

Title	SCCTC Minutes 19 October 2021
Version	1
Author(s)	Dieter Van Uytvanck
Date	2021-10-19
Status	Approved
Distribution	SCCTC
ID	CE-2021-1962

Participants

Martin Matthiesen	Chair	Finland
Matej Ďurčo	Member	Austria
Jesse de Does	Member	Belgium
Ivan Georgiev	Member	Bulgaria
Vanja Štefanec	Member	Croatia
Jan Hajič	Member	Czechia
Simon Gray	Member	Denmark
Krista Liin	Member	Estonia
Etienne Petitjean	Observer	France
Thomas Eckart	Member	Germany
Dimitris Galanis	Member	Greece
Samúel Thórisson	Observer	Iceland
Riccardo Del Gratta	Member	Italy
Roberts Darģis	Member	Latvia
Andrius Utkas	Member	Lithuania
Menzo Windhouwer	Member	Netherlands
Hemed Al Ruwehy	Member	Norway
Tomasz Naskręć	Member	Poland
Luís Gomes	Member	Portugal
Cyprian Laskowski	Member	Slovenia
Leif-Jöran Olsson	Member	Sweden
Langa Khumalo	Observer	South Africa
Martin Wynne	Observer	United Kingdom

Deputy members: Felix Helfer (DE) standing in for Thomas Eckart

Others: Dieter Van Uytvanck (CLARIN ERIC), Lene Offersgaard (assessment committee)

Action points

#	Action	By whom	By when
1962.1	Follow-up: Recommendations on PIDs outcomes: https://docs.google.com/document/d/1hLhtKwVe5fifFWvU28mukAGC0yl4gaOv6XNT4MVkioQ/edit#	Daan & Dieter	December 2021
1962.2	Check documentation status of attribute aggregator and prepare a small amendment to the B-centre requirements with a strong recommendation to configure the attribute aggregator	Martin & Dieter	November 2021
1962.3	Discuss end of terms of the chairs and potential renewal with SCCTC and CAC chairs	Dieter	November 2021

Agenda

1. Agenda: Request for changes? (2 min)
2. Approval [minutes last meeting](#) (21.9.2021) & action point status (5 min)
3. Centre assessments (10 min)
4. Proposal: Survey of Sensitive Data activities in CLARIN Centres? (15 min)
 - a. Topic was presented on Clarin Annual Conference
 - b. Is this relevant to Centres?
 - c. Should we plan presentations on the topic?
5. [Status update per country/member](#) (please provide a short bullet-wise summary in [| the Google Doc]): (15 min)
6. AOB (5 min)

The agenda is approved as is.

Approval of minutes last meeting & action points

Approved without comments

#AP4: FH reports: IG still active, good to know that the project chair (David Wilcox) is leaving and will be replaced

Centre assessments

See document CE-2021-1960 (round 17 report), distributed by email

LO reports:

Language Bank of Finland

ILC4CLARIN

Both re-assessed without comments, Språkbanken will be participating in round 18.

CE-2021-1960 (round 17 report) is approved unanimously (MM abstains).

Round 18: MPI, HuC, CLARIN-LV, Språkbanken, CLARIN:EL

KL: CELR understaffed, might not participate in round 18

Proposal: Survey of Sensitive Data activities in CLARIN Centres?

Topic was presented at the CLARIN Annual Conference. Is this relevant to Centres? Should we plan presentations on the topic?

MW: MPI is building up some knowledge, RU is working on NLP software for closed environments (without internet)

JdD: what is sensitive?

MW: mostly medical areas

MM: copyrighted material is a little easier to deal with; I was mostly referring to eg videos of children with speech

RdG: important topic for us

EP: we had to deal with such corpora (child speech/video) in Ortolang, would be nice to have a good overview of the issue

JH: agree, would be nice to have a best practice to process the data wrt anonymisation, eg which laws apply; might be good to have a K-centre on this?

LJO: second that, there are a lot of misunderstandings on this topic

MM: I would invite each centre to come up with a problem/solution description and send it by email to me. Depending on the results we could even have a separate workshop on this topic.

Status update per country/member

Short bullet-wise summaries can be found in:

https://docs.google.com/document/d/1MNArHnXyaOAsWTgfsyUV2XxMO8A_jxi7KkNE1DUZto0/edit

AOB

JdD: how to deal with requests for DL-based computationally intensive requests? (eg machine translation evaluation) Might require hardware, maintenance time.

LJO: we have standardized consultations for this, with a fixed price for an extra machine with maintenance. Consultations can have different forms, but are good to streamline the requirements.

MM: we do not have such a standardized process. We have asked projects to provide a local system/software administrator for these services, to deal with possible incidents. Often the projects do not want to do this. We have a few services though where this model works. Once a local admin is no longer available we would need to take down the service.

LJO: We have a policy to require statistics (matomo) and tests.

JH: Same approach as MM, only running services for people on our payroll. Needs: IT staff, someone who understands the service, authors of the service. We are refusing other requests. European Language Grid is relying on services hosted by the partners, Docker-based. How it will work in the future is currently unclear.

Next meeting: DVU will send a doodle for 15-18 November.