

The Centre for the Study of Manuscript Cultures (CSMC) brings together an international and cross-disciplinary research community

# TEI-based Interactive Critical Editions

Simon Schiff, Sylvia Melzer, Eva Wilden, Ralf Möller

Computer Sciences:

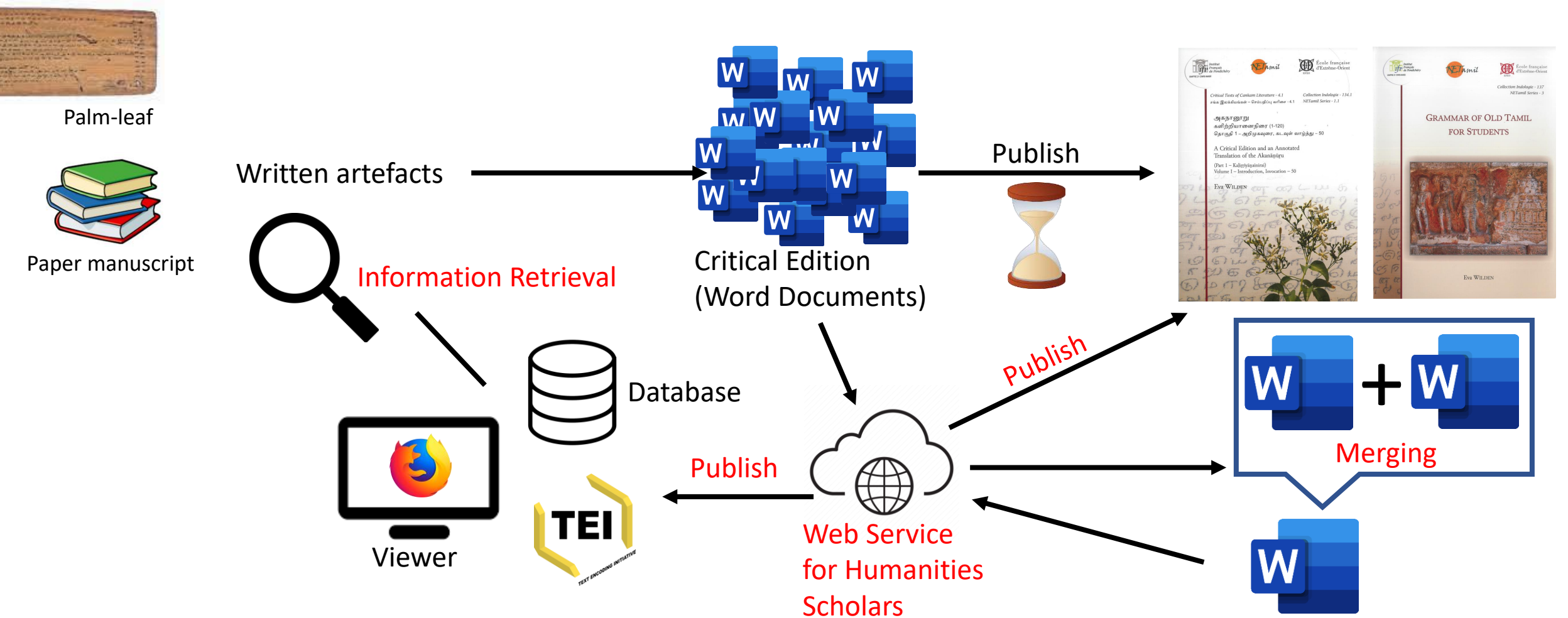
- Simon Schiff, Sylvia Melzer & Ralf Möller

Humanities Scholar:

- Eva Wilden

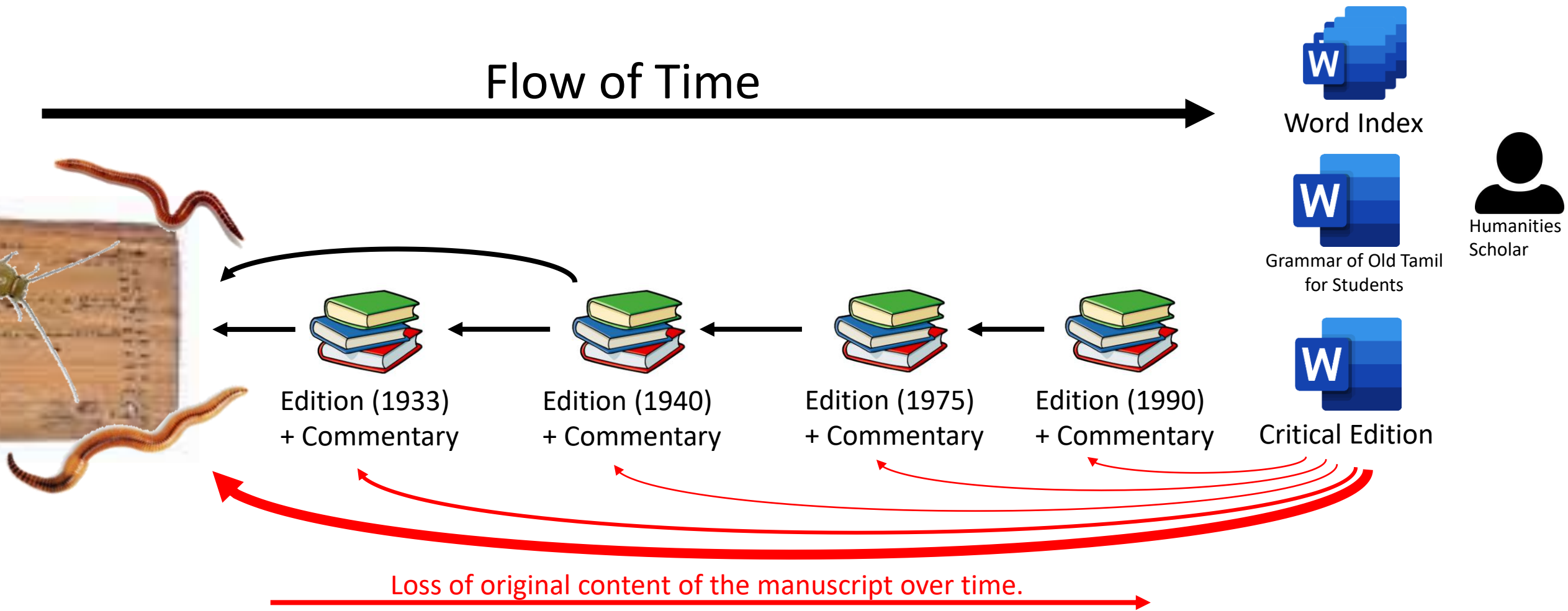
DAS 2022, La Rochelle, 23th May  
Institute of Information Systems  
University of Lübeck

# Motivation



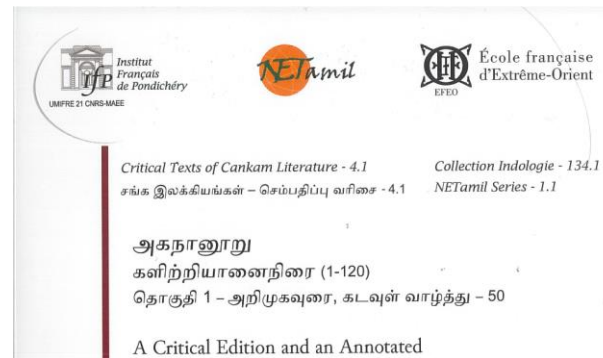
- We do not change preferred document formats and tools in use!
- Our solution is FAIR (**F**indable **A**ccessible **I**nteroperable **R**euse)

# Creation of a Critical Edition

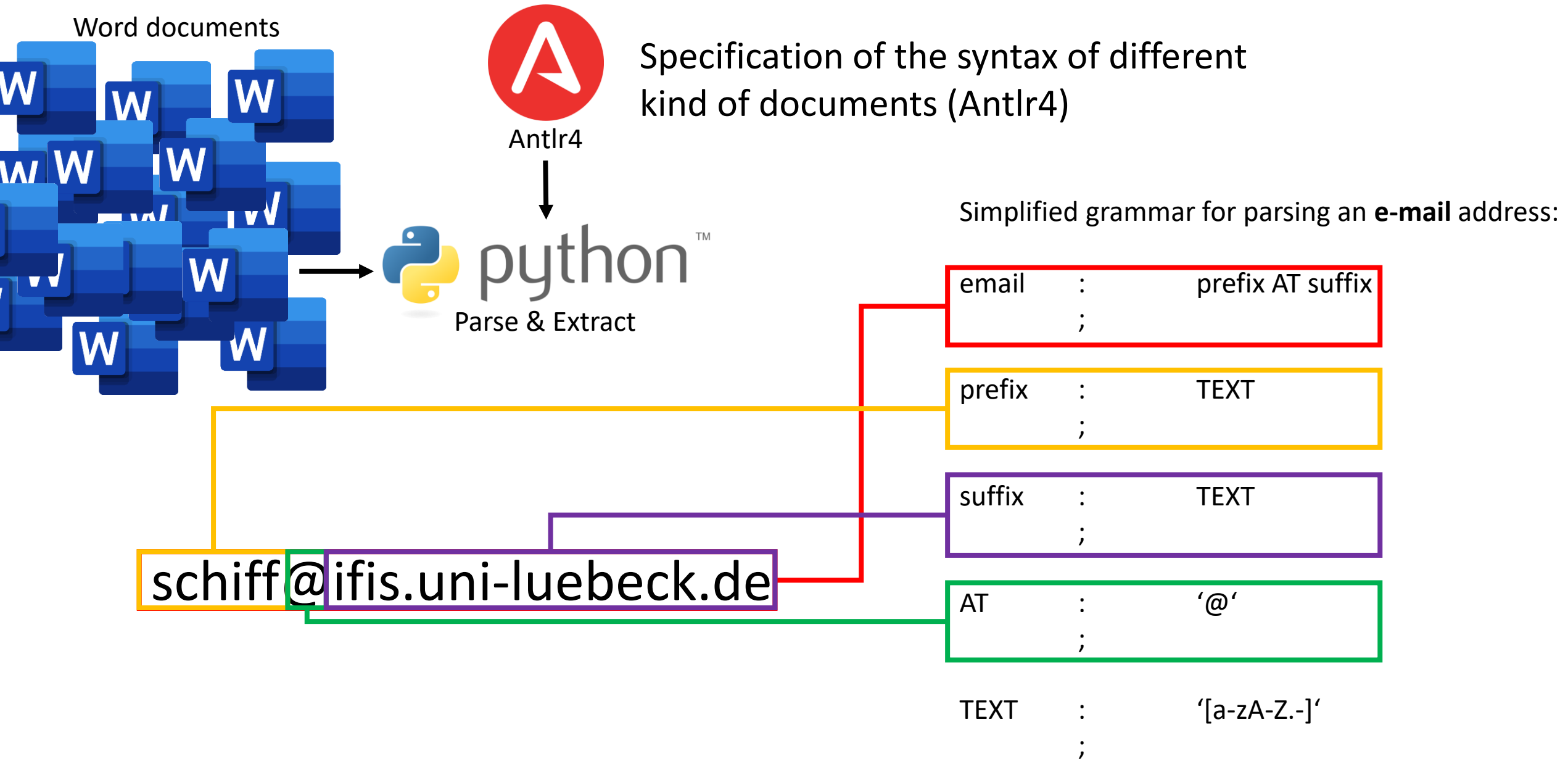


Eva Wilden, *A Critical Edition and an Annotated Translation of the Akanānūru (Part 1 - Kaḷirriyāṇainirai)*, 2018

Eva Wilden, *Grammar of Old Tamil for Students*, 2018



# Our Service (Step-by-step)



# 1. Vocabulary Index

# Grammar of the Vocabulary Index (Simplified)



<u>acā</u>	DEDR 39: n. languor 47.19 49.12° 162.16
<u>acāvu</u>	v. 5. to languish – v.r. 49.12*
<u>acāam</u>	pey. 170.9
<u>acunam</u>	TL (not DEDR, only these passages): p.n. a fabulous creatu 88.12
<u>acumpin</u>	obl. of <u>acumpu</u> DEDR 41 (Ti. 911: <u>pollānilam</u> ): n. well, mu
<u>acai</u>	DEDR 37: v. 4. to move – v.r. 40.5 77.13° 96.5v 96.6 102.4 302.2 340.22
<u>acaiyā</u>	neg. abs. 272.9
<u>acaiyinan</u>	p.a. m.sg. 102.13
<u>[acaivaral</u>	+ <u>varu</u> v.n. 96.6]
<u>acai</u>	v. 11. tr. to move; to shake
<u>acaii</u>	abs. 58.11° 398.2 398.17
<u>acaiiya</u>	pey. p.a. 58.11*
<u>acai</u>	DEDR 39: v. 4./5. to be weary ~ to rest – v.r. 29.1 99.5v
<u>acaii</u>	abs. 17.3 20.3 49.11 55.7 63.13 99.5° 120.11 190.1* 190.1
<u>acaiiya</u>	pey. p.a. 99.5* 190.1°
<u>acaiipu</u>	abs. 106.4*
<u>acaiya</u>	inf. 120.11v

```
workspace - Document_4/src/main/antlr4/de/uzl/ifu/antlr4/IndexParser.g4 - Eclipse IDE
Navigate Search Project Run Window Help
run.sh App.java App.java Reader.java Pandoc.java WordParser.g4 IndexParser.g4
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
Writable Insert 1:1:0
```

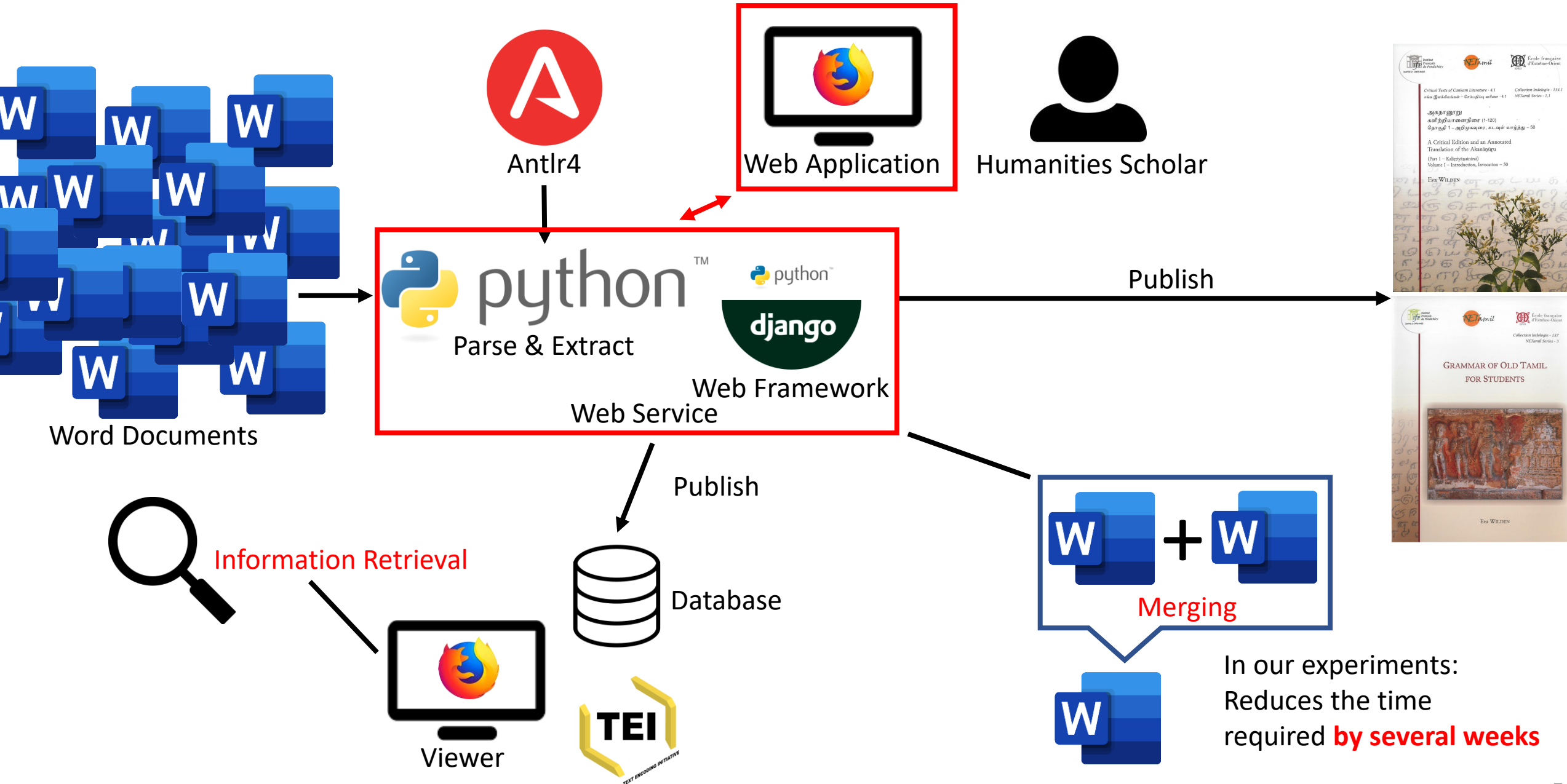
vocabulary\_index := (Title Entry\_Block)+  
Entry\_Block := Entry |  
Entry := Transcription Definition Reference\*

Title := Text  
Transcription := Text  
Definition := Text

Text := ('a'..'z' | 'A'..'Z' | 'Ä'..'ſ')+  
Reference := ('0'..'9')+ '!' ('0'..'9')+



# Our Service (Step-by-step)





# Web Application

NETamil

Upload Merge Export NETamil

Akanāṇṭṭu Index.docx ← **Word index**

AN-Maniṁṭaipavaḷam 121-200.docx ← **Critical edition**

Kolkatta (NLK)

BAY OF BENGAL

R. Narmada

R. Mahanadi

R. Godavari

R. Krishna

R. Tungabhadra

Tirupati (Sri Venkatesvara University)

Chennai (UFSL & GOML)

Pondicherry (IFP)

Chidambaram (AUL)

Tiruvaduturai (TAM)

R. Kaveri

Coimbatore (TSA)

Tanjavur (T. & MSSML)

Trivandrum (ORML)

N



# Web Application

Document Parser

Uploaded documents:

- [Akanānūru Index.docx](#)
- [AN-Manimītaipavalam 121-200.docx](#)

Select a file:

Browse... AN-Palai-Index.docx Index

Upload

Upload

Selected DOCX document to be uploaded

# 1. Vocabulary Index

# Grammar of the Vocabulary Index (Simplified)



<u>acā</u>	DEDR 39: n. languor 47.19 49.12° 162.16
<u>acāvu</u>	v. 5. to languish – v.r. 49.12*
<u>acāam</u>	pey. 170.9
<u>acunam</u>	TL (not DEDR, only these passages): p.n. a fabulous creat 88.12
<u>acumpin</u>	obl. of <u>acumpu</u> DEDR 41 (Ti. 911: <u>pollānilam</u> ): n. well, mu
<u>acai</u>	DEDR 37: v. 4. to move – v.r. 40.5 77.13° 96.5v 96.6 102.4 302.2 340.22
<u>acaiyā</u>	neg. abs. 272.9
<u>acaiyinan</u>	p.a. m.sg. 102.13
<u>[acaivaral</u>	+ <u>varu</u> v.n. 96.6]
<u>acai</u>	v. 11. tr. to move; to shake
<u>acaii</u>	abs. 58.11° 398.2 398.17
<u>acaiiya</u>	pey. p.a. 58.11*
<u>acai</u>	DEDR 39: v. 4./5. to be weary ~ to rest – v.r. 29.1 99.5v
<u>acaii</u>	abs. 17.3 20.3 49.11 55.7 63.13 99.5° 120.11 190.1* 190.
<u>acaiiya</u>	pey. p.a. 99.5* 190.1°
<u>acaiipu</u>	abs. 106.4*
<u>acaiya</u>	inf. 120.11v

```
workspace - Document_4/src/main/antlr4/de/uzl/ifu/antlr4/IndexParser.g4 - Eclipse IDE
Navigate Search Project Run Window Help
run.sh App.java App.java Reader.java Pandoc.java WordParser.g4 IndexParser.g4
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
Writable Insert 1:1:0
```

vocabulary\_index := (Title Entry\_Block)+  
Entry\_Block := Entry |  
Entry := Transcription Definition Reference\*

Title := Text  
Transcription := Text  
Definition := Text

Text := ('a'..'z' | 'A'..'Z' | 'Ä'..'Ŧ')+  
Reference := ('0'..'9')+ '!' ('0'..'9')+

# Web Application

The screenshot displays a web application interface. At the top, there is a browser window titled "Dictionary Viewer" with navigation icons (back, forward, refresh, home) and a search bar. Below the browser window is a dark navigation bar with the following menu items: "Upload", "Merge", "Export", and "NETamil".

On the left side, there is a list of uploaded files:

- Akanāṅṅūru Index.docx
- AN-Manimīṭaipavaḷam 121-200.docx
- AN-Palai-Index.docx** (highlighted with a red box and a red arrow pointing to it from the text "Uploaded vocabulary index")

The main area of the application shows a map of South India. The map includes the following labels:

- R. Narmada
- R. Mahanadi
- R. Godavari
- R. Krishna
- R. Tungabhadra
- R. Kaveri
- Tirupati (Sri Venkateswara University)
- Chennai (U.F.S.L. & G.O.M.L.)
- Pondicherry (I.F.P.)
- Chidambaram (A.U.L.)
- Tiruvaduturai (T.A.M.)
- Tanjavar (T.L. & M.S.S.M.L.)
- Coimbatore (T.S.A.)
- Trivandrum (O.R.M.L.)
- Kolkatta (N.L.K.)
- BAY OF BENGAL
- N (North arrow)

# Web Application

Dictionary Viewer

Back

## Dictionary

### Table of contents

- [Akam-Index](#)
- [Abbreviations](#)
- [1. Glossary](#)
- **a** ← Chapter a
- [ā](#)
- [i/ī](#)
- [u/ū](#)
- [e/ē/ai](#)
- [o/ō/au](#)
- [k](#)
- [c](#)
- [ñ](#)
- [t](#)
- [n](#)
- [p](#)
- [m](#)
- [y](#)
- [v](#)

# Web Application

Dictionary Viewer

← → ↻ 🏠

**a** ← Chapter a

\*alar v. 4. to gossip

**alar** DEDR 245: n. gossip [36.12](#) [45.8](#) [70.6](#) **[90.4](#)** [96.18](#) [98.25](#) [110.1](#) [113.8](#) [116.12](#) [190.4](#) [216.6](#) [218.18](#) [220.1](#) [246.7](#) [282.12](#) [296.14](#) [320.7](#) [368.19](#) [400.2](#)

alarin obl. [69.19v](#)

alarka sub. [4/1.16](#)

alar occurs at poem 90 in line 4

-a adj. suff. [34.8](#) [54.8°](#) [54.9](#) [56.5](#) [56.7](#) [62.9](#) [64.11](#) [94.4](#) [98.3](#) [104.7](#) [104.9](#) [104.10](#) [108.18v](#) [118.1v](#) [118.2](#) [144.14°](#) [154.2](#) [158.1](#) [158.13](#) [182.6](#) [184.6](#) [184.8](#) [208.2](#) [224.4](#) [228.9](#) [228.10](#) [244.2](#) [256.7°](#) [284.4](#) [296.1](#) [304.15v](#) [332.1](#) [358.5](#) [364.5](#) [364.7](#) [378.6](#) [378.20](#)

-akam loc. suff. [19.12\\*](#)

-am adj. suff. [1.17](#) [5.8v](#) [6.1](#) [7.6](#) [10.4](#) [12.6](#) [14.5](#) [15.3](#) [18.15](#) [24.4v](#) [26.4](#) [35.16](#) [37.15](#) [53.7](#) [54.8](#) [65.4](#) [70.10](#) [73.11](#) [73.14](#) [76.9\\*](#) [80.13](#) [90.3](#) [96.8](#) [97.1](#) [97.1](#) [100.13](#) [100.17](#) [100.18](#) [113.6](#) [115.4](#) [136.4R](#) [138.1](#) [138.6](#) [152.6](#) [160.8](#) [170.9](#) [200.2](#) [210.9](#) [210.12](#) [226.6](#) [240.14](#) [250.11](#) [270.4](#) [270.7](#) [284.2](#) [284.13](#) [294.7](#) [300.4](#) [302.8](#) [312.5](#) [312.8](#) [320.5](#) [328.14](#) [330.6](#) [332.9](#) [340.2](#) [346.6](#) [352.11](#) [360.2](#) [362.15](#) [368.8](#) [384.5](#) [388.7v](#) [400.19](#)

a san. [30.1](#) [60.2](#) [70.3](#) [220.16](#) [250.11](#) [290.4](#) [390.10](#)

am DEDR 162: (amai beauty, fitness) adj. pretty [3.16](#) [5.8](#) [6.2](#) [6.17](#) [9.10](#) [9.25](#) [11.2](#) [11.5](#) [15.15](#) [16.3](#) [21.3](#) [21.10](#) [21.17](#) [27.9](#) [29.7v](#) [31.6](#) [38.3](#) [38.6](#) [41.15](#) [44.14](#) [45.1](#) [46.5](#) [48.8v](#) [54.11](#) [59.7°](#) [60.5](#) [69.7](#) [71.8](#) [75.18](#) [76.12](#) [78.23\\*](#) [79.13](#) [86.11](#) [87.13°](#) [90.6](#) [96.4](#) [96.12](#) [97.2](#) [97.22](#) [99.6](#) [99.7](#) [103.3](#) [107.16](#) [107.22\\*](#) [104.12](#) [113.11](#) [113.24](#) [114.14](#) [117.2](#) [124.14](#) [126.9](#) [132.6](#) [134.12](#) [142.1](#) [154.15](#) [184.12](#) [188.12](#) [198.10](#) [198.13](#) [204.6](#) [206.8](#) [208.10](#) [226.3](#) [234.12](#) [236.7v](#) [236.11](#) [242.21°](#) [262.18](#) [266.2](#) [270.10](#) [278.8](#) [280.5v](#) [284.4](#) [288.2](#) [288.17](#) [290.7](#) [302.2](#) [320.9\\*](#) [336.7](#) [342.13](#) [368.14](#) [376.14](#) [378.3](#) [388.2](#)

aktai (not Index, not TL) p.n. of a person [76.3\\*](#) [96.12v](#) [113.4v](#)

akattu obl. [5.5](#) [8.12](#) [21.16](#) [66.15](#) [107.3](#) [124.15](#) [144.18v](#) [156.6\\*](#) [302.9v](#) [306.10v](#) [386.6](#)

akan n. inside [144.6\\*](#)

# Web Application

Poem Viewer

Back

## Tamil Poems

### kiLavi

Poem 90: Occurrence of alar at poem 90 in line 4

தோழி செவிலித்தாய்க் கறத்தொடு நின்றது.  
How the confidante stood firm in [her] duty to the foster mother.

Tamil Poem	Transliteration	Word-by-word translation into English
1 திரையு[ழ]ந் தசைஇ நிரைவளை யாய	1 tirai ~u[ant]* acaii niraī vaḷai ~āyam	1 wave borne rested row bangle attendant
2 முப்பின் குப்பை யேறி யெற்பட	2 uppiṅ kuppai ~ēri ~el paṭa	2 saltiṅ heap mounted light happen(inf.)
3 வருதிமிசெ லெண்ணுந் துறைவனோ டூரே	3 varu timil eṇṇum turai vaṇōṭ* ūrē	3 come- boat counting- ghat-he-with villageē
4 யொருதன் கொடுமையி னலர்பா டும்மே	4 ~oru taṅ koṭumaiyiṅ alar pāṭumē	4 one self- crueltyṅ gossip singing-ē
5 யலமரன் மழைக்க ணமர்ந்து நோக்கா 5	5 ~alamaral maḷai+ kaṅ amarntu nōkkā  5	5 whirling rain eye abided- look-not-she 5
6 ளாலையல் வாழிவேண் டன்னை யுயர் மிசைப்	6 alaiyal vāḷi vēṇṭ* aṅṅai ~uyar micai+	6 don't-slap live(ipt.) be-necessary(ipt.) mother be-high- elevat
7 பொதும்பிற் புன்னைச் சினைசேர் பிருந்த	7 potumpiṅ puṅṅai+ ciṅai cērp* irunta	7 groveiṅ Puṅṅai-tree branch joined been
8 வம்ப நாரை யிரிய வொருநாட்	8 vampa nārai ~iriya ~oru nāḷ	8 new wader retreat(inf.) one day
9 பொங்குவர லூதையொடு புணரி யலைப்பவ	9 poṅku varal ūtaiyoṭu puṅari ~alaippa~um	9 foam- coming cold-wind-with wave/ocean slap(inf.)um
10 முழைக்கடல் வழங்கலு வெய்ய னகன்றலை 10	10 ulai+ katal valaṅkal-um veyyan akal-talai 10	10 proximity sea roaminquum hot-he widen- area 10

# Web Application

Poem Viewer

Upload Merge Export NETamil

- Akanāṇṇūru Index.docx
- AN-Manimittapavalam 121-200.docx
- AN-Palai-Index.docx

**Merging of two uploaded word indices**

Kolkatta (NLK)

R. Narmada

R. Mahanadi

BAY OF BENGAL

R. Godavari

R. Krishna

R. Tungabhadra

Tirupati (Sri Venkateswara University)

Chennai (U/SL & GOML)

Pondicherry (IFP)

Chidambaram (AUL)

Tiruvaduturai (TAM)

R. Kaveri

Coimbatore (TSA)

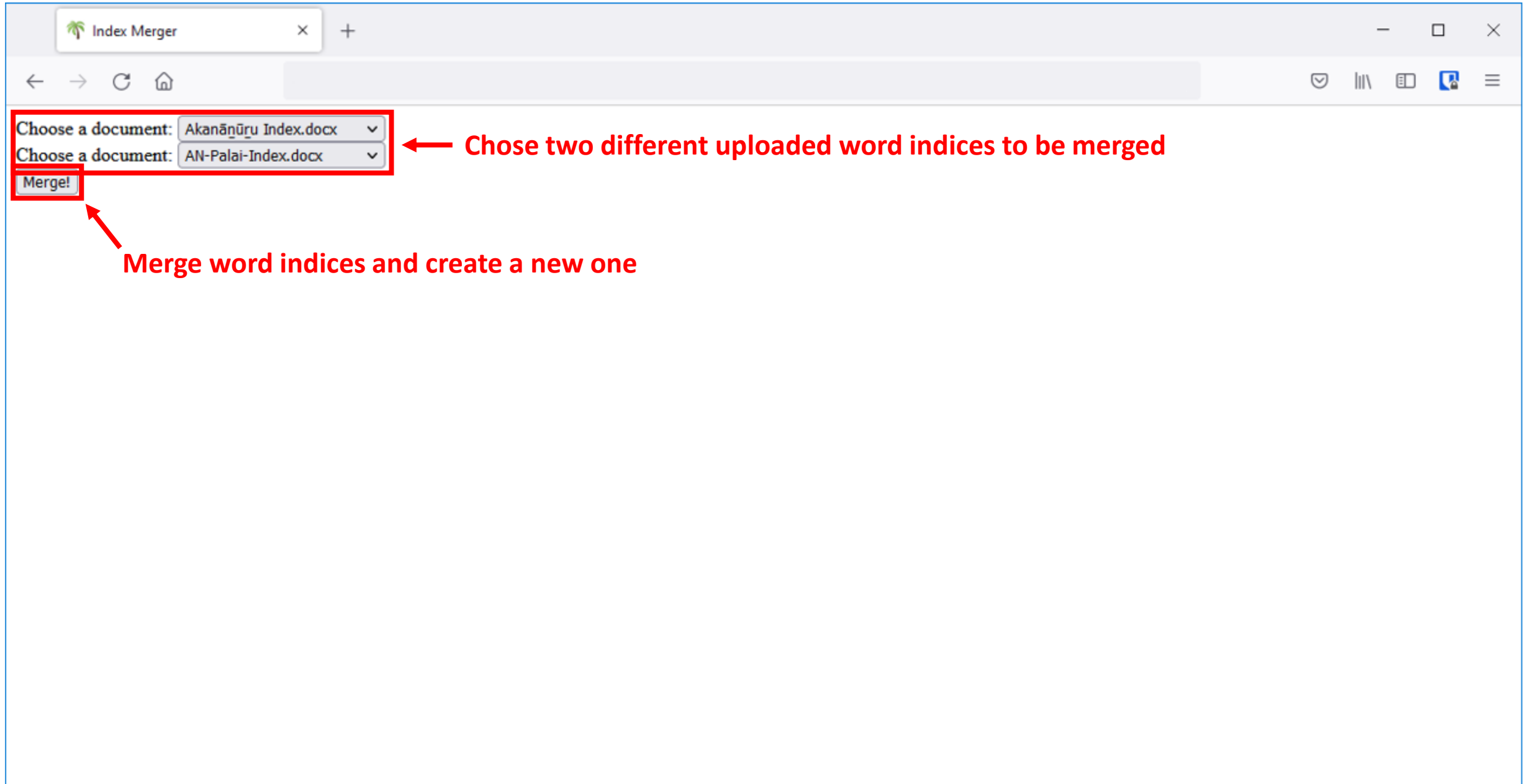
Tiruvannamalai (T & MSSML)

Trivandrum (ORML)

N



# Web Application



The screenshot shows a web browser window with a single tab titled "Index Merger". The browser's address bar is empty. Below the address bar, there are two dropdown menus for selecting documents. The first dropdown is labeled "Choose a document:" and has "Akanānūru Index.docx" selected. The second dropdown is also labeled "Choose a document:" and has "AN-Palai-Index.docx" selected. Below these dropdowns is a button labeled "Merge!". Red arrows and text annotations highlight these elements: one arrow points from the text "Chose two different uploaded word indices to be merged" to the two dropdown menus, and another arrow points from the text "Merge word indices and create a new one" to the "Merge!" button.

Choose a document: Akanānūru Index.docx

Choose a document: AN-Palai-Index.docx

Merge!

Chose two different uploaded word indices to be merged

Merge word indices and create a new one

# Web Application

NETamil

Upload Merge Export NETamil

- Akanāṅṅuru Index.docx
- AN-Maniṁṁaipavaḷam 121-200.docx
- AN-Palai-Index.docx
- Akanāṅṅuru Index.docx - AN-Palai-Index.docx**

Merged word indices

Kolkatta (NLK)

R. Narmada

R. Mahanadi

BAY OF BENGAL

R. Godavari

R. Krishna

R. Tungabhadra

Tirupati (Sri Venkatesvara University)

Chennai (UFSL & GOML)

Pondicherry (IFP)

Chidambaram (AUL)

Tiruvaduturai (TAM)

R. Kaveri

Coimbatore (TSA)

Tanjavur (T. & MSSML)

Trivandrum (ORML)

N

# Databasing on Demand with an Interface for IR

Enter a query as a sequence of words

The screenshot shows a search interface with three main panels. The left panel, titled "Dictionaries", contains a search bar with "acai" entered and a "Close x" button. Below the search bar is a list of dictionaries and orthographic groups, each with a count in a circle. The middle panel shows a list of search results for "acai", with the top result highlighted in blue. The right panel shows the details for the selected result, "acai (Akam-Index)", including two entries from the DEDR database.

**Search everything**

acai

**Dictionary**

Akam-Index	17
AN-Palai-Index-001-120	23
AN-Palai-Index-121-300	27

**Orth group**

a	42
n	17
p	6
v	2

**Results (n = 67)**

- ācai
- ācai
- acai
- acai**
- acai
- acaii
- acaii
- acaii
- acaiiya
- acaiiya
- acaiiya
- acaintanar
- acaintanar
- acaintu
- acaintu

**acai (Akam-Index)**

[DEDR 37: v. 4. to move v.r. 40.5 77.13° 96.5v 96.6 102.4 162.8 272.9 298.6 302.2 340.22](#)

[DEDR 43: v. 11. to tie – v.r. 54.7](#)

Filter results

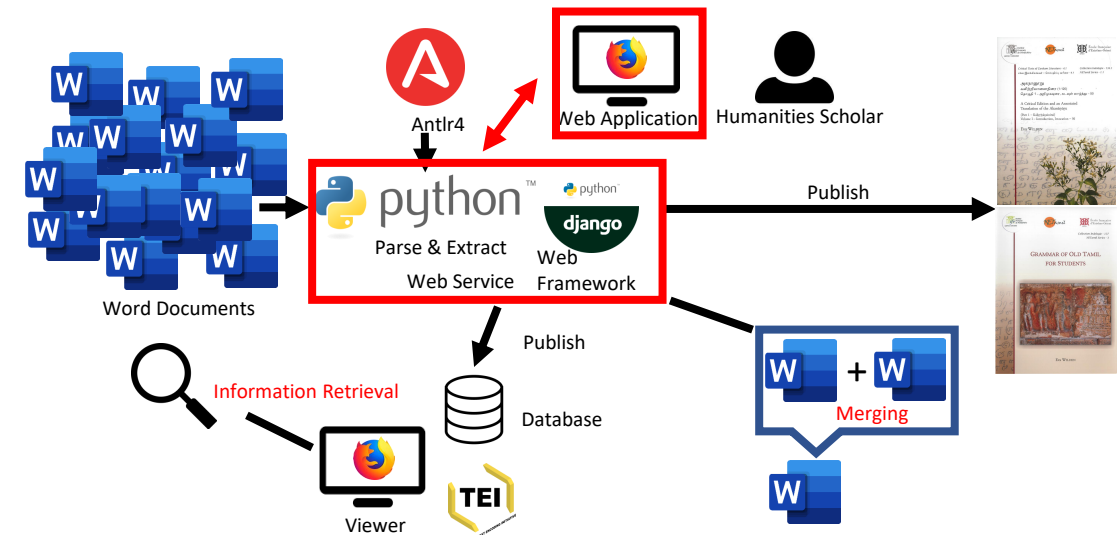
# Conclusion and Future Work

Our solution:

- Extension and not a replacement
  - Keep preferred tools and document formats
- Time needed for specific tasks reduced by several weeks
- Extensions are more likely to be accepted
- FAIR principles met (For instance Reusability<sup>1</sup>)

Future Work:

- We aim to keep an IT-expert away from the process



<sup>1</sup>Bender, Magnus, et al. "Identifying and Translating Subjective Content Descriptions Among Texts." *International Journal of Semantic Computing* 15.04 (2021): 461-485.