taisyklės paramos rodiklį

Consequent	Antecedent	Support %	Confidence %	Deployability
Aplinkos	Statistiniai	81,481	29,545	57,407
Geografiniai	Statistiniai	81,481	43,182	46,296
Moksliniai ir šviet...	Statistiniai	81,481	47,727	42,593
Finansiniai	Statistiniai	81,481	47,727	42,593
Metrologiniai	Geografiniai	48,148	26,923	35,185
Transporto	Geografiniai	48,148	26,923	35,185
Finansiniai	Geografiniai	48,148	30,769	33,333
Moksliniai ir šviet...	Geografiniai	48,148	34,615	31,481
Transporto	Finansiniai	42,593	26,087	31,481
Aplinkos	Finansiniai	42,593	30,435	29,63
Aplinkos	Moksliniai ir šviet...	40,741	27,273	29,63
Geografiniai	Finansiniai	42,593	34,783	27,778
Transporto	Aplinkos	35,185	26,316	25,926
Geografiniai	Moksliniai ir šviet...	40,741	40,909	24,074
Moksliniai ir šviet...	Finansiniai	42,593	43,478	24,074
Moksliniai ir šviet...	Aplinkos	35,185	31,579	24,074
Metrologiniai	Aplinkos	35,185	31,579	24,074
Finansiniai	Aplinkos	35,185	36,842	22,222
Aplinkos	Geografiniai	48,148	53,846	22,222
Finansiniai	Moksliniai ir šviet...	40,741	45,455	22,222
Kultūros	Transporto	16,667	22,222	12,963
Statistiniai	Geografiniai	48,148	73,077	12,963
Statistiniai	Aplinkos	35,185	68,421	11,111

Ne Lietuvoje naudojamų atvirų duomenų rūšys pagal aukščiausią įdiegiamumo rodiklį

Consequent	Antecedent	Support %	Confidence %	Deployability
Moksliniai ir šviet...	Kultūros	100,0	12,5	87,5
Metrologiniai	Kultūros	100,0	12,5	87,5
Aplinkos	Kultūros	100,0	18,75	81,25
Finansiniai	Kultūros	100,0	25,0	75,0
Transporto	Kultūros	100,0	37,5	62,5
Moksliniai ir šviet...	Geografiniai	75,0	16,667	62,5
Metrologiniai	Geografiniai	75,0	16,667	62,5
Moksliniai ir šviet...	Statistiniai	75,0	16,667	62,5
Metrologiniai	Statistiniai	75,0	16,667	62,5
Aplinkos	Statistiniai	75,0	16,667	62,5
Aplinkos	Geografiniai	75,0	25,0	56,25
Finansiniai	Statistiniai	75,0	33,333	50,0
Transporto	Geografiniai	75,0	41,667	43,75
Transporto	Statistiniai	75,0	41,667	43,75
Metrologiniai	Transporto	37,5	16,667	31,25
Geografiniai	Kultūros	100,0	75,0	25,0
Statistiniai	Kultūros	100,0	75,0	25,0
Statistiniai	Geografiniai	75,0	66,667	25,0
Geografiniai	Statistiniai	75,0	66,667	25,0
Geografiniai	Transporto	37,5	83,333	6,25
Statistiniai	Transporto	37,5	83,333	6,25

Lietuvoje naudojamų atvirų duomenų rūšys pagal aukščiausią įdiegiamumo rodiklį