

LATVIJAS UNIVERSITĀTE  
DATORIKAS FAKULTĀTE

# ATVĒRTIE DATI LATVIJĀ

BAKALaura DARBS

Autors: **Matīss Jurgelis**  
Studenta apliecības Nr.: mj10142  
Darba vadītājs: dr.dat. Uldis Bojārs

RĪGA 2014

## **ANOTĀCIJA**

Atvērto datu kustība ir relatīvi jauna, taču savā pastāvēšanas laikā tā jau paspējusi pierādīt, ka spēj sniegt tādus ieguvumus, kā valsts pārvaldes caurspīdīguma palielināšana, finanšu plūsmas palielinājums un jaunu darba vietu radīšana. Pasaules pieredze liecina, ka publiskā sektora datos slēpjas neizmantots potenciāls, kas sniedz teju neierobežotas iespējas. Arī Latvija pasaules pozitīvās pieredzes rezultātā pārņem šīs kustības pamatprincipus.

Šī darba ietvaros tiek apskats, kas ir atvērtie dati, kam tie nepieciešami un kāds ir Latvijas atvērto datu avotu saturs un tā kvalitāte uzsvāru liekot uz pozitīvajiem piemēriem. Tāpat tiek izvirzīti dažādi kvalitātes kritēriji, kas palīdz saprast cik vērtīgs ir datu avots.

**Atslēgvārdi:** atvērtie dati, atkalizmantojami dati, valsts pārvaldes caurspīdība

## **ABSTRACT**

Open data movement is relatively new, but in its age it has proven that it can give such benefits as government transparency, financial flow increase and new vacancy creating. World's experience proves that public sector data is hiding unused potential, which holds almost unlimited opportunities. Also, Latvia is taking its step to meet open data principles inspired from world's positive experience.

In this work's context there will be aim to overview what is open data, who needs it and what is Latvia's open data source and its quality. Also, there will be quality guides defined which will help detect condition of the source.

**Keywords:** open data, reusing of data, government transparency

# SATURS

APZĪMĒJUMU SARAKSTS .....	5
IEVADS .....	6
1 ATVĒRTIE DATI.....	7
1.1 Kas ir atvērtie dati .....	7
1.2 Atvērtu datu kvalitāte .....	8
1.3 Kam tas nepieciešams .....	9
1.4 Atvērtie dati pasaulē .....	10
1.5 Publiskā sektora informācija .....	11
1.5.1 PSI pamatnostādnes 2014.-2020. gadam.....	12
2 LATVIJAS ATVĒRTO DATU AVOTI.....	14
2.1 Latvijas atvērtu datu entuziasti .....	14
2.2 Atvērtu datu kvalitātes uzlabošana.....	15
2.3 Latvijā sastopamās datu kopas.....	16
2.3.1 Ceļu satiksmes drošības direkcija.....	17
2.3.2 Uzņēmumu reģistrs .....	18
2.3.3 Rīgas domes izglītības, kultūras un sporta departaments .....	20
2.3.4 Rīgas domes mājokļu un vides departaments.....	22
2.3.5 Latvijas Republikas Iekšlietu ministrija.....	23
2.3.6 Latvijas universitāte.....	26
2.4 Potenciālie datu avoti, kas datus nepiedāvā.....	27
2.5 Atvērtu datu, kopas ieviešanas priekšrocības.....	28
2.6 Atvērtu datu, kopu trūkumi .....	29
3 TENDENCES.....	31
3.1 Datu uzlabošana.....	31
SECINĀJUMI .....	33
IZMANTOTĀ LITERATŪRA .....	34

## **APZĪMĒJUMU SARAKSTS**

**PDDL** – public domain dedication licence, kas ļauj izmantot ar šo licenci aizsargātus datus kā privātiem nolūkiem tā arī komerciāliem, neierobežojot to kopēšanu un modificēšanu

**CC0** – Creative Commons 0 licence, kas ļauj datus brīvi modificēt, kopēt, izmantot tos privātiem vai komerciāliem mērķiem

**VARAM** – Vides aizsardzības un reģionālās attīstības ministrija

**IKT** – informāciju un komunikāciju tehnoloģijas

**CSDD** – ceļu satiksmes drošības direkcija

**UR** – uzņēmumu reģistrs

**ES** – Eiropas savienība

**PSI** – publiskā sektora informācija

**RDIKSD** – Rīgas domes izglītības, kultūras un sporta departaments

**RDMVD** – Rīgas domes mājokļu un vides departaments

**IM** – Iekšlietu ministrija

**LRIMIC** – Latvijas Republikas Iekšlietu ministrijas Informācijas centrs

**LU** – Latvijas universitāte

**HTML** – hiperteksta iezīmēšanas valoda, ko izmanto tīmekļa lapu attēlošanai pārlūkprogrammās

## **IEVADS**

Darba ietvaros tiks apskatīta atvērto datu kustība, kas ir relatīvi jauna. Par mērķiem tiek izvirzīts izpētīt, kas ir atvērtie dati, kā tos pieliekto, kam tie ir paredzēti un kādu ieguvumu varētu gūt sabiedrība no atvērto datu esamības. Pēc šo faktu noskaidrošanas plānots analizēt vairākus Latvijas piemērus, kur atvērtie dati ir pieejami publiski. Uzsvars tiks likts uz publiskā sektora atvērtajiem datiem, kuri ir valsts iestāžu īpašumā. Tāpat tiks apskatīti tādi datu avoti, kuros dati ir sastopami, izvairoties no gadījumiem, kur dati nav pieejami vispār, vai to saturs neatbilst atvērto datu definīcijai, tādā veida izvairoties no pētījuma par sliktiem vai neesošiem datu avotiem.

Katrs no pētījumā analizētajiem datu avotiem tiks vērtēts pēc iepriekš izvirzītiem kritērijiem, kas kopumā dos redzējumu par konkrētā datu avota kvalitātes līmeni.

Noslēgumā plānots arī apskatīt iespējamus datu attīstības un uzlabošanas scenārijus, vadoties, pēc veiktajā analīzē sastaptajiem avotu trūkumiem. Tādā veidā arī norādot uz faktoriem, kas būtu jāuzlabo esošajiem datu avotiem, kā arī jāņem vērā tiem kas plāno radīt datu avotu, par paraugu ņemot kādu no esošajiem risinājumiem.

# 1 ATVĒRTIE DATI

Katru dienu cilvēce rada milzīgu datu daudzumu[1]. Pēc “IBM” veiktā pētījuma ik dienu tiek radīti aptuveni  $2,5 \times 10^{18}$  baiti, kas ir milzīgs apjoms[1]. Šeit rodas pamatotas pārdomas par to cik ilgs ir radīto datu mūžs un kā iespējams palielināt datu vērtību. ES komisijas viceprezidente Neelie Kroes, kas atbildīga par ES digitālo programmu apgalvo: “yes, data really is the new oil” atvērtie dati ir jaunā nafta[12]. To pierāda vairāki veiksmīgi projekti, kas ļauj esošus datus izmantot vairākkārt tādā veidā no datiem iegūstot papildus labumu. Padarot sabiedrībai pieejamus dažādus datus par valsts pārvaldes iepirkumiem, budžetu, Saeimas balsojumiem var panākt lielāku nodokļu maksātāju izpratni par valstī notiekošajiem procesiem, kā arī atvērtie dati arī ir viens no veidiem, kā nodrošināt caurspīdīgu valsts pārvaldi. Tāpat atvērtie dati varētu radīt jaunas biznesa iespējas, kā arī palīdzēt uzlabot esošus servissus balstot tos uz pierādījumiem. Sekmējot valsts iekšējo saimniecisko darbību tiek sekmēta arī finanšu aprīte, likumsakarīgi palielina nomaksātos nodokļus, tādā veidā pozitīvi sildot kopējo valsts ekonomisko stāvokli.

## 1.1 Kas ir atvērtie dati

Atvērto datu pamatā ir vairākas idejas, kas definē atvērtos datus[2]. Open definition saka, ka, lai datus varētu uzskatīt par atvērtiem, tiem jāatbilst šādiem kritērijiem:

1. tiem jābūt publiski un brīvi pieejami visiem;
2. licence nedrīkst ierobežot tiesības datus pārdot tālāk vienus vai, kā citu lielāku datu sastāvdaļu;
3. licence nedrīkst aizliegt mainīt vai atvasināt datus;
4. datu formāts nedrīkst ierobežot piekļuvi pašiem datiem, formātam jābūt atvērtam, un nepiesaistītam kādai maksas programmai;
5. licence var pieprasīt atsaukties uz oriģinālajiem datu publicētājiem vai personām, kas piedalījušās datu sagatavošanā;
6. licence var pieprasīt, lai datu kopas, nosaukums vai versijas numurs atšķiras no oriģināla, ja tajā ir veiktas izmaiņas;
7. licence nedrīkst diskriminēt kādu personu vai personu grupu;
8. licence nedrīkst aizliegt apstrādāt datus un tos piedāvāt tālāk par maksu;
9. licence attiecās uz visām pusēm, kas iesaistītas datu apstrādē un tā nevar prasīt papildus licenci šīm pusēm;

10. licencei jāattiecas uz visu datu kopu un tā nedrīkst prasīt atsevišķu licenci kādai datu apakškopai;

11. licence nedrīkst aizliegt vai ierobežot datus, kas tiek izplatīti kopā ar atvērtajiem datiem.

Atvērto datu definīcija radusies salīdzinoši nesen, un tā visa pamatā varētu būt bijusi nepieciešamību pēc atvērtības un vēlme datiem piešķirt iespējami lielāku vērtību. Atvērto datu pamatnostādnes veicina datu atkal izmantošanu, kas nodrošina garāku, to dzīves ciklu.

Atvērto datu mērķis ir sniegt publisku piekļuvi datiem, tādā veidā nodrošinot informāciju, kuru vēlāk iespējams apstrādāt vai saistīt ar kādu citu informāciju un no tā uzbūvēt jaunas lietotnes, servisu vai izdarīt secinājumus, kas palīdzētu attīstīties balstoties uz pieredzi, kas iegūta no datiem.

## 1.2 Atvērto datu kvalitāte

Lai nodrošinātu datu kopas pienācīgu kvalitāti un būtu iespējams novērtēt kvalitātes pakāpi, ir svarīgi noteikt atbilstošas kvalitātes metrikas. Šādas metrikas ir ieteicis pielietot tīmekļa atklājējs Tim Berners-Lee[4]. Atvērto datu kvalitātes metrikas definē, pie kādiem standartiem jāturas datiem, lai tie būtu iespējami atvērtāki un līdz ar to arī vērtīgāki. Metrikas iedalītas 5 līmeņos jeb zvaigznēs, kur katrs nākamais ievieš jaunas prasības, bet tajā pašā laikā garantē, ka kvalitāte atbilst arī visiem iepriekšējiem līmeņiem. Kvalitātes iedalītas šādos līmeņos:

1. dati ir brīvi pieejami jebkuram tīmeklī, datu formātam nav nozīmes un tie publicēti ar atvērtu licenci (piem. PDDL vai CC0);
2. dati ir strukturēti;
3. dati ir pieejami atvērtā, ne-patentētā formātā piemēram CSV;
4. datiem iespējams piekļūt izmantojot unikālu URI, kas ļauj jebkuram norādīt uz jūsu datiem;
5. dati ir sasaistīti ar citiem tīmeklī esošiem datu avotiem, kas nodrošina papildus kontekstu.

Ievērojot visas vai vismaz daļu no augstāk minētajām metrikām dati kļūst mašīnlasāmi un ērtāk lietojami. Tāpat tiek atvieglots darbs ar datu apstrādi un iegūšanu, kas ir krietni samazinās pavadīto laiku pie datu sagatavošanas pirms tos varēs iekļaut kādā projektā un likt tiem strādāt.

### 1.3 Kam tas nepieciešams

Mūsdienās datus apkopo ļoti daudzas valstiskas un privātās institūcijas[3], piemēram par Saeimas deputātu balsojumiem, ceļu satiksmes negadījumiem, gaisa piesārņojumu, transporta sastrēgumiem ielās un auto plūsmas intensitāti pa ielām. Dažas no iestādēm arī labprātīgi publisko ievāktu informāciju. Šāda datu publiskošana aizņem laiku, ko labprātīgi velta iestāde un par to nesaņem neko pretī. Šeit rodas vairāki jautājumi, “kurš no šādu datu publiskošanas var gūt labumu”, “kāpēc tas vispār nepieciešams” utt. Iedziļinoties un analizējot uzreiz top skaidrs, ka pieeja šādiem datiem no pēc iespējas vairāk nozarēs paver plašas iespējas, kas ir saistoši kā datu publiskotājam tā sabiedrībai kopumā. Tā piemēram, ja ir pieeja noziedzīgu nodarījumu datiem uz Rīgas ielām un gaisa piesārņojumam pa Rīgas ielām vai mikro rajoniem, būtu iespējams izplānot drošāko un pēc gaisa tīrības svaigāko vakara skrējieni vai pastaigu. Tāpat norādot maršrutu varētu novērtēt cik droši veselībai ir pārvietoties pa ielām un ar kādiem potenciāliem riskiem varētu saskarties ejot pa kādu konkrētu maršrutu. Apkopojot tādu informāciju būtu lietderīgi, to pievienot arī navigāciju iekārtās, kas plāno maršrutu arī kājām gājējiem un tādā veidā gājējus varētu vest pa iespējami drošākām vietām. Pievienojot klāt vēl vairākus datu avotus paveras iespēja radīt milzīgas informatīvas, izglītojošas un uz analīzi vērstus risinājumus, kas var atvieglot ikdienu, novērst un plānot. Viss iepriekš minētais iespējams tikai tad, ja dati ir pietiekoši kvalitatīvi un tie ir mašīnlasāmi. Datus protams var pievienot arī manuāli, bet ar roku ievadītiem vai no PDF parsētiem datiem jāpatērē nesalīdzināmi vairāk laika tos pievienojot risinājumam.

Nevienam nav noslēpums, ka katrs cilvēks uz kādu lietu skatās individuāli – ar savu dzīves redzējumu, ar savu pieredzi un atziņām. Tieši tāpēc, ja dati ir publiski, katrs tajos var saskatīt un atrast kaut ko citu. Paskatīties uz tiem kaut kā savādāk un pielietot tos tā, kā viņš to redz interesantus vai noderīgus. Līdz ar to atvērto datu esamība ļauj jebkuram sabiedrības loceklim brīvi radīt dažādus pārskatus, analītiskus publicējumus, rīkus un lietotnes, kas sniegtu visiem noderīgu un saistošu informāciju.

Ne mazāk svarīgs aspekts demokrātiskā valstī ir valsts organizāciju caurspīdīgums[5], kas novērš korupciju un sniedz atbildes nodokļu maksātājiem, kuri finansē valstij pakļautās institūcijas. Līdz ar to ir pašsaprotami, ka valsts iestāžu apkopotajiem datiem būtu jābūt publiski pieejamiem. Atvērtie dati jau savā būtībā garantē datu publisku pieejamību, kas lieliski atbilst valsts iestāžu caurspīdības nodrošināšanai. Kā piemēru var minēt, Eiropas savienības tiesību aktus, kas ir pārtulkoti visās savienības valodās[13]. “Google Translate” tulkošanas dzinējs izmanto šos publiski pieejamos dokumentus, lai uzlabotu savus tulkošanas algoritmus[6], kā rezultātā uzlabojās dzinēja piedāvātais tulkošanas serviss. Tāpat atvērtie dati ir derīgi un pat

nepieciešami pašām valsts institūcijām. Par piemēru var minēt Nīderlandes izglītības ministriju, kas publiskoja visus ar izglītību saistītos datus[6] tiešsaistē un ar iespēju tos izmantot atkārtoti. Kopš tā brīža saņemto jautājumu daudzums ir krietni samazinājies, jo atbildes var atrast paša spēkiem tiešsaistē. Tādā veidā ir samazinājusies ministrijas darba slodze, līdz ar to samazinājušies arī izdevumi, turklāt ir vieglāk atbildēt arī uz pašlaik uzdotajiem jautājumiem, jo ir skaidrs, kur meklēt datus kuros slēpjas atbildes.

Ir skaidri redzams, ka atvērtajiem datiem ir potenciāls un tie var nest labumu sabiedrībai[12]. Tajā pašā laikā ir ļoti grūti prognozēt kādu, un cik lielu labumu nesīs atvērtie dati nākotnē[6], jo nevienam nav zināms kādi inovatīvi risinājumi tiks atklāti un kādi jauni paņēmieni būs iespējami.

## 1.4 Atvērtie dati pasaulē

Pasaules pieredze liecina par atvērto datu pozitīvo ietekmi uz ekonomiku un tās izaugsmi[16]. Atvērto datu kustība aizsākās, kā demokrātisks mērķis kuru uz priekšu virzīja ideja, ka valsts iestādēm būtu jāpublisko dati, kas ir to īpašumā un par kuru ievākšanu ir samaksājuši nodokļu maksātāji[17].

Valsts interesēs ir veicināt saimniecisko darbību valsts iekšienē un atvērto dati ir viens no veidiem, kā radīt jaunas darba vietas un biznesa iespējas. Apvienotās Karalistes ministru kabineta ministrs Francis Maude uzskata, ka atvērtie dati<sup>1</sup> un valdības caurspīdība ir starp 21. gadsimta spēcīgākajiem instrumentiem[16] valsts politikā.

Arī citās valstīs<sup>2</sup> ir dažādi piemēri, kā sekmēt atvērto datu kustību. Piemēram Kanādā februāra beigās tika rīkots 48 stundu programmatūras izstrādes pasākums jeb *hackatons*, kur vairāk kā 800 dalībnieki rada vairāk, kā 100 dažādus projektus izmantojot valdības atvērto datu portālu [data.gc.ca](http://data.gc.ca)[18]. Starp radītajiem projektiem tika izvēlēti 15 labākie, kas tika virzīti tālāk uz finālu, kur tie cīnījās par galveno balvu 25'000\$. Izstrādātāji, kas sanāk kopā šajos pasākumos radīja dažādas lietotnes, kas piedāvā jaunus servisu un pakalpojumus, kā arī dažādus risinājumus, kas ļauj veikt plašāku analīzi ar pieejamajiem datiem. Radītie produkti iespējams nekad neklūs par tādiem lejupielāžu rekordistiem, kā "Candy Crush Saga" taču šo produktu vērtība ir mērāma pēc pavisam citas vērtības skalas, kas sabiedrībai sniedz vairāk noderīgu funkciju.

---

<sup>1</sup> <http://data.gov.uk>

<sup>2</sup> <http://datahub.io/>

Kanādas federālās informāciju sistēmu nodaļas vadītāja Corinne Charette atzīst, ka Kanādas valdība apzinās *hakatona* ietekmi uz publiskā sektora informācijas pārvaldīšanu un izlaišanu, tāpēc Kanādas valdība aktīvi domā par atbalsta mehānismiem, kas sekmētu turpmāku progresu šajā sektorā. Šādu motīvu vadīta, Kanādas valdība izstrādā politisku instrumentu kopu, kas liktu valsts iestādēm regulāri strādāt pie to rīcībā esošo datu plānotas pievienošanas atvērto datu resursiem.

Pasaules valstu pieredze liecina, ka atvērto datu kustība ir atbalstāma un ar to saistāma pozitīva pieredze. Dažādu valstu valdības ir ieinteresētas palielināt pārvaldes caurspīdības līmeni, radīt jaunas darba vietas un dot iespēju radīt jaunas biznesa veidus padarot to īpašumā esošos datus publiski pieejamu apstrādei.

Tā piemēram ASV valdība ir viena no lielākajām datu publicētājiem, kur vienā tīmekļa vietnē<sup>3</sup> tiek piedāvāti dati par vairāk, kā 20 nozarēm, kas kopā sevī ietver aptuveni 400'000 dažādas datu kopas[19] tādos formātos, kā JSON, XML, CSV, HTML, TXT, XLS un XLSX. Publicētie dati ir par vis dažādākajām nozarēm – klimats, izglītība, enerģija, agro kultūrā, veselības aprūpe, laikapstākļi, likumi, okeāns, pilsētas un citas[20]. Tik bagāts datu avots liecina par ASV valdības ieinteresētību un patstāvīgu dabu pie šī sektora attīstības un kapacitātes palielināšanas. ASV prezidents B. Obama, arī atklāj, ka atvērto datu kustībai ir liela biznesa kapacitāte, kurai likumsakarīgi būs pozitīva ietekme uz valsts ekonomiku[17].

Vēl kāds atzīstams šaurāka spektra piemērs ir San Francisko pilsētas atvērto tīmekļa vietne, “San Francisco Data”<sup>4</sup>, kur tiek piedāvāti dati par pilsētas iekšējo infrastruktūru. Vietnē pieejami vairāk kā 500 dažādas datu kopas par tādām nozarēm, kā noziedzīgie nodarījumi, transports, kino teātri, infrastruktūra, uzņēmējdarbība un citām[21].

## 1.5 Publiskā sektora informācija

Latvijas Republikas Satversmes 100 pants nosaka “Ikvienam ir tiesības uz vārda brīvību, kas ietver tiesības **brīvi iegūt, paturēt un izplatīt informāciju**, paust savus uzskatus. Cenzūra ir aizliegta”[22]. Pants savā būtībā no “dienas kārtības” izslēdz jautājumu par to vai sabiedrībai ir tiesības uz valsts īpašumā esošajiem datiem. Kā papildus argumentu var minēt to, ka sabiedrība maksājot nodokļus jau ir apmaksājusi datu ievākšanu, kas ir valsts īpašumā.

Eiropas savienības kopējās interesēs ir veicināt publiskā sektora datu atkal izmantošanu un atvērto datu kustību kopumā. PSI atkal izmantošanu ES valstīs regulē Eiropas Parlamenta

---

<sup>3</sup> <https://www.data.gov>

<sup>4</sup> <https://data.sfgov.org/>

un padomes direktīva 2003/98/EK[7], kas papildināta ar šobrīd saistošu direktīvu 2013/37/ES[8]. Šo direktīvu mērķis ir rekomendēt pareizās pieejas, tāpēc cik ļoti PSI padarīt par atvērtajiem datiem ir katras valsts individuālā ziņā.

Svarīgākie grozījumi, ko nosaka ES direktīvas ir:

- visi dati, kas nav aizsargāti ar valsts tiesību aktiem ir atļauts atkalizmantojot;
- aktualizē jautājumu par licencēšanu liekot uzsvāru uz neierobežojošām licencēm;
- aicina datu avotus veidot ar mašīnlasāmiem un atvērtiem formātiem.

Atsaucoties uz šīm ES direktīvām Latvija, ir izstrādātas iekšējās pamatnostādnes 2014. – 2020. gadam[9], kas nosaka PSI atkalizmantošanu, kā vienu no pamatprincipiem e-sistēmu un pārvalžu plānošanā nākamajam posmam līdz 2020. gadam.

### ***1.5.1 PSI pamatnostādnes 2014.-2020. gadam***

Vides aizsardzības un reģionālās attīstības ministrijas uzraudzībā esošā darba grupa izstrādājusi informācijas sabiedrības attīstības pamatnostādnes 2014. – 2020. gadam ar nolūku turpināt attīstīt un stiprināt esošo politiku un informācijas sabiedrības attīstības sektorā. Tāpat plānots, ar šī dokumenta palīdzību noteikt informāciju un komunikāciju tehnoloģiju sektora prioritātes ES struktūrfondu plānošanas periodam līdz 2020. gadam. VARAM uzraudzībā esošās darba grupas sastāvā darbojušies pārstāvji no pilnīgi visām Latvijas ministrijām, UNESCO Latvijas nacionālās komisijas, IKT tehnoloģiju asociācijām, Latvijas tirdzniecības un rūpniecības kameras, Valsts kancelejas Darba devēju konfederācijas un Pašvaldību savienības. Pārstāvot tik plašu nozaru loku tiek veicināta iespējami kvalitatīvāku nostādņu izstrāde, kas ir saistoša no visu iesaistīto pušu viedokļa.

Dokumenta saturs liecina par iesaistīto pušu izpratni par potenciālo pienesumu, ko varētu sniegt pilnvērtīga valsts iestāžu rīcībā esošo datu izmantošana. Kā pozitīvs piemērs tiek minēta veiksmīga privātā sektora pieredze ar atvērtajiem datiem. Tāpat tiek aicināts izstrādāt vienotu un standartizētu piekļuvi atvērto datu kopām, tādā veidā atvieglojot jebkuras ieinteresētās personas darbu pie datu ieguves. Tiek uzskatīts arī, ka atvērto datu integrācija IKT, sekmēs sabiedrības pieprasījumu pēc dažādiem servisiem tajā skaitā lietotnēm, kuras pēc nepieciešamības varēs izstrādāt trešās puses tajā skaitā arī privātais sektors.

Valsts interesēs ir sekmēt izaugsmi, stimulējot dažādus procesus, tajā skaitā arī iesaistoties un iniciējot dažādas jaunas, perspektīvas iestrādes. Radot pamatnostādnes, kas orientētas uz atvērto datu integrēšanu IKT ierēdņi apzinās, ka iespējams palielināt valsts īpašumā esošo datu vērtību nododot to publiskai lietošanai. Tādā veidā stimulējot jaunu

komerc-produktu un pakalpojumu rašanu, zinātnieciskajā un pētnieciskajā darbā, sabiedrisko procesu analīzē, kas savukārt pozitīvi ietekmēs valsts ekonomiku pienesot jaunus nodokļu ieņēmumus. Šādā veidā no papildus ieņēmumiem valsts kasē varētu finansiāli sekmēt valsts īpašumā esošos atvērto datu servissus, kā arī radīt jaunus.

Lai arī nostādņu izstrāde ir atzīstama, nekur tekstā nav minēts, ka dati jāpiedāvā bez maksas vai par to sagatavošanu par noteiktu periodu tas jāveic bez maksas. Tādā veidā var rasties konflikta situācija, kad dati it kā ir pieejami taču par to jāmaksā, kas varētu radīt sabiedrības nepatiku.

VARAM elektroniskās pārvaldes departamenta vecākais konsultants, privātā e-pasta sarakstē ar autoru norāda, ka pamatnostādnes ir jāņem vērā plānojot e-pārvaldes sistēmas un tieši ar šo argumentu pašlaik tiek uzrunātas valsts iestādes, norādot, ka tas ir un būs jāievēro.

## 2 LATVIJAS ATVĒRTO DATU AVOTI

Veicot Latvijā reģistrētu domēnu izpēti tiek secināts, ka atvērto datu joma valstī ir sākuma stadijā. Pārsvārā organizācijas, kas publicē atvērtos datus veic šādas darbības salīdzinoši nesen, turklāt publicētie dati pieejami dažādos formātos, kas ir mašīnlasāmi taču to apstrādei pavadītais laiks krasi atšķiras atkarībā no formāta. Pirms datu ielasīšanas no publicētāja tīmekļa vietnes kurā atrodas sarežģīti mašīnlasāmi dati, jāizstrādā programmatūra ar kuras palīdzību būs iespējams ielasīt datus vietnei specifiskā veidā, apstrādāt tos un saglabāt datu krātuvē.

Latvijai, kā ES dalībvalstij ir saistošas visas ES direktīvas un reglamenti, kas stimulē atvērto datu kustību. Kā min vides aizsardzības un reģionālās attīstības ministrijas elektroniskās pārvaldes departamenta vecākais konsultants Toms Ceļmillers, šobrīd Latvijas atvērto datu sektoru regulē Informācijas atklātības likums[11]. Šī likuma saturam par pamatu ir ņemta Eiropas parlamenta un padomes direktīva 2003/98/EK[7] par valsts sektora informācijas atkalizmantošanu. Lai arī Informācijas atklātības likuma tekstā nekur neparādās vārdu salikums “atvērtie dati”, pēc būtības tas ir “paslēpts” aiz termina “atkalizmantošana”, norāda T. Ceļmillers.

Veicot padziļinātu izpēti var secināt, ka atvērtie dati Latvijā pastāv, taču nereti tie ir orientēti uz vietnes apmeklētājiem, kas visticamāk datus izmantos savas intereses apmierināšanai vai kādu citu statistisku pārskatu veidošanai masu medijiem. Lai arī iespēja saņemt jebkāda veida atvērtos datus jau ir atzīstams solis tajā pašā laika krietni vien lielāku labumu varētu sniegt mašīnlasāmi dati, pie tam ar patstāvīgu saiti. Līdz ar to var secināt, ka satura izvietotāji neapzinās patiesās atvērto datu iespējas un kapacitāti, kas paveras padarot datus iespējami atvērtākus.

### 2.1 Latvijas atvērto datu entuziasti

Latvijā sastopama arī “OpenData” entuziastu grupa[10], kura paši sava prieka vadīti apkopo Latvijā sastopamos atvērtos datus, padara tos mašīnlasāmus un viegli filtrējamus. Tāpat arī jebkurš cits interesents, kuram interesē atvērto datu kustība var ņemt dalību šajā grupā reģistrējoties data.opendata.lv un papildināt datu krātuvi ar jauniem datiem, izveidojot jaunu projektus, kā arī piedalīties diskusijās.

## 2.2 Atvērto datu kvalitātes uzlabošana

Lai arī iepriekš tika definētas vairākas atvērto datu kvalitātes metrikas kuras ievērojot dati paši par sevi ir ērti izmantojami, taču tas negarantē pašu datu satura piesātinājumu ar vērtīgu informāciju. Datu kvalitāti ietekmē vēl daži faktori un, lai nodrošinātu augstu datu kvalitāti būtu jādefinē vēl papildus metrikas:

- pēc iespējas detalizētāki dati – pieņemot, ka datu kopa satur informāciju par ūdens piesārņojumu publiskās peldvietās, būtu svarīgi zināt arī kādi parametri raksturo piesārņojumu, tā vietā lai zinātu tikai to, ka ūdenī drīkst vai nedrīkst peldēt;
- norāde uz ieejas datiem jeb pirmavotu – iespēja aplūkot arī sākotnējos datus no kuriem būvēta atvērtā datu kopā, piemēram ;
- datus jāatjauno tik bieži, lai tie pārklātu pēc iespējas tuvāku posmu lasīšanas brīdim – datiem jābūt pietiekoši bieži atjauninātiem, atbilstoši nozares specifikai, piemēram par peldvietu ūdens piesārņojumu dati būtu jāpublicē vismaz reizi dienā vai pat biežāk tajā pašā laikā, ja dati tiktu publicēti tikai reizi trijos gados, tad tos varētu izmantot tikai pārskatu veidošanai, bet ne peldētāju informēšanai katru dienu;
- datu formāts, lauku nosaukumi, struktūra un vērtību apgabals ir stabilas un nemainīgas vērtības, vai ir ieviests izmaiņu paziņošanas mehānisms – katra no lauku vērtībām ir aprakstīta, tā, lai izstrādātājam būtu skaidrs, kādas vērtības pieļaujamas konkrētajā laukā, tāpat, ja paredzēts ieviest jaunus laukus vai mainīt esošos, būtu vēlams ieviest izmaiņu paziņošanas mehānismu, kas brīdinātu par gaidāmajām izmaiņām un norādītu uz resursu, kur var uzzināt ko vairāk;
- datu iegūšana saite ir nemainīga un netiek mainīts arī tās formāts – ir svarīgi saites segmentiem un parametriem, ja tādi eksistē saglabāties nemainīgiem nevis pēc satura, bet uzbūves. Piemēram, ja saite izskatās šādi: [dati.gov.lv/finanses/pvn-iemaksas?gads=2014&men=03](http://dati.gov.lv/finanses/pvn-iemaksas?gads=2014&men=03), tad svarīgi ir saglabāt tās uzbūvi tādā veidā nodrošinot, ka datus varēs ielasīt automātiski, atkārtoti arī pēc sākotnējās datu ņēmēja iekārtas konfigurācijas bez papildus izstrādātāja iejaukšanās. Tas nozīmē, ka mainot saites segmentu saturu piemēram uz [dati.gov.lv/finanses/pvn?uzd=iemaksas&gads=2014&men=03](http://dati.gov.lv/finanses/pvn?uzd=iemaksas&gads=2014&men=03), dati vairs nebūs nolasāmi ar iepriekšējo konfigurāciju;
- pieejamie organizācijas dati ir viegli atrodamī un nav nepieciešamas specifiskas zināšanas par avota uzbūvi, lai atrastu datus – ieejot tīmekļa vietnē īpašniekiem būtu

jāpievieno kāds atvērto datu marķējums, piemēram “Open Definition” piedāvātās birkas<sup>5</sup>.

Ievērojot visu iepriekš uzskaitīto var iegūt ļoti augstas kvalitātes datus, kas ir nemainīgi kvalitatīvi, kā saturā tā arī pēc būtības. Uzturot šādus datus ir iespējams veidot dažādus derīgus un visiem saistošus e-pakalpojumus vai servisu.

Apskatot pieejamās datu kopas var pamanīt kādu kopīgu negatīvu iezīmi. Ja vērtējam datus tikai pēc piecu zvaigžņu skalas, tad visi datu publicētāji neiztur jau pašu pirmo līmeni, kas pieprasa, lai dati būtu publicēti tīmeklī ar atvēru licenci, tādā veidā nodrošinot rīcības brīvību ar publicētajiem datiem. Neviens no datu avotiem apskatīt licenci nepiedāvā. Citiem vārdiem, licences vienkārši nav, tāpēc šī darba ietvaros tiek uzskatīts, ka datu avots iegūst pirmās zvaigznes līmeni arī tad, ja licence nav pieejama. Pretējā gadījumā neviens no avotiem neiztur pat pirmās zvaigznes prasības un visas apskatītās atvērto datu kopas ir sliktas vērtējot pēc piecu zvaigžņu skalas.

### **2.3 Latvijā sastopamās datu kopas**

Veicot izpēti var secināt, ka ir sastopami avoti, kuros tiek publicēti brīvi pieejami dati, tomēr nav pakļauti kādam vienojošiem standartiem vai labajiem piemēriem. Lai atrastu pieejamās datu kopas tika meklēta informācija gadījumu kārtā atverot valsts iestāžu mājas lapas un meklējot starp tajā esošo saturu. Tāpat dažas datu kopas tika atrastas pateicoties iepriekš pieminētajai entuziastu grupas tīmekļa vietnei, kurā ir atrodams saraksts ar Latvijā pieejamo datu kopām.

Izvērtējot Latvijā reģistrēto domēnu saturā sastopamos atvērto datus tie tiks iedalīti pēc šādiem faktoriem:

- cik viegli atrodami atvērtie dati vietnes saturā;
- par cik ilgu laika posmu pieejami;
- datu kvalitāte raugoties caur “piecu zvaigžņu” prizmu;

Turpmākajā šīs apakšnodaļas saturā tiks apskatītas Latvijā sastopamās datu kopas un analizēta to kvalitāte. Analīze tiks veikta pēc iepriekš minētajiem kritērijiem, kā arī autora personīgā viedokļa.

---

<sup>5</sup> <http://opendefinition.org/buttons/>

### 2.3.1 Ceļu satiksmes drošības direkcija

CSDD ir valsts iestāde, kas nodarbojas ar dažādiem pakalpojumiem, kas saistīti ar transporta līdzekļiem, reģistrācija, auto ikgadējās tehniskās apskates, autovadītāju kvalificēšana un apliecību izgatavošana.

Pēc sākotnējās apskates uzreiz neizdodas atrast īpaši sagatavotus datus, taču pēc sīkākas izpētes, apakš sadaļā “Papildinformācija” iespējams atrast sadaļu “Statistika”, kas sevī ietver dažāda veida informāciju, par laika periodu sākot pat no 2000. gada.

Sākumlapa > Papildinformācija > Statistika > Ceļu satiksmes negadījumi

## CEĻU SATIKSMES NEGADĪJUMU SADALĪJUMS PA PAGASTIEM

### Pievienotie faili

- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2013.gadā \\*.xlsx](#), 33.42 Kb
- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2012.gadā \\*.pdf](#), 437.83 Kb
- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2011.gadā \\*.pdf](#), 402.77 Kb
- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2010.gadā \\*.pdf](#), 385.62 Kb
- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2009. gadā \\*.pdf](#), 567.15 Kb
- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2008. gadā \\*.pdf](#), 497.67 Kb
- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2007. gadā \\*.pdf](#), 452.04 Kb
- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2006. gadā \\*.pdf](#), 487.93 Kb
- [Ceļu satiksmes negadījumu sadalījums pa pagastiem 2005. gadā \\*.pdf](#), 449.28 Kb

### 2.1. att. CSDD publiski pieejamie dati

Dati sagrupēti pa vairākām kategorijām: ceļu satiksmes negadījumi, vadītāju apliecības, transportlīdzekļi, kuģošanas līdzekļi, autovadītāju apmācība. Zem dažām no grupām dati sadalīt pat vēl sīkāk, piemēram sadaļā “ceļu satiksmes negadījumi” sadalīta apakš sadaļās:

- ceļu satiksmes negadījumu skaits – satur statistiku par ES valstu satiksmes negadījumiem, un datus par Latvijā notikušiem satiksmes negadījumiem, kas sadalīti pa dažādām kategorijām, ievainotie, bojā gājušie, negadījumu skaits pēc laika, negadījumu skaits pēc notikuma vietas u.c.;
- ceļu satiksmes negadījumu sadalījums pa pagastiem – satur datus par satiksmes negadījumiem, kas sadalīti pa kalendārajiem gadiem.

Datu kvalitāte vērtējot to pēc iepriekš minētās “piecu zvaigžņu” skalas vērtējama ar divām zvaigznēm. Dati ir publiski pieejam jebkuram un tie ir strukturēti tabulās. Aptuveni 90% no visa informācijas pieejama PDF dokumentos, kas no lietotāja viedokļa ir ļoti ērti. Var viegli iegūt bezmaksas programmatūru PDF dokumentu lasīšanai ar kuras palīdzību var apskatīt dokumenta saturu. Taču no datu iegūšanas un atkalizmantojamības viedokļa šāda formāta dati krietni vien sarežģī apstrādes procesu. Dati, kas tiek ievietoti par pēdējiem gadiem ir formāts ir

mainījies uz .xlsx, kas no apstrādes viedokļa ir krietni vien ērtāk. Lai iegūtu nākamo kvalitātes līmeni tie būtu jāpublicē kādā no atvērtajiem formātiem piemēram CSV. Pieņemot, ka CSDD darbinieks, kas sagatavoja datus lieto MS Excell programmu, jo tieši šī programma pēc noklusējuma saglabā datus .xlsx formātā, var ļoti vienkārši saglabāt datus arī CSV formātā, ko nodrošina MS Excell programma.

Lielākais trūkums CSDD piedāvātajā pakalpojumā ir sarežģītā datu iegūšana. Šeit atrodama ļoti kvalitatīva informācija, kuru par pamatu varētu izmantot dažādiem noderīgiem komerciāliem risinājumiem. Piemēram sasaistot satiksmes negadījumu skaitu ar meteoroloģiskiem novērojumiem varētu iegūt informāciju par to, cik droši ir pārvietoties lietus laikā salīdzinot ar saulainiem vai miglainiem laikapstākļiem.

CSDD mājas lapā vienkārši un bez sarežģījumiem atrodami daudz un dažādi dati par iestādes sniegtajiem pakalpojumiem un ne tikai. Dati ir pietiekoši detalizēti un saprotami taču tie varētu būt vēl detalizētāki. Piemēram datus par iegūtajām vadītāja kvalifikācijām varētu norādīt arī personas vecumu un pilēta, kurā iegūta kvalifikācija. Tāpat dati varētu būt saistīti arī savā starpā, piemēram iegūto kvalifikāciju dati varētu būt saistīti satiksmes negadījumiem pa kategorijām. Šī valsts iestāde viennozīmīgi ir paraugs visu pārējo priekšā. Protams pie datu uzlabošanas vēl ir jāstrādā, bet kopumā piedāvāto datu plašais klāsts ir ļoti vērtīgs un interesants, turklāt dati pieejami arī par vairākiem iepriekšējiem gadiem, tāpat atsevišķās kategorijās dati pieejami arī par iepriekšējo noslēgto kalendāro mēnesi.

2.1. tabula

#### CSDD datu kvalitātes novērtējums

Kritērijs	Novērtējums
Datu pieejamība	Dati atrodami vidēji ātri taču strukturēti ļoti labi
Dati par iepriekšējiem periodiem	Atsevišķās kategorijās pieejami par periodu sākot no 2000. gada
“Piecu zvaigžņu” novērtējums	2

### 2.3.2 Uzņēmumu reģistrs

Uzņēmumu reģistrs ir iestāde, kas reģistrē uzņēmumus, komersantus, to filiāles un pārstāvniecības, masu informācijas līdzekļus, biedrības, un citus nodibinājumus, kā arī visas izmaiņas to darbības pamatdokumentos un veic citas likumdošanā noteiktās darbības.

Ieejot UR mājas lapā pietiekoši ātri var atrast datus, kas vismaz daļēji atbilst atvērto datu definīcijai, jo datu izmantošanas licence nekur nav atrodama. UR vietnes izvēlnes sadaļā “statistika” atrodas dažāda statistiska informācija par sniegtajiem pakalpojumiem un veiktajiem ierakstiem reģistros. Pati senākā informācija pieejama par veiktajiem ierakstiem sākot no

01.01.2009. Kā apgalvo vietnes uzturētāji informācija tiek atjaunota katru nedēļu un darba rakstīšanas brīdī jaunākie dati ir pieejami par iepriekšējo pilno darba dienu, bet nedēļas skatā pieejami dati par iepriekšējo pilno nedēļu. UR vietnē uzskatāmi redzams, ka dati tiek aktualizēti katru darba dienu, kas ir atzīstams fakts.

Dati pieejami lapas saturā tabulu formā (skatīt 2.2. attēlu), kas ir no mašīnlasāmības viedokļa ir pietiekoši ērti, nav mazsvarīgs arī fakts, ka datu apskates posmu var pieprasīt arī ar saitē norādītu GET parametru “start” kuram, kā vērtībā jāpadod meklējamās nedēļas pirmdienas datums formātā GGGG-MM-DD. Lai papildus atvieglotu datu apstrādi vietnes uzturētāji ir parūpējušies arī par datu iegūšanu arī CSV formātā. Zem dažām vietnē esošajām tabulām atrodama saite uz kuras nospiežot var lejupielādēt CSV datni kurā atrodas tabulas saturs.

Komerccdarbības forma	Reģistrēti	Aktīvi	Izslēgti
Akciju sabiedrība	452	391	61
Ārvalsts komersanta filiāle	661	468	193
Eiropas komercsabiedrība	3	3	0
Individuālais komersants	17397	12025	5372
Komandītsabiedrība	142	112	30
Latvijas komersanta filiāle	545	374	171
Pilsabiedrība	590	476	114
Sabiedrība ar ierobežotu atbildību	131122	121697	9425
<b>Kopā</b>	<b>150912</b>	<b>135546</b>	<b>15366</b>

Lejuplādēt statistiku [CSV](#) fomātā.

### 2.2.att. UR publiski pieejamie dati

Sākotnējā apskate liecina par detalizētiem un bieži aktualizētiem datiem. Tajā pašā laikā tīmeklī pieejama informācija, ka UR ir viena no pirmajām valsts iestādēm, kas atver savus datus piedāvājot tos lejupielādēt jebkuram bez maksas atvērto datu formā[15]. Meklējot UR tīmekļa vietnē un Google pēc dažādiem atslēgvārdiem neizdodas atrast atsevišķu vietu, kur varētu lejupielādēt visus UR datus vienuviet. Pēc ilgākiem meklējumiem tomēr izdodas atrast apakšdomēnā [dati.ur.gov.lv](http://dati.ur.gov.lv) izvietotu vietni, kurā pieejamas instrukcijas par sniegtajiem atvērtajiem datiem, kā tos importēt MS Excell programmā un PostgreSQL datu bāzē, kā arī paši atvērtie dati.

Datu kvalitāte no “piecu zvaigžņu” viedokļa vērtējama ar 4 zvaigznēm. Neskatoties uz to, ka arī šī vietne nepiedāvā apskatīt datu izmantošanas licenci, dati ir teicamā stāvoklī. Piemēram dati pavisam noteikti ir pieejami tīmeklī, dati ir labi strukturēti un pārskatāmi, dati pieejami kādā no atvērtajiem formātiem, šajā gadījumā CSV. Dati tiek publicēti unikālā adresē<sup>6</sup>, kur jebkurš var datu kopas pievienot pie savu grāmatzīmju sarakstam, un vēlāk atrast, vai kā

<sup>6</sup> <http://dati.ur.gov.lv/register/>

citādi identificēt konkrētu datu kopu. Tāpat datos pieejamas saites, uz citiem resursiem, šajā gadījumā gan uz paša UR tīmekļa lapu, kur var apskatīt informāciju par reģistrēto komercsabiedrību, ko nevar uzskatīt par pietkā līmeņa kvalitātes zīmi.

Autora personīgais viedoklis liecina, ak šīs valsts institūcijas atvērto datu serviss viennozīmīgi ir priekšzīmīgs, ne tikai pēc “piecu zvaigžņu” skalas, bet arī pēc datu detalizētības pakāpes, atjaunošanas biežuma un citiem faktoriem, kas izvirza UR priekšplānā salīdzinot ar citiem datu avotiem.

2.2. tabula

#### UR datu kvalitātes novērtējums

Kritērijs	Novērtējums
Datu pieejamība	Daļa no datiem atrodami ērti, taču UR open data serviss, kas atrodas citā domēnā atrodams ar grūtībām.
Dati par iepriekšējiem periodiem	Pieejami sākot no 2009. gada
“Piecu zvaigžņu” novērtējums	4

### 2.3.3 Rīgas domes izglītības, kultūras un sporta departaments

RDIKSD ir iestāde, kuras pakļautībā ir vairāk, kā 300 dažādas pašvaldības dibinātās iestādes, tajā skaitā pirmsskolas, pamatizglītības, vidējās izglītības un citas. RDIKSD sniegto pakalpojumu klāstā ietilpst informācijas sniegšana par izglītības pieejamību, interešu un pieaugušo neformālās izglītības programmu licencēšana, mūzikā izglītības procesa koordinēšana un citi pakalpojumi.

Ienākot RDIKSD tīmekļa vietnē<sup>7</sup> atvērto datu sadaļu nevar pamanīt uzreiz, taču pēc OpenData Latvijas entuziastu grupas vietnē esošās informācijas noprotams, ka tāda šeit eksistē. Atvērto datu kopu tomēr izdodas atrast (skatīt 2.3. attēlu) sadaļas “Iestāžu katalogs” apakš sadaļā<sup>8</sup> “Open Data”, kurā atrodami daži departamenta rīcībā esoši dati. Ņemot vērā, ka vietne sadalīta vairākās lielās pirmā līmeņa sadaļās – “Izglītība”, “Kultūra”, “Sports” un “Iestāžu katalogs”, bet dati pieejami tikai par vienu no tām, jāatzīst, ka pieejama tikai neliela daļa datu. Pašlaik departaments piedāvā datus par trijām kategorijām “Izglītības iestāžu saraksts”,

<sup>7</sup> <http://dati.e-skola.lv/>

<sup>8</sup> <http://dati.e-skola.lv/OpenData>

“Skolēnu skaita dinamika” un “Ģeogrāfiskā informācija”. Šajās sadaļās pieejama informācija par laika periodu sākot no 2003. gada.

### Atvērta pieeja datiem

Piedāvājam lestāžu kataloga datus mašīnlasāmā formā.

Datu kopa	ASCII	UTF-8
Izglītības iestāžu saraksts	lestades.csv	CSV
Skolēnu skaita dinamika Vispārizglītojošajās skolās no 2003. līdz 2011. gadam.	SkolenuSkaitaDinamika.csv	CSV
Ģeogrāfiskā informācija Uzmanību! Vienai iestādei mēdz būt vairāki punkti kartē.	Karte.csv	CSV

#### 2.3. att. RDIKSD atvērto datu avots

Datos atrodas informācija par skolām, to adresēm, kontaktinformāciju, iekšējiem identifikatoriem, un citām kategorijā. Dati viennozīmīgi varētu būt detalizētāki, taču arī esošā informācija ir pietiekoši kvalitatīva un to var apstrādāt un izmantot piemēram informējošu resursu izstrāde. Kā piemēru var minēt lietotni, kas varētu piedāvāt ērtāko skolu atvasei pēc vecāku māju un darba vietas novietojuma, tā lai veicot šo maršrutu būtu iespējams bērnus uz skolu pavadīt vai aizvest ar auto.

Pēc “piecu zvaigžņu” skalas datu serviss novērtējams ar 3 zvaigznēm, protams, ja neņem vērā to, ka arī šeit nav pieejama datu izmantošanas licence. Dati ir pieejami internetā, pieejami, kā strukturēti dati atvārtā formātā, šajā gadījumā CSV.

2.3. tabula

#### RDIKSD datu kvalitātes novērtējums

Kritērijs	Novērtējums
Datu pieejamība	Dati pieejami vidēji ērti. Nedaudz maldina datu pieejamība tikai vienā sadaļā
Dati par iepriekšējiem periodiem	Pieejami sākot no 2003. gada
“Piecu zvaigžņu” novērtējums	3

### 2.3.4 Rīgas domes mājokļu un vides departaments

RDMVD ir vadošā iestāde, energoapgādes, ūdensapgādes, un vides aizsardzības jautājumos. Iestāde sniegto pakalpojumu klāstā ir dzīvojamo platību reģistrācija un ar to saistīti pakalpojumi, ēku apsaimniekošana, arhīva dokumentu kopiju un izrakstu sniegšana un vairāki citi pakalpojumi.

Apskatot RDMVD tīmekļa vietni<sup>9</sup> jau pašā sākumlapā var redzēt daudz un dažādas saites uz informatīviem resursiem, kas potenciāli varētu būt atvērtie dati, taču šī darba ietvaros tiks apskatīti gaisa piesārņojuma mērījumu dati, kas ir pamanāmi vietnes kreisajā pusē, kā reklāmkarogs. Apakš sadaļā<sup>10</sup>, kas veltīta tieši kvalitātes mērījumiem Rīgas pilsētā pieejama informācija no trijām mērīšanas ierīcēm par laika periodu, kas nav lielāks par 10 dienām no aplūkošanas brīža un satur mērījumus ar intervālu viena stunda. Dati pieejami septiņās automātiski ģenerētās grafiskās diagrammās (skatīt 2.4 attēlu), GIF formātā, kas tiek pārzīmētas reizi 2 stundās.

01.Slāpekļa dioksīda NO2 koncentrācijas 3 stacijās	<a href="#">04 days 01.gif</a>
02.Ozona O3 koncentrācijas trīs stacijās	<a href="#">04 days 02.gif</a>
03.Sēra dioksīda SO2 koncentrācijas 3 stacijās	<a href="#">04 days 03.gif</a>
04.Benzola C6H6 koncentrācijas 3 stacijās	<a href="#">04 days 04.gif</a>
05.Oglekļa oksīda CO koncentrācijas Kr.Valdemāra ielā	<a href="#">04 days 05.gif</a>
06.Slāpekļa oksīda NO koncentrācijas Kr.Valdemāra ielā	<a href="#">04 days 06.gif</a>
07.Putekļu PM10 koncentrācijas Kr.Valdemāra ielā	<a href="#">04 days 07.gif</a>
08.Tabula - 7 piesārņojošo vielu koncentrācijas 10 iepriekšējās dienās	<a href="#">pollut.txt</a>

#### 2.4. att. RDMVD gaisa kvalitātes mērījumu dati

Dati par iepriekšējām 10 dienām visās 3 mērīšanas stacijās pieejami arī teksta datnē TXT formātā. Datne satur improvizētu tabulu, kas veidota no atstarpēm un citiem simboliem (skatīt 2.5. attēlu). Teksta datu tabula satur 6 mērījumu grupas, no trim mērīšanas ierīcēm ar vidējo piesārņojuma daudzumu katru stundu par iepriekšējām 10 dienām.

Slāpekļa dioksīds (NO2), µg/m³. R=200 µg/m³

Datums	Laiks	Brīvības 73	Kr.Valdemāra 18	Tvaika 44	
2014.05.23	00:00	68.82	41.19	34.61	(NO2)
2014.05.23	01:00	51.81	35.82	23.90	(NO2)
2014.05.23	02:00	27.54	31.38	10.98	(NO2)
2014.05.23	03:00	18.65	26.85	6.62	(NO2)

#### 2.5. att. Gaisa kvalitātes mērījumu tekstuāla tabula

<sup>9</sup> <http://mvd.riga.lv/>

<sup>10</sup> [http://mvd.riga.lv/lv/gaiss\\_sobrid/](http://mvd.riga.lv/lv/gaiss_sobrid/)

Dati ir pietiekoši detalizēti un ar tiem varētu izveidot informatīvu lietotni, kas ziņo par gaisa stāvokli pilsētas ielās vai veic analīzi pa laika periodiem. Lai arī dati pieejami dažādos attēlu formātos, kas no mašīnlasāmības viedokļa ir ļoti slikti, tomēr pieejama arī teksta datne, kas salīdzinot ar attēlu ir vienkāršāk apstrādājama. Lai būtu iespējams piedāvātos teksta datus pienācīgi izmantot būtu nepieciešams tos parsēt, izņemot no kopēja teksta atstarpes un citus nevajadzīgus simbolus tādā veidā iegūstot tīrus datus, kas ir gatavi pielietošanai. Lai arī TXT formāts ir atvērts, tomēr tā saturā var būt arī ļoti sarežģīti mašīnlasami dati. Šajā gadījumā tomēr dati ir strukturēti, taču tā tabula atdalīta ar atstarpēm un citiem simboliem, kas apgrūtina ielasīšanu un palielina kļūdu skaitu. Līdz ar to šīs datu avots no piecu zvaigžņu viedokļa iegūst 1,5 zvaigznes. Teorētiski, dati ir strukturēti, un tomēr, lai tos ielasītu būs jāpieliek vairāk pūles nekā tas būtu prasījis ar piemēram XLS formāta tabulu, tāpēc par centību tiek piešķirta 0,5 zvaigzne.

Kā jau minēts iepriekš RDMVD vietnē sastopami dati par vairākām sfērām taču ieskatoties dziļāk katrā no sadaļām var apskatīt dažādus datus, kas sevī nes lielu potenciālu. Šie dati būtībā ir neizmantojami, ja vien kāds tos manuāli nepārtaisa mašīnlasāmās tabulās. Piemēram vietnē pieejamas dažādas kartes, kurās attēlota informācija par trokšņu līmeni pilsētā<sup>11</sup>, kas publicētas, kā JPG attēli. Tāpat pieejami dažādi dati, kas atrod starp tekstu. Arī šāda veida datus ir jāapstrādā manuāli un no satura “jāizvelk” derīgie dati. Pēc apskatītās tīmekļa vietnes var secināt, ka konkrētajai iestādei ir liels datu potenciāls un tā varētu kļūt par nozīmīgu atvērto datu avotu.

2.4. tabula

#### RDMVD datu kvalitātes novērtējums

Kritērijs	Novērtējums
Datu pieejamība	Dati pieejami ērti
Dati par iepriekšējiem periodiem	Par 10 iepriekšējām dienām
“Piecu zvaigžņu” novērtējums	1,5

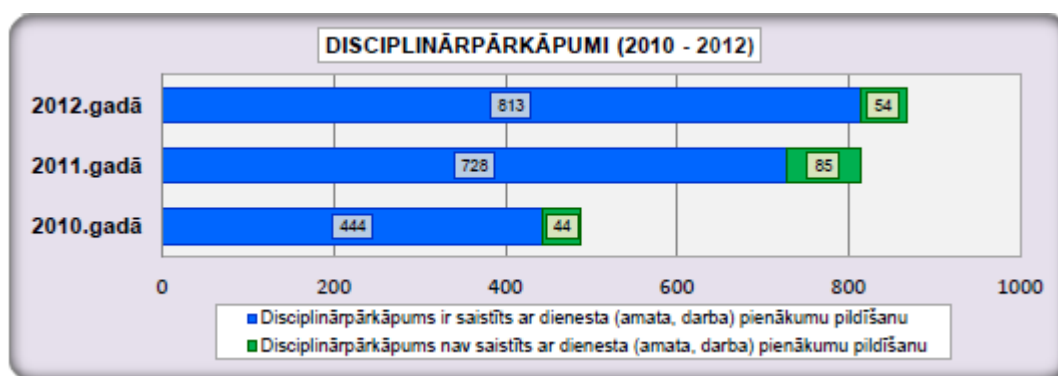
### 2.3.5 Latvijas Republikas Iekšlietu ministrija

Iekšlietu ministrija ir valsts institūcija, kas ir atbildīga par noziedzības apkarošanu sabiedriskās kārtības un drošības uzturēšanu, kā arī valsts robežu drošības uzturēšanu un citām funkcijām.

<sup>11</sup> [http://mvd.riga.lv/lv/vide/troksnu\\_kartes/](http://mvd.riga.lv/lv/vide/troksnu_kartes/)

Ienākot IEM tīmekļa vietnē pietiekoši ātri var atrast sadaļu “Statistika un publikācijas”, kur parasti atrodas atvērtie dati, tajā atrodas apakš sadaļa<sup>12</sup> “statistika”, kurā savukārt atrodas IM un tās padotībā esošu iestāžu disciplināras prakses pārskati PDF formātā. Pārskati sarindoti pēc gadiem un pieejami par laika periodu sākot no 2008. gada. Katrs no pārskatiem sastādīts, kā informējošs pārskats, kur aprakstošs teksts mijās ar grafiskām diagrammām (skatīt 2.6. attēlu) par tādām tēmām, kā:

- apbalvojumu piešķiršana;
- piešķirto apbalvojumu veidi;
- piešķirto apbalvojumu skaita dinamika;
- disciplinārsodu piemērošana;
- lēmumi par disciplinārsodu piemērošanas apstrīdēšanu;
- disciplinārpārkāpumi;
- un citas.



2.6. att. IM pārskata grafisks piemērs

Pārskatu saturs ir bagāts ar dažādiem datiem, taču no atkalizmantojamības viedokļa no tiem nav lielas jēgas, jo, lai tos varētu lietot dati ir manuāli jāizņem no pārskata un jāpadara mašīnlasāmi.

Bez šiem datiem tīmeklī pieejam informācija arī par IM informācijas centra piedāvāto informāciju, par dažādiem noziedzīgiem pārkāpumiem un citiem noziedzīga rakstura datiem, taču šādi resursu nav iespējams atrast no IM tīmekļa vietnes.

Ar meklētājdzineja palīdzību tomēr iespējams atrast LRIMIC tīmekļa vietni, kur apakš sadaļā<sup>13</sup> “Statistiskie pārskati” tiek piedāvāti dati par vairākām kategorijām, “Situācija Baltijas valstīs”, “Kriminālā statistika”, “Meklējamais transports”, “Nederīgo dokumentu reģistra statistika”, “Administratīvie pārkāpumi”. Visās iepriekš minētajās sadaļās ievietoti PDF un

<sup>12</sup> [http://www.iem.gov.lv/lat/statistika\\_un\\_publicācijas/statistika/](http://www.iem.gov.lv/lat/statistika_un_publicācijas/statistika/)

<sup>13</sup> <http://www.ic.iem.gov.lv/lv/node/69>

XLS formātu dokumenti (skatīt 2.7. attēlu), taču pēdējo 4 - 5 gadu pārskati ievietoti tikai XLS formātā. Jaunākie dati pieejami par aptuveni divus mēnešus veciem notikumiem un atsevišķās kategorijās pieejama sākot no 2003. gada.



### 2.7. att. LRIMIC publiski pieejamie dati

Lai arī dati pieejamai par dažādām kategorijām to saturs tomēr ir salīdzinoši neliels. Piemēram pārskatā par nolauptajiem un atrastajiem auto pieejami dati par automašīnu markām un gadiem, kur tiek uzskaitīts cik konkrēto auto vienību nozagtas un cik atgūtas. Šeit būtu ļoti svarīgas arī nozieguma detaļas, piemēram nozieguma vieta, laiks un tamlīdzīgs detaļas, kas varētu ļaut radīt labas un noderīgas lietotnes, kas brīdinātu par potenciāli bīstamām vietām, kur novietot savu auto. Līdzīgi ir arī citās pārskatu kategorijās, kur dati eksistē, taču tie ir ļoti virspusīgi, kas it kā iezīmē kopainas aprises taču nianšes, kas varētu krietni palielināt datu vērtību tomēr noklusē.

Kopumā IM rīcībā ir liels datu apjoms, par ko liecina arī tās pakļautībā izveidotais LRIMIC, kura tiešajos pienākumos ietilpst datu apstrāde un dažādu informatīvu ziņojumi sniegšana

Vērtējot avotu pēc “piecu zvaigžņu” skalas LRIMIC vērtējams ar 2 zvaigznēm. Dati pieejami publiski, lielākā daļa no tiem ir strukturēti XLS formāta dati, taču netiek publicēti kādā no atvērtajiem formātiem.

2.5. tabula

### LRIMIC datu kvalitātes novērtējums

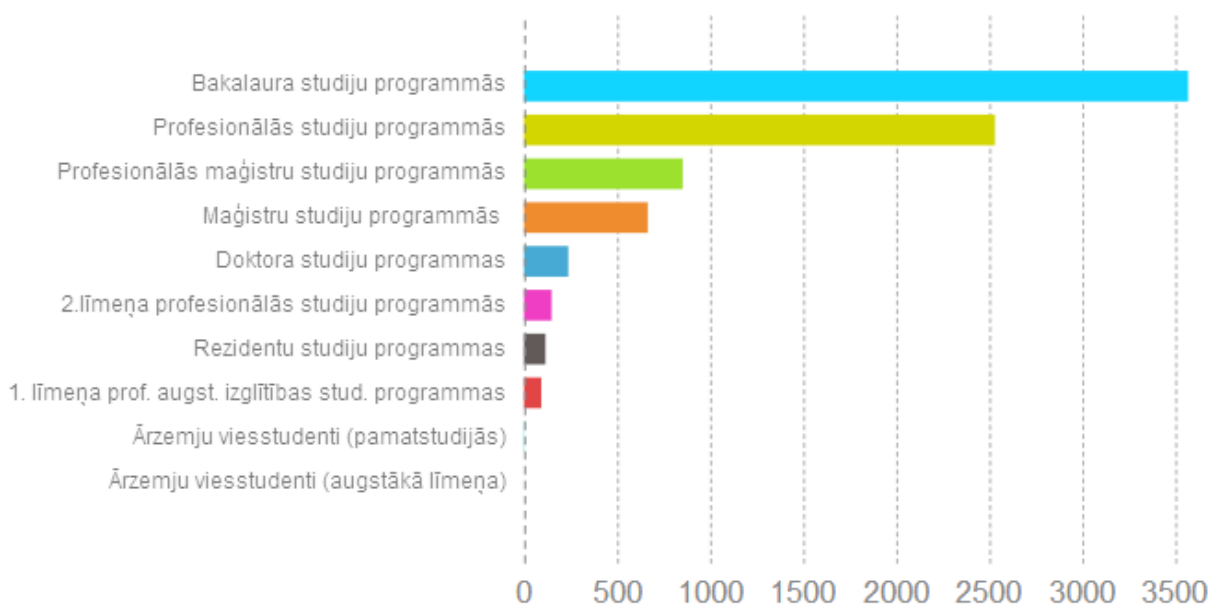
Kritērijs	Novērtējums
Datu pieejamība	Dati pieejami ļoti sarežģīti
Dati par iepriekšējiem periodiem	Sākot no 2003. gada
“Piecu zvaigžņu” novērtējums	2

### 2.3.6 Latvijas universitāte

LU ir Latvijas augstākās izglītības iestāde, kur var iegūt izglītību humanitārajās, sociālajās, dabas un veselības zinātnēs. LU sagatavo mācībspēkus citām mācību iestādēm, kā arī valsts sektora speciālistus.

LU tīmekļa lapās<sup>14</sup> izvēlne sadalīta vairākās sadaļās un apakš sadaļās, tajā skaitā arī līmeņos. Vietne ir bagātā ar dažādu informāciju, taču statistiskas vai atvērto datu sadaļu uzreiz nevar atrast. Ņemot vērā, ka LU ir viena no Latvijas vadošajām augstskolām tai tomēr būtu jāpublisko dati par uzņemtajiem studentiem, studiju programmu aizpildījumu un citiem faktoriem. Sazinoties ar LU IT servisa tehnisko darbinieku, e-pasta vēstulē tas atklāj, ka tomēr pieejamas vairākas vietas, kur publicēti dažāda veida dati par uzņemšanas<sup>15</sup> procesu, īstenotajiem ES struktūrfondu projektiem<sup>16</sup> un citiem iekšējiem rādītājiem<sup>17</sup>. Šī darba kontekstā tiks apskatīta pēdējā pieminētā vieta kurā pieejamai vairāki dažāda rakstura pārskati.

Konkrētā statistikas datu sadaļa<sup>17</sup> atrodas LU galvenās vietnes apakš sadaļā un teorētiski to ir iespējams atrast, taču, autoram to paša spēkiem neizdevās izdarīt. Sadaļā dati izvietoti pa vairākām kategorijām, piemēram “Studentu skaita sadalījums pa fakultātēm”, “Absolventi”, “Ziedojumi”, “LU darbinieku vidējais atalgojums” un citas. Dažās no kategorijām izvietota informācija ar informatīvu grafiku palīdzību (skatīt 2.8. attēlu), taču dažos ar lapas saturā esošu tabulu palīdzību.



2.8. att. LU publiski pieejamie dati

<sup>14</sup> <http://www.lu.lv/>

<sup>15</sup> <http://www.lu.lv/gribustudet/pamatstudijas/buj/konkurss/>

<sup>16</sup> <http://www.lu.lv/par/projekti/es/info/statistika/>

<sup>17</sup> <http://www.lu.lv/par/dokumenti/statistika/>

Vietnē atrodamie dati pieejami tikai saturā un atsevišķu datņu lejupielāde nav nodrošināta tādā veidā ierobežojot iespēju mašīnlasāmi apstrādāt datus, ja tie būtu pieejami kādā no atvērtajiem datu formātiem. Informācija lielākoties pieejam par pēdējiem 2 – 4 gadiem taču atsevišķās kategorijās piemēram “LU darbinieku vidējais atalgojums”<sup>18</sup> dati pieejami arī par laika posmu sākot no 2006. gada. Tabulās atrodamā informācija ir ļoti virspusēja un sniedz nelielu ieskatu par situāciju katrā no kategorijām. Sniegtie dati viennozīmīgi varētu būt detalizētāki, kas ļautu tos izmantot pilnvērtīgāk.

Vērtējot datu avotu pēc “piecu zvaigžņu” skalas, LU datu kvalitāte vērtējama ar 1 zvaigzni. Dati pieejami tīmeklī taču nevienu no tālākajām prasībām tie neapmierina. Lai arī dati ir strukturēti, tie atrodas lapas saturā izvietotās tabulās, kuru apstrāde prasa vairāk laiku, kā tas būtu nepieciešams ar piemēram XLS datnes saturu.

2.6. tabula

#### LU datu kvalitātes novērtējums

Kritērijs	Novērtējums
Datu pieejamība	Dati pieejami vidēji sarežģīti
Dati par iepriekšējiem periodiem	Sākot no 2006. gada
“Piecu zvaigžņu” novērtējums	1

## 2.4 Potenciālie datu avoti, kas datus nepiedāvā

Lai arī ir apskatītas vairākas iestādes, kas piedāvā dažādus datus par tās sniegtajiem pakalpojumiem un ne tikai, tomēr lielākā daļa nepiedāvā iespēju iegūt datus, kas pēc rakstura līdzinātos atvērtajiem datiem. Privātā sektora uzņēmumi, kas nodarbojās ar dažādu pakalpojumu sniegšanu un ir potenciāls atvērto datu avots norāda, ka var sniegt statistisku informāciju tikai individuāli apstrādājot katru pieprasījumu manuāli. Tā piemēram Latvijā lielākais sociālais tīkls SIA “Draugiem”, kas ik dienu apkalpo vairākus desmitus tūkstošus cilvēku, apstiprina, ka tiešsaistes režīmā nekādu statistiska rakstura datus nepiedāvā. Tomēr portāla sākumlapā var novērot, ka kājenē var jebkurš brīvi apskatīt šobrīd portālā esošos tiešsaistes lietotāju skaitu (skatīt. 2.3. attēlu), kas pēc būtības varētu būt atvērtie dati.

Noteikumi Palīdzība Jaunumi  
Šobrīd portālu skatās: 47 459

### 2.3 att. SIA “Draugiem” portāla apmeklētāju skaits

Latvijā un ES atvērto datu kustība ir salīdzinoši jauna, kā rezultātā sabiedrībai trūkst informācijas par tās potenciālajiem labumiem, ko tā varētu sniegt. Ir vēl kāds faktors, kas varētu

<sup>18</sup> <http://www.lu.lv/par/dokumenti/statistika/videjaisatalgojums/>

atturēt potenciālos datu avotus no šāda servisa ieviešanas. Lai izstrādātu sistēmu vai esošas sistēmas moduli, kas veic visas nepieciešamās darbības ar datiem un, lai tie būtu gatavi lejupielādei nepieciešami resursi. Darbaspēka resurss prasa finansiālus līdzekļus taču, neredzot tālāku finansiālu ieguvumu ieguldīt papildus līdzekļus uzņēmumi to darīt varētu nevēlēties. Jāņem vērā arī tas, ka uzņēmumam redzot potenciālu iespēju kādus savus datus pārdot, šāda iespēja varētu likties vilinošāka nekā datu “dāvināšana” bez maksas kāda cēlāka mērķa vārdā.

Eksistē arī citi iemesli, kas varētu atturēt no datu publicēšanas. Iespējams organizācija nenojauš savu datu vērtību, uzskatot tos par nevienam nevajadzīgiem. Dati gandrīz vienmēr kādam būs svarīgi un interesējoši, taču cits aspekts ir pareizā potenciālo interesentu piesaiste. Datu vērtību var pārbaudīt ļoti vienkāršā veidā – piedāvājot to publiski par brīvu vai par nelielu samaksu. Pēc informācijas iegūšanas par datu vērtību atliek plānot ko tālāk darīt ar organizācijas īpašumā esošajiem datiem. Piemēram, ja uzņēmums nodarbojās ar asfalta ieklāšanu, tad tās īpašumā ir informācija par to, kur uzliets jauns asfalts, bet kur lāpītas bedres. Šādu informāciju var piedāvāt publiski vai piedāvāt iegādāties kādai kompānijai, kas nodarbojās ar maršrutu plānošanu. Protams ieguvumu saskatīšana ir katras iestādes individuālā ziņā un noskaņojoties pret pret atvērto datu kustību iespējams nevar mainīt nekādi argumenti. Taču, lai kļiedētu šaubas jālūkojas pasaules piemēru virzienā.

## **2.5 Atvērto datu, kopas ieviešanas priekšrocības**

Kā jau minēts iepriekš atvērto datu esamībai ir vairāki pozitīvi faktori. Kā viens no svarīgākais būtu jāmin atvērto datu labvēlīgā ietekme, uz jaunu biznesa veidu, pakalpojumu un servisu rašanos, kas savukārt nesīs papildus ieņēmumus valsts kasē. ES komisija ir plāno, pēc atvērto datu ieviešanas valsts iestādēs visā ES, kopējā ekonomika tiks palielināta ar papildus 40 miljardiem eiro, katru gadu[23]. Tiek uzskatīts, ka katra no valsts iestādēm patiesībā sēž uz zelta monētu kaudzes un to pašlaik neapzinās. Stiprinot atvērto datu kustību ES ir vēlas kļūt par pasaules līderi PSI atkalizmantošanā.

Līdzās jau uzskaitītajiem ieguvumiem atvērtie dati ir potenciāli ļoti spēcīgs instruments korupcijas novēršanai un valsts iekšējo procesu caurspīdības uzlabošanai. Padarot datus atvērtus par valsts finanšu izlietošanu, publiskajiem iepirkumiem un citām PSI sfērām iespējams iesaistīt sabiedrību novērošanas un līdzdalības procesā, kas sekmē krāpniecisku darbību izskaušanu. Spilgts piemērs šajā ziņā ir Brazīlijas valdības veiktie uzlabojumi informācijas publiskošanā par gadījuma kārtā veiktiem valsts amatpersonu auditiem, kas samazināja amatpersonu korupciju kopumā visā valstī[24].

Kā pēdējo bet ne mazsvarīgāko ieguvumu var minēt zināšanu bāzi, kuru var iegūt uzkrāt un izmantot. Datu patstāvīga uzkrāšana, analizēšana un secinājumu izdarīšana sekmē turpmāku produktu, pakalpojumu un servisu pilnveidošanu, kas balstīta uz zināšanām.

Pēdējo 5 gadu laikā tehnoloģijas ir veikušas milzu lēcieni tāpēc, kas notiks nākotnē var tikai minēt. Līdzīgi arī ar atvērto datu ieguvumiem – ir neiespējami prognozēt, kādi ieguvumi atkāšies nākotnē, jo nav zināms, kādas tehnoloģijas būs pieejamas un kādi paņēmieni atklāti. Tikmēr valstu valdības sliecas domāt, ka atvērto datu kustībā ir milzīgs potenciāls, kas vēl pašam nav aptverts.

## 2.6 Atvērto datu, kopu trūkumi

Atvērto datu publicētāj valstīm ir vairākas priekšrocības pār tiem, kas to nedara taču pastāv arī zināmi riski, kurus ir jāapzinās un jācenšas tos ierobežot un minimizēt.

Kā viens no atvērto datu trūkumiem tiek minēta vēsturiskā pieredze ar datu apkopošanu. Iestāžu rīcībā bieži ir milzīgi datu apjomi, kuru apstrāde prasa milzīgas pūles, resursus un darbietilpību. Lai persona apstrādātu šādus datus, padarītu tos mašīnlasāmus un atvērtā formātā tas var aizņemt daudz laika pie tam ja dati ir jāpublicē katru nedēļu vai katru dienu, tas var izrādīties pārāk dārgi, lai to darītu manuāli. Lai šo problēmu risinātu par to jādomā brīdī, kad iestādei tiek izgatavot informācijas sistēma ar kuru plānots veikt ikdienas darbus. Mūsdienā datoru jauda un tādas tehnoloģijas, kā makoņskaitļošana<sup>19</sup> ļauj lielus datu apjomus apstrādāt un uzturēt daudz efektīvāk un ātrāk. Datu sagatavošanai un publicēšanai noteikti jānorisinās automātiski, bez cilvēka iejaukšanās, šajā procesā.

Pieņemot, ka kādas sfēras dati tiek atvērti un padarīti publiski pieejami, ļoti iespējams, ka šajos datos arī varētu parādīties informācija, kas ir aizskaroša, kādai personai vai personu grupai. Piemēram, ja tiek publicēti dati par tiem Rīgas daudzdzīvokļu namiem, kas ir parādā siltuma piegādātājam par iepriekšējām apkures sezonām, tad šāda veida informācija varētu būt aizskaroša vai nepatīkama tiem, kas dzīvo šajos namos, bet visus siltuma rēķinus ir nomaksājuši. Tāpat iespējams, ka dati, kuros pieejama dažāda ģeoloģiska informācija norāda par kādām krasām novirzēm no normas, kas varētu likt sabiedrībai pārlietu uztraukties vai pat radīt paniku. Piemēram informācija, kas varētu izraisīt nevēlamu satraukumu varētu būt par strauju ūdens līmeņa celšanos ūdens tilpnēs, kritisku gaisa piesārņojuma palielināšanos, strauju maksātnespējīgo uzņēmumu skaita pieaugumu vai kādiem citiem potenciāli bīstamiem mierīgu ikdienu ietekmējošiem rādītājiem.

---

<sup>19</sup> <http://www.pcmag.com/article2/0,2817,2372163,00.asp>

Autora personīgais uzskats ir tāds, ka iespējams valsts iestādes uzskata, ka to rīcībā esošie dati ir milzu vērtē un tos atdot kādam bez maksas būtu nepieļaujami, tāpēc dati tiek turēti iestāžu īpašumā un iespējams tā īsti nemaz netiek izmantoti.

### 3 TENDENCES

Latvijas atvērto datu kustība ir apņēmības pilna attīstīt un pierādīt, ka atvērtie dati ir demokrātiskas un ilgtspējīgas valsts sastāvdaļa. Pasaules pieredzes ietekmē arī vietējās valsts un pašvaldību iestādes apsver iespēju uzlabot piedāvāto datu klāstu un to kvalitāti. Par Latvijas atvērto datu kustības attīstību liecina arī darba izstrādes laikā valsts iestāžu izveidotie datu avoti, Tā piemēram dažas dienas atpakaļ darbu uzsāka UR atvērto datu avots, kas ir viens no pirmajiem valsts PSI datu avotiem, turklāt datus UR piedāvā kvalitatīvā līmenī vadoties pēc “piecu zvaigžņu” skalas. Privātā e-pasta sarakstē ar autoru UR darbiniece Sandra Notrums norāda, ka: ”Šie atvērtie dati, ko šobrīd piedāvājam, ir pirmais solis šajā virzienā. Laika gaitā mēs plānojam attīstīt šo jomu. Tālākais solis varētu būt Maksātnešpējas reģistrs, bet, lai to piedāvātu interesentiem ir jāveic sagatavošanās darbi. Prognozējams, ka tas varētu notikt ne ātrāk kā nākamajā gadā.”. Darbinieces teiktais apliecina, ka iestādes mērķis ir attīstīt un uzturēt atvērto datu avotu.

#### 3.1 Datu uzlabošana

Apskatītie Latvijas datu avoti atklāj vairākas nepilnības, kas nākotnē noteikti būtu jānovērš. Tā piemēram, lai arī iestāde uztur datus bieži tos ir ļoti sarežģīti atrast. Dažās vietās bija nepieciešams sazināties arī ar iestādes darbiniekiem, kas palīdzēja atrast pieejamos datus. Šo problēmu noteikti būtu jāatrisina attīstot atvērto datu servisu. Kā vienkāršāks risinājums, ko jau piedāvā “OpenDefinition” organizācija ir informatīvas birkas<sup>20</sup> jeb pogas pievienošana tīmekļa vietnes kājēnē, kur pēc klikšķa uz pogas, lietotājs nonāk konkrētās iestādes atvērto datu publicēšanas vietā. Šādā veidā iespējams panākt ieinteresēto personu pietiekoši ātru informēšanu par saturā pieejamajiem atvērtajiem datiem. Pogas dizains ir standartizēts un tāpēc visās vietās tam būtu jābūt vizuāli vienādam. Pieejami arī dažādi citi birku veidi (skatīt 3.1 attēlu), kas informē lietotāju par atvērto datu saturu, kas iespējams nav tieši dati, bet gan kāds cita veida noderīgs saturs.



3.1. att. Informatīvas atvērto datu birkas

<sup>20</sup> <http://opendefinition.org/buttons/>

Otrs paņēmiens, kas risinātu datu pieejamības problēmu varētu būt atsevišķa servisa izveide, kur visas valsts institūcijas varētu publicēt savus datus, līdzīgi, kā to jau dara ASV<sup>21</sup> un Apvienotās Karalistes<sup>22</sup> valdības. Šajā vietnē iespējams noteikt vienotus standartus, datu formātus un citas ietekmējošu faktorus, kas izstrādātājam var atvieglot datu iegūšanas procesu. Kopējā valsts datu krātuvē var izveidot, dažādas kategorijas, piemēram “vide”, “transports”, “velelība” un citas, kur datus iespējams ātri uziet.

Kā papildus noteicošs faktors datu pieejamībai būtu iespēja jau meklētāj dzinēja rezultātos pamanīt, ka saturs, kas atrasts atbilst atvērtai datu kopai. Šāda efektu varētu panākt vietnes saturā ievietojot HTML birku ar saturu aprakstošu informāciju (skatīt 3.2 attēlu), ko savukārt meklētājdzinēja rāpuļi varētu nolasīt un noglabāt pie vietnes aprakstošajiem indeksiem. Tālākā satura attēlošana meklētājdzinēju rezultātos protams ir individuālas dabas jautājums taču pareizāk būtu lietot jau iepriekš pieminētās “OpenDefinition” birkas, kas nodrošinātu vienotu izpratini par vietnē atrodamo. Šādi meklētājdzinēju rezultāti protams ir katra individuāla izvēle, tāpēc attēlot vai slēpt būtu katra individuālie meklēšanas uzstādījumi.

```
<meta name="open-data" content="open-data, open-knowledge">
```

### 3.2. att. Atvērto datu meta dati

---

<sup>21</sup> <http://www.data.gov/>

<sup>22</sup> <http://data.gov.uk/>

## SECINĀJUMI

Apskatot atvērto datu kustību var secināt, ka pasaulē ir vairāki pozitīvi piemēri, kas pozitīvi ietekmē to izaugsmi. Pasaules valstis, tajā skaitā arī Latvija iespaidojoties no veiksmīgās pieredzes, ko parādījis, kā privātais tā arī publiskais sektors, ievieš arī savā dienas kārtībā atvērto datu servisu izstrādi.

Atvērto datu kustībai viennozīmīgi ir potenciāls, pie kam līdz šim brīdim tas nav īsti novērtēts. Cilvēce jau gadu desmitiem uzkrāj dažādus datus, kas glabājās arhīvos un kartotēkās taču par to atkārotu pielietošanu un vērtības palielināšanu publiski sāk runāt tikai relatīvi nesen. Lai arī Latvijā atvērto datu avoti tikai tagad tiek radīti, var secināt, ka “valdības gaitēnos” atvērtie dati tiek uzskatīti par vērā ņemamu ieguvumu, kas Latvijai ir jāaktualizē. Par šādu novērojumu liecina valsts izstrādātās pamatnostādnes 2014. – 2020. gadam, kas turpmākajiem valsts IT projektiem uzliek par pienākumu integrēt atvērto datu servisu. Šādu pamatnostādņu izstrādi noteikti ietekmējušas arī saistošās ES direktīvas, kas nosaka datu atkalizmantojamību un publisku pieejamību akcentēt, taču nenoliedzams ir fakts, ka arī valdība šādu pieeju atbalsta. To pierāda e-pasta sarakste ar VARAM amatpersonu T. Ceļmilleru un viņa izteikumi par atvērto datu pozitīvajiem piemēriem, ar kuriem pašlaik tiek uzrunātas citas valsts iestādes, lai stimulētu datu publicēšanu ar esošajiem līdzekļiem.

Atvērto datu avotiem nenoliedzami Latvijā vēl ir jāaug, jo pieejamie risinājumi ir tikai ceļā uz kvalitatīva datu servisa veidolu, pēc “piecu zvaigžņu” vērtējuma. Par pamatu attīstībai jāuzsver pozitīvie ietekmējošie faktori, kā piemēram valsts pārvaldes carspīdības palielināšana, jaunu darba vietu, biznesa veidu, pakalpojumu un servisu rašanos, kā arī finansiālo līdzekļu pieplūdumu valsts kasē. Šie ietekmējošie faktori liek secināt, ka atvērtajos datos ir neizmantots potenciāls, ko pierāda pasaules prakse.

## IZMANTOTĀ LITERATŪRA

1. What is big data? [tiešsaiste]. IBM [atsauce 02.05.2014]. Pieejams internetā: <http://www-01.ibm.com/software/au/data/bigdata/>
2. The Definition [tiešsaiste]. Open Definition [atsauce 02.05.2014]. Pieejams internetā: <http://opendefinition.org/od/>
3. Atvērtā valstī nepieciešami atvērti dati [tiešsaiste]. LATA [atsauce 05.05.2014] Pieejams internetā: <http://lata.org.lv/atverta-valsti-nepieciesami-atverti-dati/>
4. Is your Linked Open Data 5 star [tiešsaiste]. World Wide Web [atsauce 08.05.2014]. Pieejams internetā: <http://www.w3.org/DesignIssues/LinkedData.html>
5. Transparency [tiešsaiste]. Wikipedia [atsauce 11.05.2014]. Pieejams internetā: [http://en.wikipedia.org/wiki/Transparency\\_\(behavior\)#Politics](http://en.wikipedia.org/wiki/Transparency_(behavior)#Politics)
6. Why open data? [tiešsaiste]. Open data handbook [atsauce 11.05.2014]. Pieejams internetā: <http://opendatahandbook.org/en/why-open-data/index.html>
7. Eiropas parlamenta un padomes Direktīva 2003/98/EK par valsts sektora datu atkalizmantošanu [tiešsaistē] EUR-Lex Piekļuve ES tiesību aktiem [atsauce 16.05.2014]. Pieejams internetā: <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32003L0098>
8. Eiropas parlamenta un padomes Direktīva 2013/37/ES, kas groza Direktīvu 2003/98/EK [tiešsaistē] EUR-Lex Piekļuve ES tiesību aktiem [atsauce 16.05.2014]. Pieejams internetā: <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013L0037>
9. Informācijas sabiedrības attīstības pamatnostādnes 2014. – 2020. gadam [tiešsaistē]. Politiskās plānošanas dokumentu datu bāze [atsauce 18.05.2014] Pieejams internetā: <http://polsis.mk.gov.lv/view.do?id=4518>
10. Par datu atvērtību [tiešsaistē]. [atsauce 21.05.2014]. Pieejams tiešsaistē: <http://opendata.lv/>
11. Informācijas atklātības likums [tiešsaistē]. Latvijas republikas tiesību akti [atsauce 24.05.2014] Pieejams tiešsaistē <http://likumi.lv/doc.php?id=50601>
12. Neelie Kroes: Information is new oil [tiešsaiste]. Webzine of the Young European federalists [atsauce 25.04.2014] Pieejams internetā: <http://www.thenewfederalist.eu/Neelie-Kroes-Information-is-the-new-oil,04732>
13. Piekļuve ES tiesību aktiem [atsauce 25.05.2014]. Pieejams internetā: <http://eur-lex.europa.eu/>
14. Valsts pārvaldes iekārtas likums [tiešsaistē]. Latvijas republikas tiesību akti [atsauce 26.05.2014]. Pieejams internetā: <http://likumi.lv/doc.php?id=63545>

15. Uzņēmumu reģistra dati būs pieejami ik vienam [tiešsaistē]. Tvnet ziņu protāls [atsauce 26.05.2014] Pieejams tiešsaistē: [http://financenet.tvnet.lv/zinas/511656-uznemumu\\_registra\\_dati\\_bus\\_pieejami\\_ikvienam](http://financenet.tvnet.lv/zinas/511656-uznemumu_registra_dati_bus_pieejami_ikvienam)
16. Open Government and Transparency: The UK experience [tiešsaistē]. World bank [atsauce 27.05.2014]. Pieejams internetā: <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTINFORMATIONANDCOMMUNICATIONANDTECHNOLOGIES/EXTDEVELOPMENT/0,,contentMDK:23101784~pagePK:210058~piPK:210062~theSitePK:559460,00.html>
17. Joel Gurin, *Open Data Now*, McGraw-Hill, 2013
18. Treasury Board has 15 apps that prove its open data strategy is working[tiešsaistē]. IT World Canada [atsauce 26.05.2014]. Pieejams internetā: <http://www.itworldcanada.com/article/treasury-board-has-15-apps-that-prove-its-open-data-strategy-is-working/90719>
19. 30 places to find open data on the web [tiešsaistē]. Original cisual contens for brands [atsauce 28.05.2014]. Pieejams internetā: <http://blog.visual.ly/data-sources/>
20. The Home of the U.S. Government's Open Data [atsauce 28.05.2014]. Pieejams internetā: <http://www.data.gov/>
21. San Franisco Data [atsauce 28.05.2014]. Pieejams internetā: <https://data.sfgov.org/analytics>
22. *Latvijas Republikas Satversme*, Satversmes Sapulce, Feb. 1922.
23. Digital agenda: Turning government into gold[tiešsaistē] EU press releases database [atsauce 28.05.2014]. Pieejams internetā: [http://europa.eu/rapid/press-release\\_IP-11-1524\\_lv.htm](http://europa.eu/rapid/press-release_IP-11-1524_lv.htm)
24. The benefits of open data [tiešsaistē]. Open Knowledge Foundation Blog [atsauce 28.05.2014]. Pieejams internetā: <http://blog.okfn.org/2012/10/05/the-benefits-of-open-data-evidence-from-economic-research/>

Bakalaura darbs "Atvērtie dati Latvijā" izstrādāts LU Datorikas fakultātē.

Ar savu parakstu apliecinu, ka pētījums veikts patstāvīgi, izmantoti tikai tajā norādītie informācijas avoti un iesniegtā darba elektroniskā kopija atbilst izdrukai.

Autors: Matīss Jurgēlis

Rekomendēju/nerekomendēju darbu aizstāvēšanai (nevajadzīgo izsvītrot)

Vadītājs: dr.dat. Uldis Bojārs

Recenzents:

Darbs iesniegts Datorikas fakultātē 02.06.2014.

Dekāna pilnvarotā persona: vecākā metodiķe Ārija Sproģe

Darbs aizstāvēts bakalaura gala pārbaudījuma komisijas sēdē

prot. Nr.

Komisijas sekretār\_\_: