

Title Minutes SCCTC 25 May 2021
Version 1
Author(s) DVU, MM, LS
Date 2021-05-31
Status Approved
Distribution Centre Committee
ID CE-2021-1881



Participants

Martin Matthiesen (Chair, FI), Matej Ďurčo (AT), Pavel Stranak (CZ), Thomas Eckart (DE), Bart Jongejan (DK), Krista Liin (EE), Vanja Štefanec (HR), Roberts Dargis (LV), Leif-Jöran Olsson (SE), Etienne Petitjean (FR), Marinos Ioannides (CY), Luís Gomes (PT), Tomasz Naskret (PL), Lene Offersgaard (assessment committee), Dieter Van Uytvanck (CLARIN ERIC)

0 Action points

#	Action	By whom	By when
1	Follow-up: Recommendations on PIDs outcomes: https://docs.google.com/document/d/1hLhtKwVe5fifFWvU28mukAGC0yl4gaOv6XNT4MVkioQ/edit#	Daan & Dieter	March 2021

1 Agenda

1. Agenda: Request for changes?
2. Approval [https://office.clarin.eu/v/CE-2021-1858-SCCTC_Minutes_14_April_2021_v1.docx | minutes last meeting] (14.4.2021) & action point status
3. Centre assessments
4. Centre Meeting 2021, suggestions for discussion (~20min):
 - a) Hold it on 9 and 10 June (taskforces can make it)
 - b) Discuss how to fill the slots: [<https://docs.google.com/document/d/1mtMpjuqs3z-2WMs1DvGwl9nvsObDm3NS8nTz6tAB77s/edit#> | please add your input to the Google Doc!]
 - c) Have the regular SCCTC meeting end of June (there is no reason to couple it since it will be virtual), set date via doodle
5. Status update per country/member (please provide a short bullet-wise summary in [https://docs.google.com/document/u/0/d/1MpVX9II7o-oPiOp4-lbxTSjPH_1dYB3M_jSZVP9YIvl/edit | the Google Doc]):
6. AOB

2 Approval minutes last meeting & action points

The meeting minutes of 14 April 2021 ([CE-2021-1858](#)) are approved.

3 Centre assessments

Lene: Report of 16th round ([CE-2021-1857](#))

Recommendation to regrant B-centre label to ARCHE and TalkBank.

SCCTC approves the report.

Date for 18th and 19th round not yet set, assessment committee will propose exact dates for publication at <https://www.clarin.eu/content/assessment-reports>

4 Centre meeting

Centre meeting takes place on 9 and 10 June.

Draft agenda to be published on Thu 27 May, depending on interest per topic, then broader distribution (action point DVU).

Next SCCTC: 30 June at 14:00 CEST

5 Status update per country/member

Austria

- [Tour de CLARIN: CLARIN Centre-B Vienna](#)
- ARCHE-Repository: dissemination service for TEI documents
=> allows to offer (simple Digital Editions directly from the repository)
- Intensive work in SSHOC
 - Further development of SSH Open Marketplace (towards curation)
 - Integration of ARCHE with LR Switchboard and (potentially) VCR)
 - Conversion hub: <https://conversion-hub.acdh-dev.oeaw.ac.at/> (WIP)

Bulgaria

- No report

Croatia

- CLARIN ERIC and HR-CLARIN consortium has been presented at the thematic block “Initiatives for the Open Science” at the Days of E-infrastructures conference organized by SRCE -- Zagreb University Computing Centre (<https://dei.srce.hr>).
- A series of coordinating meetings with other ERIC and ESFRI project national coordinators was held in the Ministry of Science and Education where full support has been acknowledged for further mid-term period.

Czech Republic

- Preparation for the CZ CLARIN Cafe (April 15).
- Negotiation to join the Infrastructure Emergency funding call
- Preparation of European Language Equality surveys (includes CLARIN as partner)

Cyprus

- No report

Denmark

- The Danish Parliament Corpus 2009 - 2017, v2, with metadata about the MPs, timing of the speeches and subject annotation of each agenda item is made available (<http://hdl.handle.net/20.500.12115/44>).

Estonia

- Organizing user guides, documentation and helpdesk
- Automation of updating our services, moving towards Kubernetes.
- Webseminar on our services April 27th: data repository services, (federated) content search.

Finland

- CTS and CLARIN B Applications submitted
- FIN-CLARIAH funding application for 2022-23 submitted
- Preparing for update of CLARIN Agreement

France

- CTS certification in progress for the Ortolang C-Center (more documentation produced)
- Presentation of CLARIN at the National Library (BnF) mainly in the context of the “BnF Data Lab” (See <https://www.cnrs.fr/en/node/4520> ... in French sorry)
- Presentation of CLARIN at [AUF](#) (Agence Universitaire de la Francophonie)

Germany

- Funding recommendation for next project phase (“Text+”), final decision: July 2021
- Centres start active work on preparation for CTS/CLARIN recertification (due by: 10/2021 - 05/2022; 18th or 19th assessment round)

Greece

- Preparing the next version of clarin:el software (backend and UI)

Hungary

- No report

Iceland

- CLARIN-IS was one of the organizers of an online conference on language technology, “Máltækniþytingin” (The LT Revolution), on May 18. The CLARIN coordinator gave a talk and introduced CLARIN-IS.

Italy

- ILC4CLARIN
 - The repository is increasing its number of lexical resources and tools for DH. We continued describing the Latin resources from ALIM (Italian Archive for Medieval Latin (resources)). Currently 11 out of 49 volumes collecting documents from VIII to XI Century. This activity is still in progress, but we plan to end by Summer. This completes the activities on ALIM (for the time being). We also started collecting Latin resources from the ERC project "LiLa: Linking Latin": they are of different types, including Latin treebank; derivational, valency and polarity lexica; annotated data for lemmatizers and PoS taggers; etymological dictionaries and a large collection of Latin lemmas in LLOD

formats. Some of them are already available from the VLO (<https://vlo.clarin.eu/?0&fq=collection:CIRCSE&fqType=collection:or&q=CIRCSE>). The repository hosts Greek resources as well. The last two are a collection of Cretan inscriptions (along with its web interface) and a TEI corpus which is an outcome of seminars on the digitization of fragments in ancient Greek (held at the University of Parma). It's important to cite the polimodal corpus (<http://hdl.handle.net/20.500.11752/OPEN-534>), a speech transcription of a journalistic TV program (Annotation has been done using XML as markup language and following the TEI standard for Speech Transcripts in terms of utterances). It is part of the language resource family.

- B recertification and CTS submitted.
- Brainstorming on the integration of CLARIN-IT/MS data in European Language Grid
- Activities on ALIM in CLARIN-IT still continue;
- Still awaiting from CTS.

Latvia

- DHN 2020 post-proceedings published [CEUR-WS.org/Vol-2865 - Digital Humanities in the Nordic Countries \(DHN 2020\)](https://ceur-ws.org/Vol-2865-Digital_Humanities_in_the_Nordic_Countries_DHN_2020)
- Participation at WINLP 2021 [Panel on Low Resource Languages and Diversity in NLP Research](#)

In collaboration with the State research programme “Digital Resources for Humanities: Integration and Development” presentation of Latvian language corpora at [Demonstration seminar on the use of Digital Resources in the Humanities](#) as part of the "[DIGITAL WEEK 2021](#)" (March 24)

Lithuania

- Keynote speech “Open Access Resources and Tools for Language Learning” at an international conference "[Diving into Challenges of Online Language Learning and Teaching. Benefits from OER](#)”
- Preparation of the “Questionnaire for the analysis of the impact of Lithuania's membership in international research infrastructures” requested by the Research Council of Lithuania

The Netherlands

- CLARIAH Lunch Lecture on 17th Century Newspapers by Nicoline van der Sijs and Joris van Zundert (April 8, 2021)
- Lecture by Jan Odijk at the Experimental Methods in Language Acquisition Research (EMLAR) workshop on Semi-Automatic Assessment of Spontaneous Language (SASTA). April 14, 2021
- Demo of Semi-Automatic Assessment of Spontaneous Language (SASTA) by Jan Odijk and Jelte van Boheemen at the UiL-OTS Language Science Day (April 16, 2021)

Norway

- The National Library of Norway has designed and implemented a new version of the national resource catalogue
- The plan for the ELMCIP (Electronic literature knowledge base) upgrade to Drupal and integration in CLARINO has been finalized

Poland

- Created facebook website(<https://www.facebook.com/Clarín-PL-Infrastruktura-naukowa-107248931464578>) and newsletter

Portugal

- Contribution to “Chatting with BoD around a coffee”
- Masterclass to PhD students in Linguistics of Universidade Nova de Lisboa about PORTULAN CLARIN workbench
- Dissemination activities contributing to the celebration of the World Day of the Portuguese Language (May 5)
- Improvements on the workbench technical infrastructure and bug fixing

Slovenia

- Released ParlaMint V2.0 (<http://hdl.handle.net/11356/1388>, <http://hdl.handle.net/11356/1405>)
- Now working on ParlaMint 2.1, the final version in the scope of the project

South Africa

- No report

Sweden

- Preparing workshops and new releases of resources Superlim (superglue), lexical resource Coderoomor
- Preparing for reassessment, updating repo software, updating handle server
- Issued recommendation on Matomo for statistics, switching helpdesk from otrs to zammad

United Kingdom

- OTA repository: A bug in DSpace means that our Matomo installation isn't working, and so we are not getting the detailed stats on usage that we were expecting.
- CLARIN-UK website: search box for VLO added on home page
- A funding proposal is under preparation to support the network coordination costs of the CLARIN-UK network in 2022.

USA (third party TalkBank - CMU)

- No report

6 AOB

Questions from Dieter:

- CLARIN repos that host training materials? (eg DigiLing -> <http://hdl.handle.net/11356/1243>)

Matej has some interesting input for Lulianna on training material repos, Dieter will connect both [done].

- User requirements. Any relevant reports from consortia or centres where these are described?

Pavel: Maybe not a classic user requirement, but the question we are asked most frequently is whether researchers can be given access to or can deposit literary texts. Unfortunately, this is not possible due to copyright restrictions. This was also brought up in CLIC. Also, the current exceptions (in the TDM regulation) are not enough to allow sharing with other researchers.

Dieter will report about this to the NCF.

- Any other centre (apart from IDS) interested in handle prefix migration from CNRI to EPIC/GWDG? (Final call)

No further centres indicate interest.

Next meeting:

Wednesday 30 June 2021 at 14:00 CEST.