

Anonymization Service for Texts and Semantic Finlex

Saara Packalén (MJF), Arttu Oksanen (Edita, Aalto), Minna Tamper (Aalto),
Aki Hietanen (MJF), Jouni Tuominen (HELDIG), Eero Hyvönen (HELDIG, Aalto)

HELDIG Digital Humanities Summit 2018

23.10.2018



MINISTRY OF JUSTICE

FINLAND

Content

- The project
 - The project organization
 - Aims & objectives
 - Motives
- Technical realization
 - The anonymization service
 - Semantic Finlex

The project organization

- 1.10.2018 - 30.9.2020
- Launched and coordinated by the Ministry of Justice Finland
- Cooperation with Helsinki Centre for Digital Humanities, Aalto University and Edita Publishing Oy
- Funded by the Ministry of Finance's special funds for projects that aim to increase productivity by means of robotics, analytics, artificial intelligence or other growing technologies (<https://vm.fi/robohaku2018-kevat>)

Aims & objectives

- To create and productize
 - Service for anonymization of documents/records that contain personal data
 - Service for content description and annotation of documents/records
- To implement the services
- In order to
 - Improve transparency of legal system
 - Enhance legal protection of citizens
 - Enable utilization of previous solutions in public administration and research

Motives

- Huge amounts of information and data are produced by public sector organizations – useful to other authorities, companies and citizens
 - Data protection regulation does not allow publishing data containing personal data
- The project focuses on court decisions
 - Effect on citizens yet only a small part of them are published
- Essential obstacles for increasing the amount of decisions anonymized are
 - Lack of resources
 - Lack of practical tools
- The documents/data accessible online today does not provide a representative picture of court decisions

The anonymization service

- Web application
 - Anonymization workflow:
 - Import a document that contains names into the application
 - Perform automatic anonymization (find named entity candidates)
 - Manually locate and fix errors
 - Export anonymized document
- Web service to be integrated with 3rd party software

The anonymization service

- User interface

Tuomiolauselma

Hovioikeuden tuomiota muutetaan.

Sähkötoimitusehtoa sovitellaan siten, että sen jälkeen, kun edellä selostetun, yhtiön puolesta kihlakunnanoikeudelle 15.1.1986 jätetyssä sanottuna päivänä päivätyssä kirjelmässä vaihtoehtona c esitetyn ehdotuksen edellyttämät muutostyöt sekä liittymis- ja sähkötoimitussopimus on tehty, **Sähköyhtiö Oy:n** on toimitettava tilaa varten sähköä ilmaissähköosuutena sovitut 5.000 kilowattituntia vuodessa ja sen ylittävän määrän yleistariffin hinnan mukaan perusmaksuineen.

Matti Meikäläinen vapautetaan korvaamasta yhtiön oikeudenkäyntikuluja kihlakunnanoikeudessa ja hovioikeudessa. Muutoin hovioikeuden tuomio jää pysyväksi.

Maija Meikäläisen vaatimus hänelle täällä aiheutuneiden oikeudenkäyntikulujen korvaamisesta hylätään.

Ratkaisuun osallistuneet: oikeusneuvokset Salervo, Ådahl, Lehtonen, Haarmann ja Huopaniemi

Henkilöt

Matti Meikäläinen Henkilö A

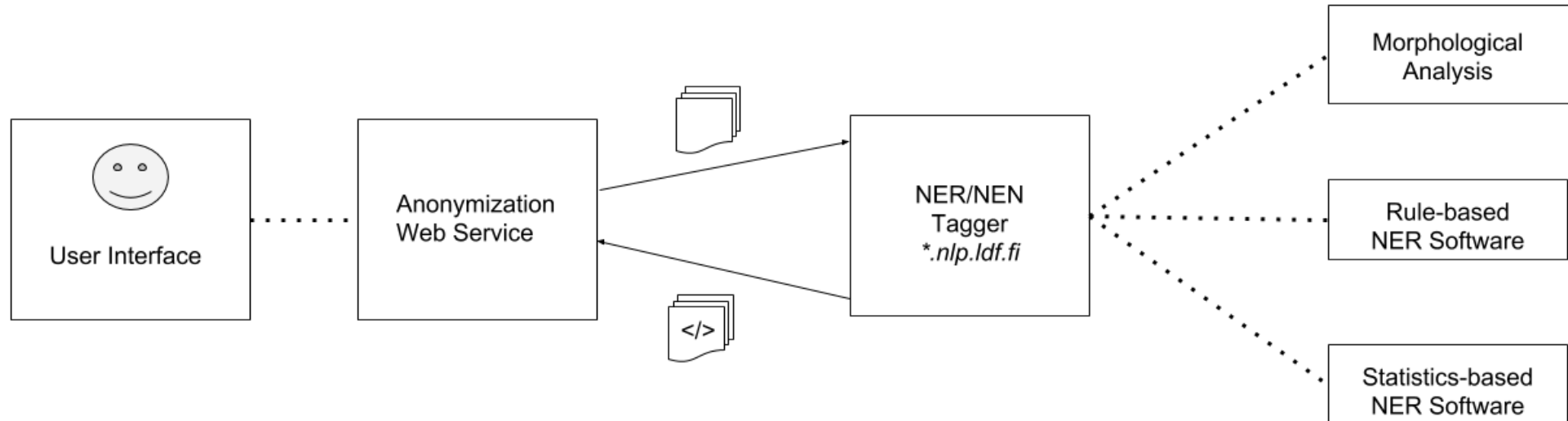
Maija Meikäläinen Henkilö B

Organisaatiot

Sähköyhtiö Oy Yritys A

The anonymization service

- Architecture



Annotating Semantic Finlex datasets

- Annotation service based on the same building blocks as the anonymization service
 - Can be used to enrich the documents published in the Semantic Finlex service by linking them to
 - Other legal documents
 - Legal vocabularies
 - Other datasets published as Linked Data
- Semantic search and recommender system based on the annotations
- Machine learning methods applied to find similar judgments

Thank you!

finlex@om.fi