
ESRA: AN END-TO-END SYSTEM FOR RE-IDENTIFICATION AND ANONYMIZATION OF SWISS COURT DECISIONS

A PREPRINT

Joel Niklaus

Research Center for Digital Sustainability
Institute of Computer Science
University of Bern
Bern, Switzerland
joel.niklaus@inf.unibe.ch

Matthias Stürmer

Research Center for Digital Sustainability
Institute of Computer Science
University of Bern
Bern, Switzerland
matthias.stuermer@inf.unibe.ch

September 29, 2020

ABSTRACT

The publication of judicial proceedings is a corner-stone of many modern democracies. It enables the court system to be audited and be made accountable by ensuring that justice is made in accordance to the laws in effect. Equally important is privacy, as a fundamental human right (Article 12 in the Declaration of Human Rights). Therefore, it is important that the parties (especially minors, victims or witnesses) involved in these court decisions be anonymized securely. Today, anonymization of court decisions in Switzerland is performed either manually or semi-automatically using primitive software. While much research has been conducted on anonymization for tabular data, the literature of anonymization for unstructured text documents is thin and virtually non-existent for court decisions.

In 2019, it has been shown that the manual anonymization is not secure enough. In 21 of 25 attempted Swiss federal court decisions related to pharmaceutical companies, pharmaceuticals and legal parties involved could be manually re-identified. This was achieved by linking the decisions with external databases using regular expressions. An automated re-identification system serves as an automated test for the safety of existing anonymizations and thus promotes the right to privacy. Manual anonymization is very expensive (recurring annual costs of over CHF 20M in Switzerland alone according to an estimation). Consequently, many Swiss courts only publish a fraction of their decisions. An automated anonymization system reduces these costs substantially, further leading to more capacity for publishing court decisions much more comprehensively.

For the re-identification system Topic Modeling with Latent Dirichlet Allocation is used to cluster an amount of over 500K Swiss court decisions into meaningful related categories. A comprehensive knowledge base with publicly available data (such as social media, newspapers, government documents, geographical information systems, business registers, online address books, obituary portal, web archive etc.) is constructed to serve as an information hub for re-identifications. For the actual re-identification a general-purpose language model is fine-tuned on the respective part of the knowledge base for each category of court decisions separately. The input to the model is the court decision to be re-identified and the output is a probability distribution over named entities constituting possible re-identifications.

For the anonymization system Named Entity Recognition (NER) is used to recognize the tokens that need to be anonymized. Since the focus lies on Swiss court decisions in German, a corpus for Swiss legal texts will be built for training the NER model. The recognized named entities are replaced by the category determined by the NER model and an identifier to preserve context.

This work is part of an ongoing research project conducted by an interdisciplinary research consortium. Both a legal analysis and the implementation of the proposed system design ESRA will be performed within the next three years.

This study introduces the system design of ESRA, an end-to-end system for re-identification and anonymization of Swiss court decisions. Firstly, the re-identification system tests the safety of existing anonymizations and thus promotes privacy. Secondly, the anonymization system substantially reduces the costs of manual anonymization of court decisions and thus introduces a more comprehensive publication practice.

Keywords Artificial Intelligence (AI) · Natural Language Processing (NLP) · Law · Courts · Legal Tech · Privacy · Named Entity Recognition (NER) · Topic Modeling