

Clustering of Analogies for Inter-Language Similarities

Software project - 7th presentation

Justine Diliberto, Cindy Pereira, Anna Nikiforovskaja

Université de Lorraine, IDMC

03.02.2022



UNIVERSITÉ
DE LORRAINE



Institut des
sciences du Digital
Management & Cognition

What was done before

- Defined 6 languages to study
 - Rule extraction of languages
 - Trained the model on the languages
-
- Rules comparison and conversion
 - Predicted the similarity
 - Built the website
 - Wrote the report

Language inflections

Language	#Inflections
English	5
Mezquital Otomi	18
Swedish	34
German	35
Finnish	91
Karelian	161

Table: Number of different inflections per language

Similarities

- Karelian and Finnish
 - 29 identical inflections
 - Other similar rules for nouns and adjectives
- German and Finnish/Karelian
 - Similarities between indicative, nominative and accusative
- Swedish and English
 - Some similar rules

JSON

JSON format: convenient for website

```
{ "VERBS":  
  { "INFINITIVE": "no change",  
    "INDICATIVE":  
      { "PRESENT":  
        { "Singular":  
          { "3rd person": "-s" } } },  
    "PARTICIPLE":  
      { "PRESENT": "-ing",  
        "PAST": "-ed" },  
    "PAST": "-ed" } }
```

Figure: Rules for English

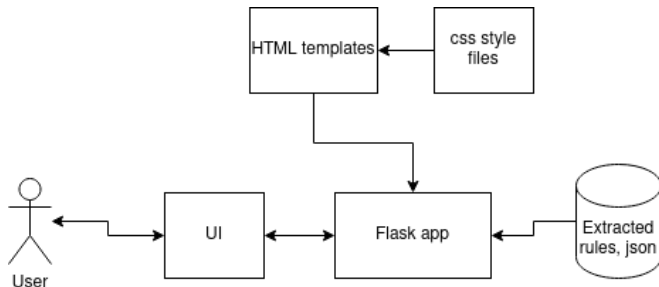
Why do we need a website?

- Make rules comparison more visible
- Make the findings easy to explore
- Present languages rules, their transfer scores, analogies where the model performs well

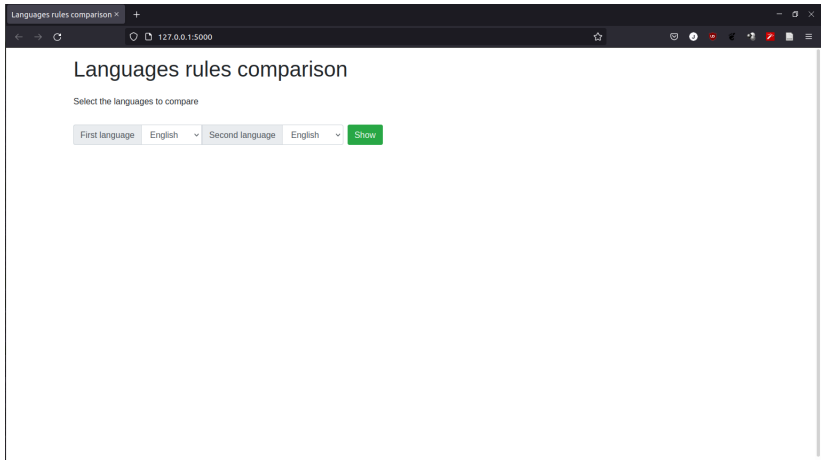
General architecture

Tools used:

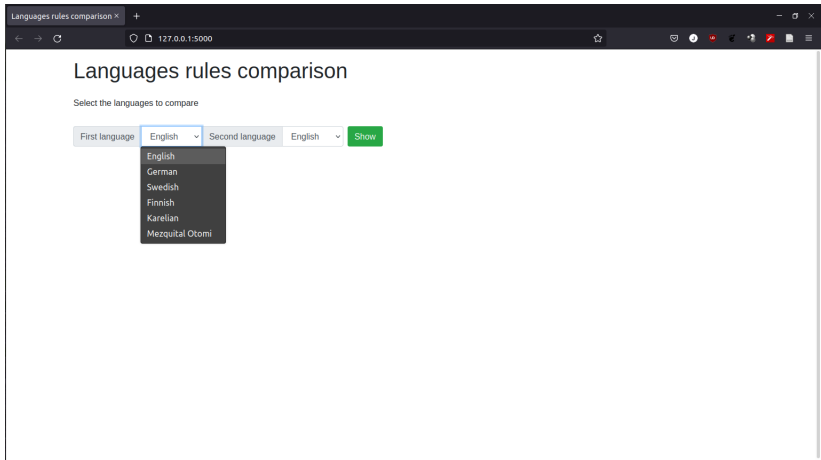
- Microframework Flask
- HTLM, CSS to make it beautiful



Website



Website



Website

Languages rules comparison

Select the languages to compare

First language English Second language English Show

German

VERBS

- INFINITIVE: -en
- INDICATIVE
 - PRESENT
 - Singular
 - 1st person: -e
 - 2nd person: -est
 - 3rd person: -et
 - Plural
 - 1st person: -en
 - 2nd person: -t
 - 3rd person: -en
 - PAST
 - Singular
 - 1st person: -te
 - 2nd person: -test
 - 3rd person: -te
 - Plural
 - 1st person: -ten
 - 2nd person: -ten/-en
 - 3rd person: -ten

Swedish

VERBS

- INFINITIVE
 - Active: no change
 - Passive: -s
- PARTICIPLE
 - PRESENT: -nde
 - PAST: -d | -en
- CONVERB
 - Active: -t
 - Passive: -ts
- IMPERATIVE: no change
- INDICATIVE
 - PRESENT
 - Active: -r
 - Passive: -s
 - PAST
 - Active: -de
 - Passive: -es

INFINITIVES

What is almost done:

- Report
- Website

To add:

- Comparison highlights
- Information on transferability
- Example analogies with the best performance

Thank you for your attention!