UNIVERSITY OF TARTU
Institute of Computer Science
Computer Science Curriculum

Arkadi Statsenko

Estonian Data Tracker (Andmejälgija) Notifier

Bachelor's Thesis (9 ECTS)

Supervisor:
Daniel Würsch, MSc

Tartu 2025

Estonian Data Tracker (Andmejälgija) Notifier

Abstract:

Andmejälgija is a protocol developed by Information System Authority (RIA), the purpose of which is to provide a uniform interface for querying Estonian residents' data access logs. There is also an *Andmejälgija* web-view accessible from the state-portal Eesti.ee.

The purpose of this thesis is to create a mobile application that would notify it's users of updates in the access logs, letting them know that their data in some state database has been accessed. Implementation choices of different aspects of the solution are also going to be covered together with advantages and disadvantages of each. Additionaly, the overview of the existing state databases will be provided, including whether they provide access logs or not.

Keywords: Andmejälgija

CERCS: CODE Code name

Andmejälgija teavitaja

Lühikokkuvõte:

Käesoleva lõputöö eesmärk on luua rakendus, mis teavitaks kasutajaid juurdepääsulogide uuendustest, andes neile teada, et nende andmeid mõnes riigi andmebaasis on kasutatud. Rakendusel on olnud erinevad rakendusvariandid, mida käsitletakse ka koos igaühe eeliste ja puudustega. Lisaks antakse ülevaade olemasolevatest riiklikest andmebaasidest, sealhulgas sellest, kas nad pakuvad juurdepääsulogisid või mitte.

Võtmesõnad: Andmejälgija

CERCS: KOOD Koodi nimi eesti keeles

Contents

1. Introduction

In Estonia there are a lot of state databases holding users' data, like Population Registry (*Rahvastikuregister*) and Health Portal (*Terviseportaal*). People's data in these databases is accessed by different parties for the variety of purposes. Usually those are legitimate purposes, like a doctor accessing person's health data, or even people themselves accessing their data in some information systems. However, sometimes the purpose of data access is not clear.

For the purposes of making the process more transparent, the Estonian Information System Authority (RIA) created a special service, Data Tracker (*Andmejälgija*), in 2017, through which users can check which parties have accessed their data. [1]

The Data Tracker is a people-oriented service on the state portal eesti.ee, which aims to ensure transparency in the processing of personal data in the public sector. The data tracker relies on the ability of each data repository to store the data processing taking place within itself in the form of logs, in order to later display it to the individual, i.e. the data subject, via the service on eesti.ee.[1]

Architecturally, it is a fully distributed system, i.e. the information displayed to the user comes directly from the database that implemented the Data Tracker service. At the user's request, eesti.ee makes a query to each of the Data Tracker services and displays the query response without saving it.[1]

The Data Tracker should display to the individual information about data processing taking place locally in the database (activities of officials-employees with personal data) as well as an overview of when data has been transferred to a third party (via X-Road to another government agency, company, etc.).[1]

The Data Tracker doesn't notify, however, when the data is accessed by someone. In order to learn about the update in the data access logs, the person has to go to the eesti.ee web-view and manually query access logs from specific databases.

The primary objective of this thesis is to solve this problem by creating a mobile phone app that would notify it's users about near-real time updates in the data access logs.

---

[1] https://www.err.ee/590454/leht-rahvastikuregistris-nuhitakse-ebaseaduslikult-inimeste-andmetes

Additionally I would like to examine existing state databases, including whether they provide access logs or not.

2. Andmejälgija

2.1 The protocol

Andmejälgija is a protocol that state databases are responsible for implementing themselves. In order for the database to offer an Andmejälgija service, they have to create an X-Road interface according to RIA specification[2].

X-Road is a REST-based protocol which is used for secure data exchange between Estonian information systems over the Internet.

The Andmejälgija X-Road interface is expected to have the following endpoints:

findUsage

A query searches the data recorder database for usage records that match the constraints given in the input. The output of the query returns all records found.[2]

usagePeriod

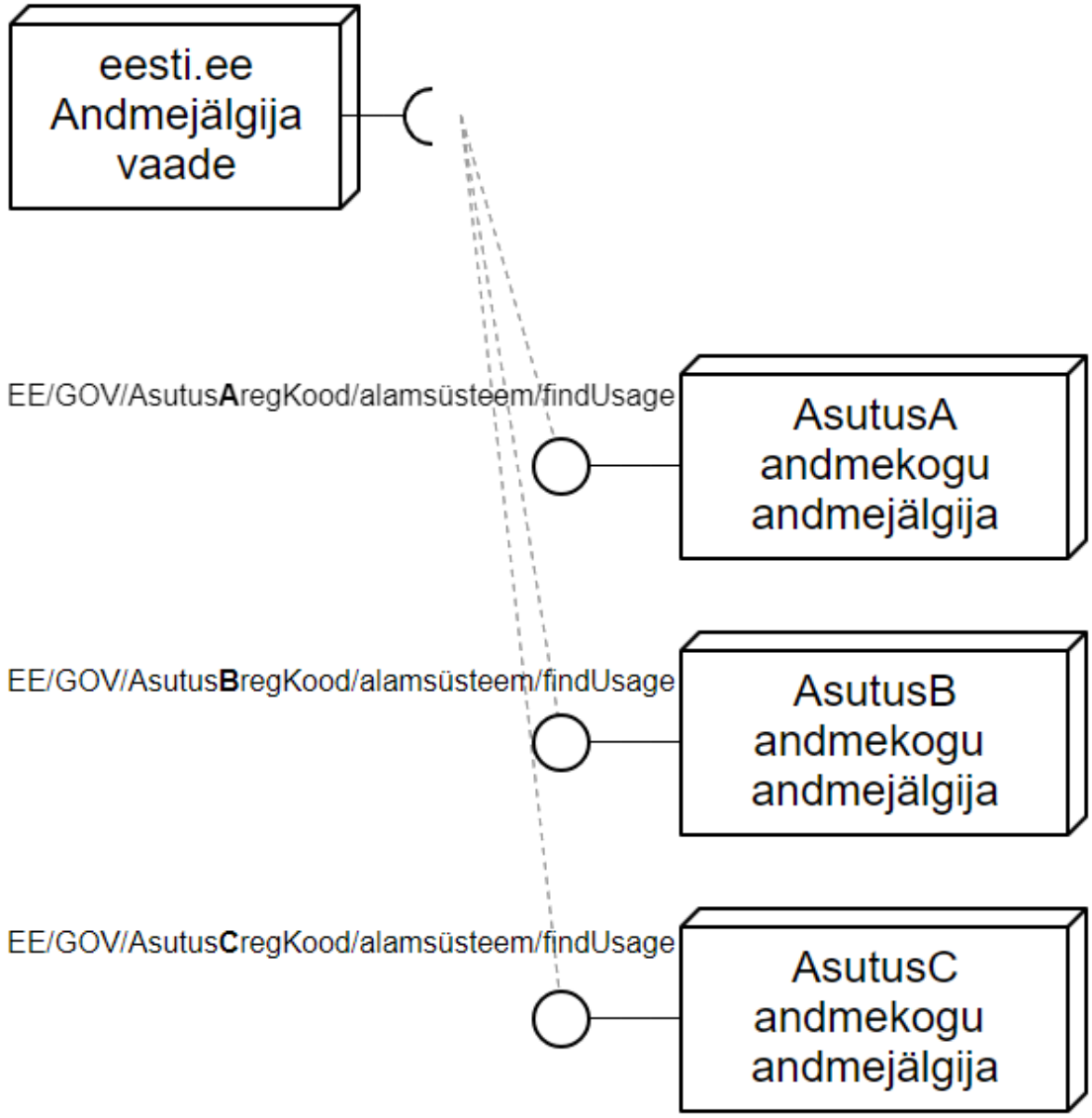
The time period for which usage information can be requested.[2]

heartbeat

Requesting the availability status of the tracker's usage information.[2]

---

[2] https://github.com/e-gov/AJ/blob/master/doc/spetsifikatsioonid/Kasutusteabe_esitamise_protokoll.md



2.2 Usage

There are several ways for end-users to access their Data Tracker data. State portal eesti.ee provides a web-view for Andmejälgija where end users can access their data access logs. Recently, RIA have also published a mobile app for eesti.ee, that also features a view of data access logs.

Another way to access Andmejälgija data is through Rahvastikuregister, although, for some reason, not all state databases that are covered by Eesti.ee are covered by Rahvastikuregister.

2.3 Adoption

The current adoption of Andmejälgija is flawed, to say the least. Very often it is not at all clear why your data have been accessed, at least from the first glance. Explanation messages are

vague and confusing, often being similar to ”data access by personal code”, making it difficult to understand the reason behind the data access, even if it was you who accessed it.

11.04.2025 20:44	HARIDUS- JA NOORTEAMET	Rahvastikuregister	ISIKUKOODI ALUSEL ISIKUANDMED

There is even an information sheet with recommendations for services implementing the Andmejälgija protocol, and it states that providing poor quality explanations for data access is a bad practice.[3]

2.3 Näiteid heast ja halvast selgitusest

Hea:

Isikuandmete töötlemise aeg	Tegevus	Isikuandmeid vastu võtnud osapool
13.01.2015 10:20:27	Retsepti vaatamine arsti poolt; retsepti number 1018472350	Arst Viktor Pihlakas
19.01.2018 10:58:23	Isiku päring kehtivate juhilubade kohta läbi riigiportaali eesti.ee	Jaan Kask 32405023456

Halb:

Isikuandmete töötlemise aeg	Tegevus	Isikuandmeid vastu võtnud osapool
13.01.2015 10:20:27	ISIKUKOODI ALUSEL ISIKUANDMED	ASUTUS X
19.01.2018 10:58:23	ISIKU LAIENDATUD INFO PÄRING ISIKUKOODI JÄRGI	SIHTASUTUS Y

Apparently, the advice is often ignored.

Furthermore, currently there is no law requiring institutions to implement the Andmejälgija protocol, meaning that its use is pretty much voluntary.

In the following chapters I will cover different state databases and their implementations of Andmejälgija, and whether they implement the protocol at all.

[3] https://www.ria.ee/sites/default/files/documents/2022-11/Soovitusi-Andmejalgija-rakendamiseks.pdf

3. Overview of state databases

3.1 Databases implementing Andmejälgija

- Digiregistratuur
- Elamislubade ja töölubade register
- Kinnistusraamat
- Kutseregister
- Maksukohustuslaste register
- Politsei taktikalise juhtimise andmekogu
- Põllumajandusloomade register
- Põllumajandustoetuste ja põllumassiivide register
- Rahvastikuregister
- Retseptikeskus
- Sotsiaalkaitse infosusteem
- Sotsiaalteenuste ja toetuste register
- Tööinspektsiooni tooelu infosusteem
- Töötuskindlustuse andmekogu

POLIS Information System, while in the list of the Andmejälgija view on Eesti.ee, actually doesn’t show any entries even if they should be there.

3.2 Databases providing other form of data access tracking

- E-toimik: provides it's own web-view for displaying requests made about you.

Kuupäev	Päringu liik	Päringu sooritaja	Eesmärk
15.04.2025	Minu kohta tehtud päringute vaatamine	Arkadi Statsenko, 50305170834 (E-toimik)	Isik enda kohta
18.03.2025	Päring kõikide karistuste kohta	Kaitseressursside Amet, 70007647 (Kaitseväekohustuslaste register)	Automaatkontroll
18.02.2025	Päring kõikide karistuste kohta	Kaitseressursside Amet, 70007647 (Kaitseväekohustuslaste register)	Automaatkontroll

3.3 Databases not providing any kind of data access tracking

- Schengen Information System
- Piirikontrolli andmekogu (PIKO)
- Infosusteem POLIS

4. Raw access logs vs Andmejälgija

5. Implementation

5.1 Description

One of the objectives of this thesis has been to develop a notifier for Andmejälgija data. I decided to create a mobile app for that.

5.2 Discussion of implementation choice

In addition to mobile app, there are other ways to develop such a solution. In this section I will discuss alternatives, as well as why I eventually decided to settle with the current approach

5.2.1 X-Road service

Andmejälgija specifications requires databases to implement an X-Road interface, as described in 2.1, so one option would be to create an X-Road service that would query Andmejälgija data over X-Road. The main advantage of this approach would be the freedom on how to notify the users of changes to the access logs. The service could support various channels of communication, including instant messenger bots, e-mail and others. The list of requirements in order to operate such a service is daunting, however.

- In order to join X-Road, legal entity is needed
- Permission has to be requested from every X-Road service you want to query data from
- As part of X-Road network, you need to operate a Security Server. It can be self-hosted anywhere for testing, but for production you need to have a Hardware Security Module (HSM), that costs around 10000€ (or >200€/month rent)

Satisfying this criteria is difficult and expensive. Additionally, even if I succeeded, that would make me a data controller and force people using my service to trust me with their data. That's whi prefer the standalone approach.

5.2.2 Standalone approach

This approach uses Eesti.ee session for accessing Andmejälgija data. Once you are logged in on Eesti.ee, certain internal API endpoints become available. Namely GET https://www.eesti.ee/andmejalgija/api/v1/usages endpoint can be used to query Andmejälgija data. The endpoint requires a parameter dataSystemCodes with which specific databases can be specified. For example GET request

/usages?dataSystemCodes=digiregistratuur&dataSystemCodes=rahvastikuregister would request access logs from Digiregistratuur and Rahvastikuregister.

The main advantage of this approach is the abscence of all disadvantages of the X-Road approach: there is no need for any kind of bureaucracy and the solution could be an open-source project, available for anybody to compile and use. There arises a problem, however. What about the notification part? Do I expect users to set everything up on their hardware, including relevant communication channels? That would narrow down the project's user base to technical people knowing how to self-host, and having a server.

That's why I thought that creating a mobile app would be the most optimal approach. The app would run the eesti.ee session and poll the API. This approach would combine the ease of setting up and use with solution remaining standalone, without a central server.

5.3 User guide

5.4 Software distribution

5.5 Known problems

6. Conclusion

7. Discussion

References

[1] Information System Authority (RIA). Andmejälgija / Data Tracker. Version 1.0.3. Apr. 15, 2025. https://github.com/e-gov/AJ.

[2] Information System Authority (RIA). Andmejälgija / Data Tracker protocol specification. Version 1.4.1. Apr. 15, 2025. https://github.com/e-gov/AJ/blob/master/doc/spetsifikatsioonid/Kasutusteabe_esitamise_protokoll.md.

License