

# CLARIN 2021 Bazaar Session

Chair: Frieda Steurs

Day 2

Tuesday 28 September

15.00 - 16.00 (CEST)



# CLARIN ANNUAL CONFERENCE

Clarín Bazaar - chair: F. Steurs  
28/09/2021 - 15:00

TRIPLE

UPSKILL

AFOR



**TRAINING &  
KNOWLEDGE  
EXCHANGE**



**DIGITAL  
HUMANITIES**

AIUCD's  
ONLINE  
EXPERIENCE

LITHUANIAN  
COLLOCATIONS



**DATA CURATION  
USING NLP**

GREEKSCHOOLS

FONTI 4.0

WAGEINDICATOR  
COLLECTIVE  
AGREEMENTS DB

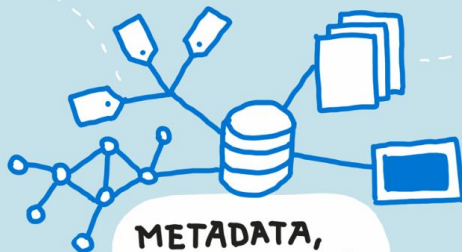
RESOURCE  
& TOOL  
FAMILIES

DH COURSE  
REGISTRY

CLARIN'S  
ONLINE  
EXPERIENCE

FORMAT  
RECOMMENDATIONS

CMDI



**METADATA,  
CITATION, RDM**

THE SWITCHBOARD



**ADVANCED TECHNOLOGIES  
& RESOURCE HARMONISATION**

DATA  
CITATION

DELEAD

UMBRELLABIRD

# Breakout Rooms

## Room 1: Training and Knowledge Exchange

- Put Yourself on the Map: Join the DH Course Registry!
- How CLARIN Went Virtual – 1.5 Years Online, in Data and Images
- UPSKILLS: Towards New Teaching Resources for a New Professional Profile
- TRIPLE Training on Open Science

## Room 2: Digital Humanities

- AIUCD2021 Conference Went Virtual through the Support of CLARIN: Experience, Opportunities and Suggestions
- AFor: Collecting, Analysing, Sharing Memories through a Multidisciplinary Open Archive of Oral Sources
- FONTI 4.0: Towards an Unsupervised Transcription Chain for Accessible Analogue Oral Archives

## Room 3: Data Curation Using NLP

- Lithuanian Arbitrary Collocations: Recognition Criteria and Methods
- Domain-Specific Languages on Editing Papyri: The GreekSchools Case Study
- The WageIndicator Collective Agreements Database and its Machine Learning-Based Developments

## Room 4: Metadata, Citation, RDM

- Component Metadata Infrastructure (CMDI) Help Desk, News, Discussion and Rumours
- Much Ado About Data Citation
- Ask your SIS: Collecting Centre Recommendations on Data Deposition Formats
- Catching up with the DELAD Initiative to Share Corpora of Speech of Individuals with Communication Disorders

## Room 5: Advanced Technologies and Resource Harmonisation

- The Switchboard: Demo and Discussion of Existing and New Features
- CLARIN Resource and Tool Families
- UmbrellaBird, Building a Data Stewardship Organisation for an International Multilingual Large Language Model Training Corpus

## Breakout Rooms

- **Room 1: Training and Knowledge Exchange**
- **Room 2: Digital Humanities**
- **Room 3: Data Curation Using NLP**
- **Room 4: Metadata, Citation, RDM**
- **Room 5: Advanced Technologies and Resource Harmonisation**
- **Room 6: (Extra Room) Training and Knowledge Exchange**
- **Room 7: (Extra Room) Digital Humanities**
- **Room 8: (Extra Room) Data Curation Using NLP**
- **Room 9: (Extra Room) Metadata, Citation, RDM**
- **Room 10: (Extra Room) Advanced Technologies and Resource Harmonisation**

# Division of Roles between Stall Holders within One Breakout Room

- Stall Holders:
  - Follow the order of stalls in the programme for pitch/demo
  - Be concise with your first pitch - no longer than 2-3 minutes
  - Share your screen where needed
  - Feel free to moderate the Q&A
- Audience:
  - Raise your questions in the chat.

# Use of Breakout Rooms (Main + Extra)

- 15.00 - 15.15
  - Initial pitches/demos in the topical breakout rooms (1-5)
- 15.15 - 16.00
  - If the room is not too crowded, the discussion will take place there
  - If the room is too crowded, please move back to the plenary space to coordinate with the host which stall-specific discussion is moved to which of the additional breakout rooms (6-10).

Enjoy the fruitful interaction!