



# CLARIN-PLUS

## Work package 5.2

### Bringing in New Users

*Martin Wynne*

*CLARIN ERIC Director for User Involvement &  
University of Oxford*



The potential user community is large, fragmented and diffuse, and hard to reach by methods that try to address them all at once.

In order to allow for a focused approach in bringing in new users **four topics** will be identified that will be used to **demonstrate** how the application of language and speech technology tools and services on digital language material can **advance humanities and social sciences research** in fields other than linguistics.

# Workshops: procedure



For each of the four topics the following steps will be taken:

- Identify, through **desk research**, a number of successful applications of language and speech technology in social sciences and humanities, as described in journals and conference proceedings
- Organise a **3 day workshop** with 1 day of presentations, 1 day of tutorials and 1 day of practical exercises, preferably in the margin of an international SSH conference. All presentations will be recorded on video.
- Ask the **authors** whether they are prepared to attend the workshop where they present their work, teach a class on how it works and prepare **practical exercises** directly related to their work
- After the workshop collect and **publish all relevant material** on the website for on-line use by others. The structure should be such that it can be ported to other languages and **re-used at national events**, in order to avoid language barriers.



Planned topics (subject to change, based on user interest, overlap with other events):

- Work on digital collections of **newspapers** (as a dataset highly relevant to many SSH communities, and there are typically few or no IPR problems with older material)
- **Oral history archive materials** (a topic which combines language and speech technology, and is hard to do without digital tools)
- **Parliamentary records** (from national assemblies and the European Parliament, and which are normally freely accessible and available for re-use)
- Analysis of **social media**

Approximate timing: months 6, 12, 18 and 24, but co-location with other events may require slightly different scheduling once the starting date of the project has been established.



Candidates have been identified for the following workshops (and named in the proposal):

- Ineke Schuurman (University of Leuven): Newspaper collections
- Petya Osenova (Bulg Acad of Sciences, Sofia): Parliamentary records
- Jolanta Kovalevskaitė (Vytautas Magnus University, Kaunas): Social media

And a further offer is in preparation:

- John Coleman (University of Oxford): Oral history archive collections

# CLARIN-PLUS User Workshops

## CLARIN+ User Workshops

- 📍 Kaunas
- 📍 Leuven
- 📍 Oxford
- 📍 Sofia

Proposed venues for workshops







**Subtask T5.2.2** – *Building a sustainable registry of courses and other educational material related to the use of digital language resources and tools in HSS research.*

For scholars in the humanities and social sciences who want to educate themselves as well as for educators of such scholars it is hard to know which courses and educational material exist. An initial collection has been made over the last two years, but the material needs to be systematized and published in such a way that it is easily accessible and maintainable. CLARIN ERIC and DARIAH ERIC have already expressed their intention to initiate collaboration in the field of education, and for this specific activity collaboration will be sought with DARIAH ERIC, where similar initiatives exist, serving largely the same community, albeit with a slightly different focus.

*Timetable:* months 6-18

*Deliverable:* Operational course and education material registry M18

*Resource:* 1 person month