

TALLINNA TEHNIKAÜLIKOOL  
Infotehnoloogia teaduskond  
Informaatikainstituut

IDU40LT  
Kajar Karuauk 042373IABB

**Masinloetavate avaandmete esitamine  
Lennuameti näitel**

bakalaureusetöö

Juhendaja: Enn Õunapuu  
Phd  
dotsent

Tallinn  
2016

---

## **Autorideklaratsioon**

Kinnitan, et olen koostanud antud lõputöö iseseisvalt ning seda ei ole kellegi teise poolt varem kaitsmisele esitatud. Kõik töö koostamisel kasutatud teiste autorite tööd, olulised seisukohad, kirjandusallikatest ja mujalt pärinevad andmed on töös viidatud.

---

*(kuupäev)*

*(allkiri)*

## **Annotatsioon**

Käesoleva bakalaureusetöö eesmärgiks on analüüsida avaandmeid ja Lennuohutuse järelvalve infosüsteemi avaandmete avalikustamise rakendust.

Bakalaureusetöös leiab käsitlemist Berners-Lee viie täрни süsteem, mille kolmanda täрни nõudeid tuleks rakendada vastavalt avaandmete rohelisele raamatule avaliku sektori andmekogude avaandmete avalikustamisel. Käesolevas töös analüüsib autor samuti eelnimetatud kolmanda täрни nõutele vastavaid andmevorminguid – CSV, XML ja JSON.

Bakalaureusetöö keskseks uurimisküsimuseks on, kas Lennuohutuse järelvalve infosüsteem vastab avaliku sektori avaandmete avalikustamise nõuetele.

Töö praktiliseks tulemuseks on soovitud Lennuohutuse järelvalve infosüsteemi täiendamiseks.

Lõputöö on kirjutatud eesti keeles ning sisaldab teksti 25 leheküljel, 3 peatükki, 7 joonist.

## **Abstract**

### **Presentation of the machine-readable open-data by example of Civil Aviation Administration**

The objective of this Bachelor's thesis is to analyze open-data and the publishing of open-data of the Aviation Safety Information System.

In this thesis author examines the five stars' system of Berners-Lee that encompasses requirements of the third star that should be applied according to the Green Book to the publishing of open-data of the public sector. The author also analyzes formats that meet those requirements referred to above – CSV, XML and JSON.

The main research question of the Bachelor's thesis is whether the Aviation Safety Information System is in compliance with the requirements of the publishing of open-data of the public sector.

The practical result of the Bachelor's thesis are suggestions how to amend the Aviation Safety Information System.

The thesis is in Estonian and contains 25 pages of text, 3 chapters, 7 figures.

## Lühendite ja mõistete sõnastik

<b>Avaandmed</b>	<b><i>Open data</i></b> Avaandmed on masinloetavas formaadis andmed, mis on mõeldud kõigile vabalt ja avalikult kasutamiseks ning millel puuduvad kasutamist ning levitamist takistavad piirangud.
<b>Avaliku sektori avaandmed</b>	<b><i>Open government data</i></b> Avaliku sektori poolt publitseeritud juurdepääsupiiranguteta avaandmed.
<b>Avaandmete viie täрни süsteem</b>	<b><i>5 * deployment scheme for open data</i></b> Tim Berners Lee poolt pakutud vormingute ja kodeeringute süsteem avaandmete avalikustamiseks.
<b>AvTS</b>	<b><i>Public Information Act</i></b> Avalik teabe seadus
<b>CSV</b>	<b><i>Comma Separated Values</i></b> Failivorming, kus iga rida tähendab ühte kirjet ja veerud eraldatakse üksteisest enamasti semikoolonitega või komadega.
<b>JSON</b>	<b><i>Javascript Object Notation</i></b> Andmevahetuse formaat, mis koosneb nimi-väärtus paaride kollektsioonidest.
<b>LOIS</b>	<b><i>ASIS (Aviation Safety Information System)</i></b> Lennuohutuse järelvalve infosüsteem
<b>Masinloetav formaad</b>	<b><i>Machine-readable format</i></b> Failiformaat, mis on struktureeritud selliselt, et tarkvararakendused suudavad andmeid ja nende sisemist struktuuri kergelt tuvastada, ära tunda ja välja lugeda.
<b>Teek</b>	<b><i>Library</i></b>

Korduvkasutamiseks mõeldud funktsioonide, klasside, makrode kogum.

**Windows-i teenus**     *Windows service*

Arvuti programm, mis töötab Windows operatsioonisüsteemi taustal.

**XML**     *Extensible MarkupLanguage*

W3C poolt loodud laiendatav märgistuskeel struktureeritud info jagamiseks erinevate infosüsteemide vahel.

**XSD**     *XML Schema Defination*

Fail, milles on kirjas reeglid, mida XML fail sisaldada tohib.

## Sisukord

1. Sissejuhatus .....	9
1.1 Taust ja probleem .....	9
1.2 Ülesande püstitus .....	9
1.3 Metoodika .....	10
1.4 Ülevaade tööst .....	10
2. Avaandmete analüüs .....	11
2.1 Avaandmete portaal .....	12
2.2 Avaandmete formaat .....	13
2.2.1 CSV .....	13
2.2.2 XML .....	14
2.2.3 JSON .....	16
2.3 Lennuameti näide .....	19
2.3.1 Lennuohutuse järelevalve infosüsteem .....	19
2.3.2 LOIS-i avaandmete seis praegu .....	20
2.3.3 Ärianalüüs .....	20
2.3.4 Disaini kavand .....	22
3. Kokkuvõte .....	24
Kasutatud kirjandus .....	25

## **Jooniste nimekiri**

Joonis 1. CSV näide .....	14
Joonis 2. XML-i näide .....	15
Joonis 3. XML-i skeemi (XSD) näide .....	16
Joonis 4. JSON-i näide .....	17
Joonis 5. JSON-i skeemi näide .....	18
Joonis 6. LOIS-i arhitektuur .....	20
Joonis 7. Arhitektuuri kavand.....	23



# 1. Sissejuhatus

Avaliku sektori asutused toodavad ning koguvad üsna suurel hulgal andmeid, mida nad vajavad enda igapäevatööks. Neid andmeid kasutades on loodud mitmeid avalikke teenuseid. Teenustele on võimalik ligi pääseda läbi riigiportaali eesti.ee ([www.eesti.ee](http://www.eesti.ee)), selle kaudu saavad nii füüsilised kui ka juriidilised isikud neid teenuseid tarbida. Näiteks on kodanikul võimalik selle portaali kaudu esitada erinevaid taotlusi, vaadata enda sõidukite andmeid või kontrollida juhiloa kehtivust. Samu andmeid, mida avalik sektor kogub, sooviksid enda huvides kasutada ka füüsilised ja juriidilised isikud. Näiteks luues nende andmete abil uusi teenuseid.

## 1.1 Taust ja probleem

Vastavalt avaliku teabe seadusele peavad avaliku sektori asutused juurdepääsupiiranguteta teabe avalikustama veebis. Enamasti näeb veebis avalikustamine välja nõnda, et andmed on loetavad vaid veebilehekülgede kaudu, mis tähendab seda, et neid andmeid on üsna raske kui mitte võimatu taaskasutada. Seoses sellega peavad avaliku sektori asutused vastavalt AvTS-ile alates 2015. aasta 1. jaanuarist enda andmekogude avaandmed tegema kättesaadavaks masinloetaval kujul. Vastavalt seadusele tegi enda hallata andmekogu Lennuohutuse järelevalve infosüsteemi andmed avalikuks ka Lennuamet.

## 1.2 Ülesande püstitus

Käesoleva bakalaureusetöö peamiseks eesmärgiks on analüüsida Lennuohutuse järelevalve infosüsteemi avaandmete avalikustamise praegust lahendust ning teha ettepanekuid selle parandamiseks.

Lisaks analüüsitakse avaandmeid üldiselt. Mis on avaandmed ning nende avalikustamise vormingud. Millised oleks sobivad vormingud andmete avalikustamisel masinloetaval kujul lähtudes Berners-Lee viie täрни süsteemist.

### **1.3 Metoodika**

Käesolevas töös kasutatakse eesmärkide saavutamiseks tegevusuuringut. Võrreldes teiste uurimisviisidega eristub tegevusuuring sellega, et selle tulemusi on võimalik kohe praktikas rakendada. Antud töö tulemusi on võimalik üsna kiiresti realselt rakendada selleks, et parandamaks LOIS-i avaandmete avalikustamise hetkeseisu.

### **1.4 Ülevaade tööst**

Töö esimeses osas analüüsitakse, mis on avaandmed. Samuti käsitletakse uurimisküsimusi, mida kujutab endast avaandmete portaal ning mis on Berners Lee viie täрни süsteem. Töö esimeses osas analüüsitakse samuti levinumaid avaandmete avalikustamise vorminguid.

Töö teises osas analüüsitakse Lennuohutuse järelevalve infosüsteemi avaandmete praegust rakendust ning pakutakse välja lahendused selle täiendamiseks.

## 2. Avaandmete analüüs

Avaandmed on masinloetavas formaadis andmed, mis on mõeldud kõigile vabalt ja avalikult kasutamiseks ning millel puuduvad kasutamist ning levitamist takistavad piirangud. Kuna avaandmeid on võimalik luua ka teisel sektoritel peale avaliku sektori, siis on võetud kasutusele mõiste avaliku sektori avaandmed. Avaliku sektori avaandmed on avaliku sektori poolt publitseeritud juurdepääsupiiranguteta avaandmed.

Avalikustamisele peaks kuuluma kõik riigiametite kui ka kohalike omavalitsuste kogutud andmed, mille kasutamine avalikult ei ole keelatud ning mis ei sisalda vaid isikuandmeid. Juhul kui andmed sisaldavad nii isikuandmeid kui ka muid andmeid, siis avalikustamisele kuuluvad vaid mitte-isikuandmete osad. Andmete avalikustamisel tuleks järgida 2007 aastal kirja pandud kaheksat printsiipi (Open Government Data Principles), mis defineerivad, millised andmed on avatud. Andmehulk on avatud, kui on täidetud järgmised tingimused:

- terviklik – avalikuks tehakse kõik andmed, välja arvatud juurdepääsupiirangutega andmed;
- algallikast – andmed on töötluseta algallikast ning säilitanud originaalkuju ja detailsuse;
- ajakohane – andmed tehakse avalikuks võimalikult kiiresti;
- kättesaadav – andmed on mõeldud laiale kasutajate ringile võimalikult laia kasutuseesmärgiga;
- masinloetav – andmed on masinloetavas formaadis ja automaatselt töödeldavad;
- mitte-diskrimineeriv – andmed on kättesaadavad kõigile, nendele ligipääsemiseks ei ole vaja kuhugi end registreerida ega taotleda juurdepääsu;
- vabu standardeid kasutatav – andmed on esitatud vabas formaadis;
- litsentsi vaba – andmed ei ole kaitstud autoriõiguste, patendi, kaubamärgi ega ärisaladuse regulatsioonidega;

Avalikustamisel on nii positiivseid kui ka negatiivseid mõjusid. Positiivsetest mõjudest võiks välja tuua:

- majanduse elavdamine andmete taaskasutuse läbi – andmete avalikustamine annaks erasektorile võimaluse andmeid kasutades luua uusi kasumlikke äriteenuseid;
- avaliku sektori läbipaistvuse suurendamine – andmete avaandmetena masinloetaval kujul publitseerimine suurendab avaliku sektori läbipaistvust;
- avaliku sektori tõhustamine – näiteks teabenõuete arvu vähendamine.

Negatiivsetest mõjudest ühed olulisemad on avalikustamisega seotud kulud. Paratamatult kaasnevad andmete masinloetaval kujul kättesaadavaks tegemisega kulud. Kulude suurus sõltub sellest kui keeruline on andmete avalikustamine. Näiteks mõne andmebaasi päringu kirjutamine ja tulemuste eksport XML faili väga palju ei maksa. Samas kui soovitakse veebiteenust, kus saaks andmeid pärida erinevate parameetrite järgi, siis selle loomine on nii rahaliselt kui ka ajaliselt kallim. Lisaks avalikustamisega kaasnevate kuludega võib avalikustamine vähendada ka tulusid. Näiteks mitmed infosüsteemid (nt. Äriregister) pakuvad enda kogutud andmeid tasu eest.

## 2.1 Avaandmete portaal

Avaliku sektori poolt pakutavatest avaandmetest ülevaate saamiseks on loodud avaandmete portaal aadressiga <https://opendata.riik.ee>. Portaali eesmärk on koondada avaliku sektori avaandmetega seotud informatsioon ühte kohta. Antud kodulehelt on võimalik leida andmekogud, kus avaandmeid pakutakse ning vajadusel on võimalik andmed alla laadida. Lisaks saab portaalis lugeda avaandmete teemalisi uudiseid, leida andmete põhjal loodud rakendusi ning osaleda foorumis antud teemaga seotud diskussioonides.

Portaalist on võimalik allalaadida ka avaandmete roheline raamat (<https://opendata.riik.ee/roheline-raamat>). Avaandmete roheline raamat on välja antud Majandus –ja Kommunikatsiooniministeeriumi riigi infosüsteemide osakonna poolt. Raamat on mõeldud neile, kes soovivad antud valdkonnaga lähemalt tutvust teha.

## 2.2 Avaandmete formaat

AvTS ütleb, et avaandmed peavad olema kättesaadavad masinloetaval kujul. Seda seadus ei ütle, et millises masinloetavas formaadis avaandmed avalikustatud peaksid olema. Avaandmete roheline raamat annab vormingu osas soovitusi, et andmed võiks avaldada 3 täрни formaadis. Tärnide all peetakse silmas Berners-Lee poolt 2010. aastal soovitatud viie täрни süsteemi andmete avalikustamisel. Berners-Lee poolt välja pakutud tärnide tähendused on järgmised:

- 1 tärn – andmed on veebis kättesaadavad, formaat ei ole oluline.
- 2 täрни – andmed on veebis kättesaadavad masinloetavas formaadis, kuid vorming on seotud mingi kindla tootega (nt. Excel).
- 3 täрни – andmed on veebis kättesaadavad masinloetavas formaadis ja formaat ei ole seotud mingi konkreetse tootega.
- 4 täрни – andmed on veebis kättesaadavad masinloetavas formaadis, kuid objektide identifitseerimiseks kasutatakse unikaalseid URI-sid.
- 5 täрни - andmed on veebis kättesaadavad masinloetavas formaadis ning on URI-ide abil lingitud teiste andmekogude andmetega.

Kuna avaandmete roheline raamat soovitab avaandmete puhul vähemalt 3 täрни ehk selliseid formaate, mis ei ole seotud mõne konkreetse tootega, siis järgnevalt võtan vaatluse alla CSV, JSON ja XML formaadid.

### 2.2.1 CSV

CSV (Comma Separated Values) on failivorming, kus iga rida tähendab ühte kirjet ja veerud eraldatakse üksteisest enamasti semikoolonitega või komadega (vt joonis 1).

CSV vormingu plussid:

1. andmetest on kerge aru saada;
2. CSV-d on lihtne genereerida;
3. palju teeke erinevates programmeerimiskeeltes.

CSV vormingu miinused:

1. liiga lame, puudub hierarhia;
2. keeruline valideerida;
3. andmetüübid puuduvad – arvulisi ja tekstilisi väärtusi on raske eristada.

```
Year,Make,Model,Description,Price
1997,Ford,E350,"ač, abs, moon",3000.00
1999,Chevy,"Venture ""Extended Edition""",",",4900.00
1999,Chevy,"Venture ""Extended Edition, Very Large""",,5000.00
1996,Jeep,Grand Cherokee,"MUST SELL!
air, moon roof, loaded",4799.00
```

### Joonis 1. CSV näide

#### 2.2.2 XML

XML (Extensible Markup Language) on W3C poolt loodud laiendatav märgistuskeel struktureeritud info jagamiseks erinevate infosüsteemide vahel (vt joonis 2). Erinevalt JSON-ist kasutatakse XML-i üsna tihti ka väljaspool veebikeskkonda. Näiteks eksporditakse andmeid andmebaasist mõnda XML faili.

```

<?xml version="1.0"?>
<x:raamatud xmlns:x="urn:raamatud">

  <raamat id="1">
    <autor>Kivirähk</autor>
    <pealkiri>Mees, kes teadis ussisõnu</pealkiri>
    <kategooria>ilukirjandus</kategooria>
    <hind>44.95</hind>
    <avalikustamisKp>2000-10-01</avalikustamisKp>
    <ylevaade>Tore raamat ussidest.</ylevaade>
  </raamat>

  <raamat id="2">
    <autor>Kivirähk</autor>
    <pealkiri>Mees, kes teadis ussisõnu</pealkiri>
    <kategooria>ilukirjandus</kategooria>
    <hind>44.95</hind>
    <avalikustamisKp>2000-10-01</avalikustamisKp>
    <ylevaade>Tore raamat ussidest.</ylevaade>
  </raamat>

</x:raamatud>

```

## Joonis 2. XML-i näide

XML-i plussid:

1. skeemi tugi; skeem (vt jooni 3) on XML fail, kus on kirjas reeglid, mida XML fail sisaldada tohib; vastu seda skeemi toimub XML faili valideerimine;
2. andmetüübid on olemas; lisaks on võimalik ise defineerida uusi andmetüüpe;
3. hierarhiline;
4. palju teeki erinevates programmeerimiskeeltes.

```

<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
            targetNamespace="urn:raamatud"
            xmlns:bks="urn:raamatud">

    <xsd:element name="raamatud" type="bks:raamatudType"/>

    <xsd:complexType name="raamatudType">
        <xsd:sequence>
            <xsd:element name="raamat"
                        type="bks:raamatType"
                        minOccurs="0"
                        maxOccurs="unbounded"/>
        </xsd:sequence>
    </xsd:complexType>

    <xsd:complexType name="raamatType">
        <xsd:sequence>
            <xsd:element name="autor" type="xsd:string"/>
            <xsd:element name="pealkiri" type="xsd:string"/>
            <xsd:element name="kategooria" type="xsd:string"/>
            <xsd:element name="hind" type="xsd:float" />
            <xsd:element name="avalikustamisKp" type="xsd:date" />
            <xsd:element name="ylevaade" type="xsd:string"/>
        </xsd:sequence>
        <xsd:attribute name="id" type="xsd:string"/>
    </xsd:complexType>
</xsd:schema>

```

**Joonis 3. XML-i skeemi (XSD) näide**

### 2.2.3 JSON

JSON (Javascript Object Notation) on andmevahetuse formaat, mis koosneb nimi-väärtus paaride kollektsioonidest (vt joonis 4). Kuigi JSON pärineb Javascripti keelest, on ta siiski programmeerimiskeelest sõltumatu formaat. Peamiselt leiab JSON kasutamist andmete vahetuseks serveri ja veebilehitseja vahel. Väljaspool veebikeskkonda kasutatakse JSON-i praegu veel üsna harva.



```

[
  {
    "id": 2,
    "name": "An ice sculpture",
    "price": 12.50,
    "tags": ["cold", "ice"],
    "dimensions": {
      "length": 7.0,
      "width": 12.0,
      "height": 9.5
    },
    "warehouseLocation": {
      "latitude": -78.75,
      "longitude": 20.4
    }
  },
  {
    "id": 3,
    "name": "A blue mouse",
    "price": 25.50,
    "dimensions": {
      "length": 3.1,
      "width": 1.0,
      "height": 1.0
    },
    "warehouseLocation": {
      "latitude": 54.4,
      "longitude": -32.7
    }
  }
]

```

#### Joonis 4. JSON-i näide

JSON-i plussid:

1. Javascriptis lihtne kasutada;
2. hierarhiline;
3. skeemi tugi (vt jooni 5); sarnaselt XML-iga on ka JSON-il olemas skeemi tugi, mis võimaldab JSON-i valideerida;
4. palju teeki erinevates programmeerimiskeeltes.

```

{
  "$schema": "http://json-schema.org/draft-04/schema#",
  "title": "Product set",
  "type": "array",
  "items": {
    "title": "Product",
    "type": "object",
    "properties": {
      "id": {
        "description": "The unique identifier for a product",
        "type": "number"
      },
      "name": {
        "type": "string"
      },
      "price": {
        "type": "number",
        "minimum": 0,
        "exclusiveMinimum": true
      },
      "tags": {
        "type": "array",
        "items": {
          "type": "string"
        },
        "minItems": 1,
        "uniqueItems": true
      },
      "dimensions": {
        "type": "object",
        "properties": {
          "length": {"type": "number"},
          "width": {"type": "number"},
          "height": {"type": "number"}
        },
        "required": ["length", "width", "height"]
      },
      "warehouseLocation": {
        "description": "Coordinates of the warehouse with the product",
        "$ref": "http://json-schema.org/geo"
      }
    },
    "required": ["id", "name", "price"]
  }
}

```

## Joonis 5. JSON-i skeemi näide

JSON-i miinused:

1. Javascriptis reserveeritud sõnu ei tohi kasutada.

## 2.3 Lennuameti näide

### 2.3.1 Lennuohutuse järelevalve infosüsteem

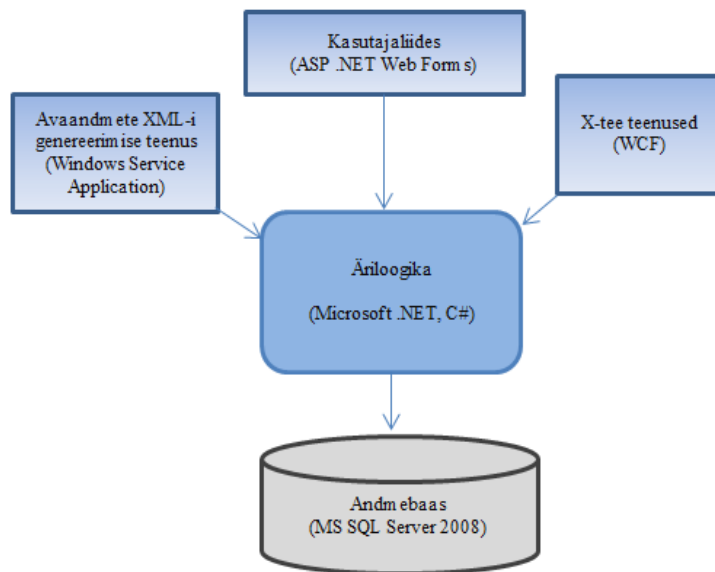
Lennuameti järelevalve infosüsteem (LOIS) on Lennuameti hallatav andmekogu, mis võimaldab ettevõtjatel ja eraisikutel esitada elektroonilisi taotlusi tegevuslubade ja sertifikaatide saamiseks, samuti saab kasutaja jälgida temaga seotud menetluste käiku ja teha erinevaid päringuid teda puudutavate andmete kohta. Lisaks võimaldab LOIS läbi viia erinevaid järelevalve menetlusi. Näited LOIS-is teostatavate menetluste kohta:

- Lennuvälja / kopteriväljaku sertifikaadi taotlemine;
- Lennuvälja / kopteriväljaku sertifikaadi muutmise taotlemine;
- Lennuvälja / kopteriväljaku sertifikaadi järelevalve.

LOIS-i tehniline informatsioon:

- raamistik – Microsoft .NET 3.5;
- andmebaas – Microsoft SQL Server 2008;
- rakendusserver – Internet Information Services 7 (IIS 7);
- realiseeritud C# programmeerimiskeeles.

LOIS on kolme-kihiline rakendus, kus on eraldatud kasutajaliidese, ärioloogika ja andmebaasi kiht (vt. Joonis 6).



## Joonis 6. LOIS-i arhitektuur

### 2.3.2 LOIS-i avaandmete seis praegu

LOIS-i avaandmed on allalaetavad aadressilt: <https://lois.ecaa.ee/avaandmed/aa.xml>. Andmed on kättesaadavad ühe suure XML failina, mida uuendatakse kord kuus. XML fail sisaldab infot LOIS-is läbiviidud erinevate menetluste.

### 2.3.3 Ärianalüüs

Praegune lahendus vastab Berners-Lee viie täрни süsteemis kolmele tärnile, kuid siiski on praegusel lahendusel mõned olulised miinused:

- andmete leitavus ja atraktiivsus;
- andmete filtreerimise võimalus puudub, andmed on kättesaadavad ühe suure XML failina ning puudub võimalus pärida andmeid filtreid kasutades;
- avaandmed ei ole aktuaalsed, sest avaandmete XML faili värskendatakse kord kuus.

### **2.3.3.1 Andmete leitavus ja atraktiivsus**

Otsides internetist infot LOIS-i avaandmete kohta selgus tõsiasi, et ega seda infot väga ei ole. Majandus – ja kommunikatsiooniministeeriumi kodulehelt (<https://www.mkm.ee/et/avaandmed>) võib leida viite, et Lennuohutuse järelevalve infosüsteem pakub avaandmeid, kuid näiteks avaandmete portaalist ei leia ühtegi märget Lennuameti poolt pakutavate avaandmete kohta.

Andmete atraktiivsuse all pean silmas seda, et kui vaadata praegust avaandmete XML faili, siis sellest on üsna raske aru saada. Failist võib lugeda, et LOIS-is on tehtud palju erinevaid menetlusi, sisestatud erinevaid dokumente, väljastatud erinevaid sertifikaate ja lubasid, kuid puudub informatsioon selle kohta, et milleks need andmed head on ja kes nendest andmetest kasu võiks saada.

Lennuameti kodulehel ja avaandmete portaalis peaks olema rohkem infot selle kohta, mis andmeid on LOIS-ist võimalik pärida.

Lisaks võiks kaaluda ka seda, et teha rohkem infot avalikuks. Praegu on avalikud vaid taotluste ja muutmiste menetlused, milledega näiteks taotletakse sertifikaate või muudetakse sertifikaatide kehtivuse aega, kuid LOIS-is on olemas ka järelevalve menetlused. Järelevalve menetlusi teostatakse selleks, et kontrollida sertifikaatide omanike vastavust nõuetele. Kui järelevalve käigus leitakse rikkumisi, siis need rikkumised fikseeritakse ja antakse aega nende kõrvaldamiseks. Minu arvates need leitud rikkumised on andmed, mis võiks potentsiaalsetele kasutajatele huvi pakkuda. Näiteks võib mõni lennundus huviline soovida teha statistikat leitud rikkumiste. Need andmed võiks olla avalikud, ei pea konkreetselt viitama mõnele ettevõttele või isikule. Järelevalve menetlusi tuleks analüüsida sellise pilguga, et äkki siiski on võimalik seal mingeid andmeid avalikuks teha.

### **2.3.3.2 Andmete filtreerimine**

Ühe suure XML faili miinuseks on tema suurus. Pean siin silmas nii füüsilist suurust kui ka andmete hulka. Praegune avaandmete lahendus eeldab suuresti seda, et kui keegi soovib enda teenuses kasutada LOIS-i avaandmeid, siis ta laeb selle XML faili enda masinasse ning siis hakkab seda töötleva. Loomulikult on võimalik ehitada teenus ka nõnda, et päritakse andmeid jooksvalt, aga arvestades seda, XML pidevalt suureneb, siis võib hakata see pikapeale mõju avaldama loodud teenuse töökindlusele (nt. päringud muutuvad aeglasemaks).

Lisaks on ühte XML faili pandud liiga palju infot ja vajaliku info kättesaamiseks tuleb andmeid eelnevalt palju töödelda.

Praegune suur XML fail tuleks jaotada osadeks, milleks on minu arvates järgnevad võimalused:

- luua mitu erinevat staatilist XML faili (nt. eraldi failid LOIS-is teostatud menetluste kohta ja väljastatud sertifikaatide kohta). Sellega tagame selle, et andmete hulk ühes XML failis ei ole enam nõnda suur;
- ehitada veebiteenus mitme erineva päringuga, kus on võimalik andmeid parameetrite järgi filtreerida.

### **2.3.3.3 Andmete aktuaalsus**

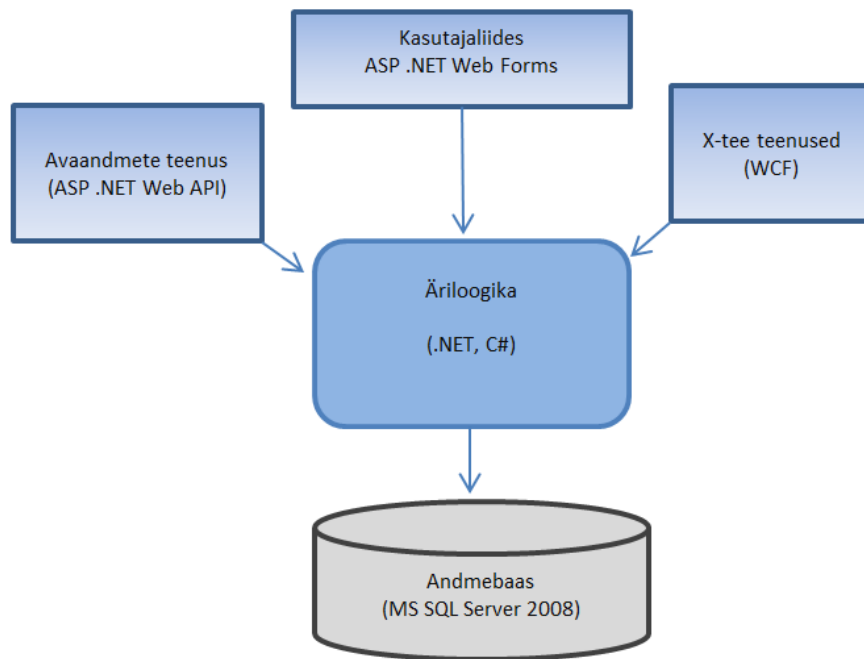
Avaandmete XML-i genereerimine kord kuus põhjustab selle, et andmed ei ole aktuaalsed. Kui keegi soovib luua teenuseid LOIS-i avaandmeid kasutades, siis ta peab arvestama sellega, et andmed ei ole kõige värskeamad.

Kõige kiirema võimaliku lahendusena võiks XML faili genereerida kord ööpäevas, kuid luues veebiteenuse, millega on võimalik otse LOIS-i andmebaasist andmeid pärida, langeb andmete aktuaalsus probleem samuti ära.

### **2.3.4 Disaini kavand**

Kui mõned asjad (nt. XML faili genereerimise tihedus või andmete kättesaadavus avaandmete portaalis) on võimalik parandada võimalikult väikese aja- ja ressursikuluga, siis näiteks veebiteenuse ehituseks kulub veidi rohkem aega ja vahendeid. Kuid arvestades seda, et praegune rakendus on realiseeritud .NET raamistikus ning on olemas infrastruktuur selle töös hoidmiseks, siis oleks mõistlik uus lahendus samuti realiseerida .NET raamistikus. Näiteks ehitada uus teenus LOIS-i kõrvale (vt. Joonis 7) ning kasutades selleks ASP .NET Web API raamistikku.

Praegune avalikustamise formaat, milleks on XML, on igati sobiv. Kuid samas annaks lahendusele lisandväärtust näiteks JSON formaadi tugi. Realiseerides uue lahenduse ASP .NET Web API raamistikus on nende kahe formaadi koos kasutuselevõtmine väga lihtne (parameetriga juhitud).



**Joonis 7. Arhitektuuri kavand**

### **3. Kokkuvõte**

Töö eesmärgiks oli analüüsida avaandmeid, mis need on ning tutvustada erinevaid formaate, milles võiks avaandmed avalikuks teha.

Avalikustamiseks sobiva formaadi valikul peaks alustama sellest, et mis tasemel Berners-Lee viie täрни süsteemis andmeid avalikustada soovitakse. Kui eesmärk on neljas või viies tärn, siis tuleks valida XML formaat, sest see on kõige sobivam andmete linkimiseks. Juhul kui piirduakse kolme tärniga, siis oleks soovitatav valida JSON või XML formaadi vahel.

Teine eesmärk oli analüüsida praegust Lennuohutuse järelevalve infosüsteemi avaandmete avalikustamise rakendust.

Kokkuvõtvalt võib öelda, et avaandmete avalikustamise nõuded on täidetud, formaat on masinloetav ning vastab Berners-Lee viie täрни süsteemis kolmele tärnile, mis on ka avaandmete rohelise raamatu soovitus. Kuid praegusel lahendusel on ka mõned olulised miinused, mis on takistuseks uute teenuste loomisel, kasutamaks LOIS-i poolt pakutavaid avaandmeid.



## Kasutatud kirjandus

<http://opengovdata.org/>

<https://opendata.riik.ee/>

[https://public.resource.org/8\\_principles.html](https://public.resource.org/8_principles.html)

<http://riso.ee/et/avaandmete-roheline-raamat>

<https://open-data.europa.eu/et/data>