CoLABnet

Competencies Network for Liberal Arts on the Balkans
CoLABnet

- Multi-level and interdisciplinary cooperation
- Teaching, Training, Research
- Linguistics, History, Sociology, Computer Science, Mathematics, Literature
The CoLABnet concept

- Making sense of Erasmus+ Key Action 1
- Maintaining eligibility for Erasmus+
- Integrating teaching and research
- Adding third-party funding
CoLABnet incentives

- Paving the way for
  - Bigger projects
  - Joint programmes
  - Better collaboration
The CoLABnet Gear

CoLABnet

CoLABbase

ERASMUS+

CoLABprojects
Erasmus+

- International Credit Mobility (KA107)
- Funding for Western Balkan 2,7 – 4 mio € per Call for Germany
- Total funding for CoLABnet 2015-2020: 935,912 €
Network partners
Mobilities within CoLABnet

![Mobilities Graph](image)

- **Student**
- **Staff**
Participating staff

Staff members involved in CoLABnet exchange: Multiple and First involvement
Perspective

- Further funding via KA107
- Capacity building (KA2)
- Joint study programme?
CoLABprojects

- Interdisciplinary project lines
- Filter for Erasmus+ mobilities
- Integration of research and teaching
CoLABprojects

- Each project line defines
  - Topics
  - Academic fields involved
  - Mobility formats
  - Output

- Open to new suggestions and grassroots initiatives
Humboldt Balkan Cosmos

- Interdisciplinary research on the ethnical, political, religious, linguistic and historical diversity of the Balkans
- Providing best-practice examples for the combination of research and teaching
- Implementing the university`s third mission: internationalising university education and integrating civil society
<table>
<thead>
<tr>
<th>Year</th>
<th>Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>Field trip with students from Berlin and Belgrade to Muslim minorities in Macedonia and Albania</td>
</tr>
<tr>
<td>2018</td>
<td>Two workshops on discoursive integration and segregation of Muslims and Jews in BiH and Montenegro</td>
</tr>
<tr>
<td>2019</td>
<td>Field trip with students Sandžak</td>
</tr>
</tbody>
</table>
Output

- Website
  www.humboldt-balkan-cosmos.net
- Check the exhibition on the South-Slavic floor!
Internationalization

- Internationalization of curricula, teaching methods and mentoring
- Empowerment for international third-party funding abilities
- Institutional collaboration within the region
- Internationalization of (young) scholars
Activities

- Workshops on EU-fundraising
  - Tirana, Montenegro, Sarajevo, Kragujevac, Niš, Prishtina
- Career service for young scholars
Corpus Linguistics

- Focus on the **creation of spoken corpora**
- Incentive:
  - Use the multi-faceted aspects of the creation of spoken corpora for teaching and research
- Aims
  - Knowledge transfer for corpus creation
  - Skills transfer for corpus tools
  - Creation of research resources
  - Joint projects
  - Join forces
Corpus Ling as CoLABproject

- **Topics:**
  - Tool creation, Corpus design, Material collection, Corpus compilation, Analysis

- **Dimensions:**
  - Teaching – Research – Outcome

- **Academic fields involved:**
  - Linguistics, Sociology, Mathematics, Computer Science, Anthropology, Digital Humanities
Tool creation

- Linguistics, Mathematics, Informatics
- Teaching: Scripting, Statistics
- Research: Tools
- Output: Tools for ABCMS
  *e.g.* Tagset for Albanian, GATE plugins, FrameNet
Corpus design

- Linguistics, Sociology
- Teaching: Sampling, Demography,
- Research: Representativity, Architecture, Language use
- Output: Corpus Templates, Corpus Designs e.g. Demographically balanced spoken corpus for Bosnia-Herzegovina, Intertextual Corpus of Serbian Law Language in the 19th century
Material sampling

- Linguistics, Anthropology, Sociology
- Teaching: Fieldwork, Collection, Processing
- Research: Ethnolinguistic research
- Output: Material
e.g. Language of heritage speakers in Berlin
Corpus compilation

- Linguistics, Computer Science, Mathematics
- Teaching: Databases, Scripting, File formats
- Research: Processing
- Output: Corpus
e.g. CRONUS – Corpus for the Research on Narratives and their Use in Speech, Spoken Corpus of Serbian in Hungary
Corpus analysis

- Linguistics, Sociology, Mathematics, Computer Science, Anthropology
- Teaching: Statistics, Tools
- Research: Sociolinguistic Corpus Analysis
- Output: Papers, Theses
Mobility formats

- Trainings for students and staff (train the trainer) on corpus creation and corpus linguistics by HU scholars
- Workshops on corpus linguistics didactics, development of similar or joint courses and harmonized curricula in Berlin
- Workshops on corpus design, corpus creation in Berlin
- Individual research/Study stay/Training
Activities

- Basic trainings in corpus creation (including the usage of the tools Exmaralda and Gate)
  - Sarajevo, Kragujevac, Prishtina
- Individual two-week trainings at HU
- PhD-projects with research stays at HU
  - Jelena Pavlovic (Kragujevac): Law discourse corpus
  - Dalian Zogaj (Prishtina): MULTEXTeast tagset for Albanian
  - Andrea Hrstić (Zagreb): Croatian FrameNet
- Staff week / Workshop
The CoLABnet Gear
CoLABbase

- Platform for the exchange of
  - Project outputs
  - Teaching material
  - Project proposals
  - Literature
  - Etc.

- At the moment only at HUbox ->
  - Presentations
  - Libraries (General Corpus Linguistics, Spoken Corpora)
  - Training materials
  - Templates in the corporate design
Website

- https://www.slawistik.hu-berlin.de/de/fachgebiete/suedslawsw/colabnet
This workshop

- **Aims**
  - Bring together (future) agents in corpus linguistics
  - Provide a basic common level of knowledge about spoken corpora
  - Focus on design and creation
  - Develop future corpus projects
  - Gain an overview over existing experience and knowledge and define fields where training is necessary
Your experience

Your Previous Experience in Corpus Linguistics

- Existing Corpora
- Recording Language
- Editing Recordings
- Transcription
- Tagging
- Tagsets
- Annotation
- Corpus Engines
- Data Formats
- Other Resources

Yes ▼ No

Graph showing the percentage of participants with experience in various aspects of corpus linguistics.
Your expectations

- Making a feasible plan that could connect all these institutions in a joint project, new possibilities of extending the present project, define steps for future collaboration in the field of corpus design, network with other colleagues in the field and, hopefully, embark on a national or regional project in the area

- get acquainted with the experience my colleagues have
Your expectations

- learn more about the available tools, improve my knowledge of corpus linguistics
- creation and use of the corpora
- higher level of knowledge in the area of (spoken) corpus design
- agreement about applying the same methodology and standards in corpus creation, criteria for transcription, annotation and corpus representativeness
Programme Monday

- 11:00
  - A basic introduction into types and possible usages of corpora, existing corpora, and topics in corpus analysis

- 12:00
  - Lunch

- 13:00 UdL6
  - IO Registration and Cash desk
Programme Tuesday

- **09:00** UdL6, 1066e
  - Spoken Corpora

- **11:00**
  - The sociolinguistic dimension in copora

- **12:30**
  - Lunch

- **14:00 – 16:00**
  - Experiences in Corpus Creation: Input from Participants

- **19:00** Bistro Ribelle
  - Workshop Dinner
Programme Wednesday

- 9:00 DOR65, 5.57
  Desiderata in Corpus Linguistics and Corpus Creation

- 11:00
  Current issues in (spoken) corpus design

- 12:00
  Lunch

- 13:30
  Developing project ideas
<table>
<thead>
<tr>
<th>Time</th>
<th>Session</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00</td>
<td>Slavic Diachronic Corpora</td>
</tr>
<tr>
<td>10:45</td>
<td>Standardizing sociolinguistic annotation and metadata</td>
</tr>
<tr>
<td>12:00</td>
<td>Lunch</td>
</tr>
<tr>
<td>14:00</td>
<td>Linguistic tagsets and the MULTEXT-East specifications</td>
</tr>
<tr>
<td>14:45</td>
<td>A MULTEXT-East tagset for Albanian</td>
</tr>
<tr>
<td>14:45</td>
<td>A MULTEXT-East tagset for Historical Texts</td>
</tr>
<tr>
<td>17:00</td>
<td>Outlook</td>
</tr>
<tr>
<td>18:00</td>
<td></td>
</tr>
</tbody>
</table>
Registration

- Registration is in the main building, International office
  Room 2258
- Please go there one-by-one!
- You’ll receive a cash note for the cashier
- The Cashier is also in the main building
  1. Floor, room 1002
  Passports needed!
- Tomorrow, we will also be in the main building!