

WG3 2nd Meeting Minutes - 2023-09-08

The second WG meeting was mainly devoted to the two tasks/deliverables outlined in the MoU, documentation of multilingual language tools and multilingual evaluation campaigns. In addition, we discussed ways of working in the WG.

Documentation of Multilingual Language Technology Tools

The morning session on documentation of multilingual tools was led by the task leaders A. Seza Doğruöz, Kristina Kocijan and Teresa Lynn and featured the following presentations:

- Introduction to European Language Grid (ELG) (Teresa) ([slides](#))
- Survey on preferences concerning documentation platforms (Kristina) ([slides](#))
- Overview of ELG (invited presentation by Maria Giagkou) ([slides](#))
- Exploration of data in ELG (Kristina) ([slides](#))
- Other platforms and future directions (Seza) ([slides](#))

There was a lively discussion covering many topics but focusing on the choice of documentation platform and the scope of the documentation task. Here is a summary of the main points and some of the issues raised in relation to them:

- It makes sense to use on an existing platform for several reasons.
 - We don't have the resources to develop our own platform (or maintain it after the end of the action).
 - We don't need to start from scratch because many tools will already be documented.
- Of the available platforms, ELG seems to be the best fit for what we want to do, although all platforms have pros and cons. For example:
 - ELG harvests information from other platforms and sources, but it does not provide direct access to tools and resources in the same way that CLARIN does.
 - It is not clear how ELG will be maintained in the future.
- We need to define (and possibly narrow) the scope of our documentation task along several dimensions, including the following:
 - What is meant by multilinguality?
 - Do we include only strictly multilingual tools, or do we want to document monolingual tools for many languages?
 - What languages do we include?
 - ELG focuses on languages in Europe, but can easily be extended.
 - In practice, we may be limited by the expertise available within the action.
 - How do we deal with dialects and language variants?
 - How do we take users into account?
 - Do we restrict ourselves to tools or do we also include (other) resources?
 - Should we try to integrate typology resources? How?

Multilingual Evaluation Campaigns

The afternoon session was organized by the WG3 leaders and featured short invited presentations from WG members who had submitted ideas for evaluation campaigns:

- D. Zeman: UD Shared Task Inside Out ([abstract](#))
- A. Wróblewska: NLPRe Benchmarking Platform ([abstract](#))
- O. Goldman, L. Weissweiler and R. Tsarfaty: Morpho-Syntactic Analyso-Parsing ([abstract](#), [slides](#))
- T. Samardzic: Subword Pooling

The proposals were discussed in breakout groups, followed by a general plenary discussion where several other ideas for evaluation campaigns were proposed as well. The WG leaders will call a meeting with the proposers to discuss next steps. Anyone else who would like to be involved in this task are welcome to contact the WG leaders. The goal is to organize at least one evaluation campaign during 2024 and a different one during 2025.

Ways of Working

A short session was devoted to ...

From:
<https://unidive.lisn.upsaclay.fr/> - **Universality, diversity and idiosyncrasy
in language technology
CA21167 COST Action**

Permanent link:
https://unidive.lisn.upsaclay.fr/doku.php?id=wg3:wg3_meeting_2023-09-08&rev=1695237135

Last update: **2023/09/20 21:12**

