

Title Minutes SCCTC 16 May 2017
Version 1
Author(s) LST
Date 2017-05-18
Status Final
Distribution Centre Committee
ID CE-2017-1030



Participants

Matej Ďurčo (AT), Wolfgang Sauer (AT, observer), Jan Hajič (CZ), Pavel Straňák (CZ), Dirk Goldhahn (DE), Lene Offersgaard (DK), Krista Liin (EE), Martin Mathiessen (FI), Marc Kemps-Snijders (NL), Marcin Pol (PL), Riccardo Del Gratta (IT), Leif-Jöran Olsson (SE), Dieter Van Uytvanck (CLARIN ERIC, Chair), Linda Stokman (CLARIN ERIC, minutes), Twan Goosen (CLARIN ERIC), Franciska de Jong (CLARIN ERIC, observer), Alexander Koenig (EURAC, Observer)

Excused: Kiril Simov (BG), Harris Papageorgiou (GR), Andrius Utka (LT), Roberts Dargis (LV), Freddy Wetjen (NO), João Silva (PT), Cyprian Laskowski (SI), Martin Wynne (UK)

0 Action points

#	Action	By whom	By when
1	Provide centre logos and guidelines on how to use these in the style guide (also add examples how to link these to the relevant page)	CLARIN Office	May 2017
2	Set up small author task force to summarize proposals made for archiving software (agenda point 4)	Lene	Next SCCTC meeting
3	Submission of evaluation documents and web pages must be accompanied by a short summary in English. Add this requirement to the B-centre criteria	Lene	May 2017
4	CLARIN ERIC will monitor the progress, and gather general guidelines and common answers to questions for DSA	DVU	a.s.a.p.
5	Propose new centre labels as a more human-readable alternative to the letters	DVU/Karolina	a.s.a.p.

1 Agenda

1. Agenda
2. Approval minutes last meeting & action point status ([CE-2017-0994](#))
3. B-Centre assessments
 - a. Outcome of the 8th assessment round ([CE-2017-1015](#))
 - b. Status update, planning of the upcoming assessment rounds
 - c. How to handle expiration of centre certification
4. A-services: discussion slot ([CE-2017-1029](#))

5. Report about Europeana integration ([CE-2017-0996](#))
6. [Reports from the task forces](#): (taskforce coordinators)
 - [CMDI](#)
 - [Metadata curation](#)
 - [Federated Content Search](#)
 - [Persistent Identifiers](#)
 - [AAI](#)
7. Status update per country/member (all) -> **please add a snippet for your country to the [reporting google doc](#)**
8. Any other business

2 Approval minutes and action points

The minutes of the March 2017 SCCTC meeting have been approved.

3 B-Centre assessments

- a. **Outcome of the 8th assessment round** ([CE-2017-1015](#))

AAI

Good recommendation about guiding users at the login stage:

<https://lindat.mff.cuni.cz/en/how-do-i-sign-up>

Use of public CMDI profiles

The CMDI profiles, that a centre uses, have to be public to be accepted in assessment. In the current version of CMDI, public profiles cannot get updates of links to CCR. The development status of profiles might offer a solution to this. CMDI taskforce will have a look at this (fixing CCR links for public profiles).

Approval of prolongations B-Centre status

The prolongations of the B-Centre status for the Meertens Institute and University of Copenhagen are approved as stated in the CE-2017-1015 document.

Logos

General CLARIN logo should link to CLARIN website and can be used on webpage to show consortia is affiliated with CLARIN.

B-centre logo could link to [B-centre page](#)

C-centre logo could link to page with explanation on what it means to be a C-centre (<https://www.clarin.eu/content/clarin-centres>)

CLARIN ERIC should provide logos and some guidelines on how to use them and some example on how to link to them.

Insights on new DSA procedure

- Marc Kemp Snijders-> It is an improvement, more focus on procedures, but: real hassle -> not all procedures very clearly specified, technical stuff mostly covered in

CLARIN (e.g. version control) all in all: good exercise, but takes lot of time, forces policies/procedures-> more specific than previous version

- Lene Offersgaard -> more writing
- Matej Durco -> positive note on new DSA: nice guideline towards stabilization of procedures, complete reshuffle

b. Status update, planning of the upcoming assessment rounds

Assessment for Meertens and Copenhagen are approved.

9th assessment round: assessments for CMU, IVDNT have been submitted

Waiting for response from ING and assessment Språkbanken is expected soon.

10th assessment round: Vienna, Text archive Oxford, CLARIN-EL and CLARIN-IT (ILC-CNR) have expressed willingness to participate.

The deadline for the 10th assessment round is mid October 2017. If you start a new re-assessment now you have to use the new DSA 17-19 procedure.

c. How to handle expiration of centre certification

If the centre has been informed beforehand about the expiration of the B-status and has not yet renewed we will bring down the status from B to C (temporarily if the centre starts a new assessment procedure with the aim of regaining the B status), inform the Centre about this and ask them to remove the B-centre logo.

4 A-services: discussion slot ([CE-2017-1029](#))

There are two discussions:

1. The labelling of the centres is not transparent (this is also a question for our communication officer Karolina Badzmierowska).
2. Conceptual questions, like:
 - a. Is the role of providing the service purely about hosting?
 - b. How do you reward the specific development of these services?

Lene Offersgaard: Many are puzzled by B-centre and A-service (why go for B if there is an A, there is confusion). Think about the labelling. View: A= Infrastructure

Marc Kemp-Snijders: A-service is the backbone / core of what CLARIN is. View: A-service does not require software development. B-centre could be seen as hierarchy difference

Marcin Pol: Hosting (maintenance, servers, updates, update new versions) should be sufficient for an A-service.

Maciej Āurčo: no benefits, more work for no money.

Jan Hajič: Instead of labelling the services give them names (backbone, community services). For some B-centres you have to have A services or not. Do not use a B-centre logo but something other. Look at it from 2 dimensions: one dimension = core CLARIN Centre, other dimension = what services you provide (and by name, not by a, b, c, d).

Pavel Straňák: The CLARIN Centres that are part of the consortia might ask why they have to pay if there can be labelled centres outside of the consortia as well at no cost.

DVU: These services are of a different dimension (A vs E = external computing centres, eg providing EPIC services).

Franciska de Jong: Take a few more steps in finding common ground and then report to the BoD and NCF. First try to summarize what the committee already agrees on, make a division between core services and other services and create a document for input to the NCF. The outcome of the summarization can be taken up in the BoD June meeting first and then be discussed in the f2f NCF September meeting.

A short tour de table, asking for potential A-service candidates and willingness/interest to host them, results in the following list:

- AT: curation module
- CZ: Dspace repository + deposit services
- PL: good hardware to run A-services, eg PID service
- SE: plan to host A services, SND -> repository, will harvest from centres + national node
- FI: REMS, (no interested users so far) + sensitive data service
- DE: helpdesk, weblicht
- NL: OpenSKOS/CCR (not completely sure about making it an A-service, depends on SLA/risk assessment)

The discussion is concluded by the planning of the next steps in this process: create a document describing the outcomes of this discussion (DVU, with input from JH and FdJ), putting this on the agenda of the BoD and NCF (most probably the face-to-face meeting in Budapest).

5 Reports from the task forces [Report about Europeana integration \(CE-2017-0996\)](#)

Report by Twan Goosen:

- Digital second stage of development is planned for August. Europeana wants to share their data (mostly from libraries) with research communities.
- Taking up Europeana data in the VLO and making it available to third party infrastructures. Conversion to CMDI is complete and the outcome can be incorporated into the VLO. Made a selection of what we want to include. Language Resource Switchboard hackathon will also include some Europeana resources.
- An abstract explaining the work in detail has been submitted for the CLARIN annual conference.

6 Reports from the task forces

CMDI:

- Current main task is the CMDI best practices document (with curation task force), see [recent meetings](#) for details. A first draft for a large part of document is done, but there still is quite some work left as well. A number of members from both taskforces have submitted an abstract for the CAC describing the rationale behind the document and a summary of the progress. Development on core CMDI related software components (Component Registry, [VLO](#), CLAVAS) continues with a number of task force members involved.

Metadata curation:

- Work on best practices with CMDI task force (see above)
- Picked up development and maintenance for the curation module again after departure of Davor Ostojic, who has now been replaced by Wolfgang Sauer.

Federated Content Search:

- See deliverable CLARINPLUS-D2.9: [Federated Content Search Engine v2 \(software\)](#)

Persistent Identifiers:

- No report

AAI:

- No meeting in Utrecht due to lack of members present
- Currently under discussion: how to deal with self-registering IdPs

7 Status report of the countries

These reports are a *public* extract of the common NCF and SCCTC monthly report for April 2017, which is also available as *internal* document CE-2017-1024.

Austria

- Failed to submit for latest re-assessment round due to complexities and delays refurbishing (in fact completely rebuilding) the repository solution based on Fedora 4. But doing good progress and definitely ready for the next round.
- Re-launched DH course registry: <https://registries.clarin-dariah.eu/courses/> at the DARIAH annual event.
 - As first tangible collaborative service of CLARIN and DARIAH CLARIN is now supporting further development of the application and CLARIN and DARIAH will collaboratively work on curation of the content (by delegating national moderators).
- Development (and maintenance) of curation module and vlo-curation instance stalled due to main developer Davor leaving ACDH (as of end of March). However since May new colleague Wolfgang Sauer taking over curation agendas, as well as CLARIN in Parthenos concerns
- Pushing for first practical use case NERLIX in PARTHENOS project ([description](#)) allowing to apply services on resources similarity to LRS, but in different technical setup.

Bulgaria

No written report

Czech Republic

- Meeting in Amsterdam about CLARIN-EUDAT cooperation
- Preparation for EUDAT-CLARIN repository connector over HTTP API
- Improvement of TreeQuery service: Bookmark lists for queries

- Implementing new features to CLARIN Dspace: upload from cloud storage (in progress)
- Co-organization of UD Shared Task (for CoNLL 2017, Aug. 2017)
- Preparation of EAMT conference (co-organization), takes place in Prague, May 2017

Denmark

- DSA work and DM documentation
- Restructuring clarin.dk frontpage
- Creating fixed workflows for using tools, to make it easier for users to use them

Dutch Language Union

No written report

Estonia

- Upgrading our KORP and adding corpora.
- Work on getting Estonian (EstNLTK) into Weblicht (by Sven Laur)
- Adding the new consortium member (Literary Museum) to CELR consortium, submitting appropriate paperwork for funders and integrating their resources.
- Working on adding more IdP (through TAAT, not directly to eduGAIN).
- Mart Klanberg (our server administrator and technical person) is leaving - documenting and redistributing his workload.
- Some problems adding Discovery Service to our SP. Also still working on the DSA documents, so we decided not to take part in the 8th assessment round.

Finland

Done

- [English translation](#) for Defining usage of PIDs for updates, accruing corpora
- Workflow manager [Mylly](#) presented at [Nordic Clarin Network Workshop in Gothenburg](#).

In Progress

- More co-operation with Sweden in Korp development planned. Sweden moves to Github, we fork.
- We are moving towards agile methods like Kanban, to improve communication in the team.
- B2SAFE-pilot: No news.
- Sensivite Data: CSC had a workshop on the topic. There is still a lot of confusion. At least Finnish national legislation does explicitly allow the use of personal data for research purposes, in some cases even without consent. Reclassification of classified data is also possible.
 - CSC is piloting a [REMS](#) / [OAuth](#) / [FUSE](#) based access control mechanism. The underlying data can be encrypted. The Language Bank might take part in the pilot.
 - ISO 27001 for the Language Bank is progressing, we assessed the requirements. One challenge is that the Language Bank is a joint service of CSC and U Helsinki and not a separate organization.

Germany

- Central Monitoring:
 - There are a few failing probes due to different reasons (small changes at services, malfunctioning probes, actual problems of the services), We will get in contact with the different institutions.
 - There will be monthly reports for central services and centres of CLARIN-D soon (average uptime last month, last year...). Is this necessary or desirable in a CLARIN-EU-wide version.

- Assessment:
 - CLARIN-D will create a document on changes in DSA-procedures in the next weeks, to allow for German centres to prepare accordingly. Is there any existing documentation on changes? Who to contact at DSA that in November CLARIN-D will have reassessment?
 - All CLARIN-D centres will request B-centre-reassessment early 2018
- MIME types (in WebLicht): Discussion in Saarbrücken - (temporary) agreement on first list of MIME types and format variants used in CLARIN-D. Agreement on quick and simple solution, which needs minor changes to WebLicht (fixed list of possible combinations)
- Computing Centre Jülich: Working on Memorandum of Understanding on cooperation. Currently we are using hosting services of EUDAT in Jülich.

Greece

No written report

Hungary

No written report

Italy

- Report from ILC4CLARIN Centre: worked hard to adopt the next release of clarin-dspace repository software (2017.02), reported and closed some issues to the clarin-dspace github, especially on the send-file feature; The DSA procedure for the repository began on April (26/4)

Latvia

No written report

Lithuania

On April 3-4, Andrius Utkas gave a seminar for doctoral students “Language Technology, Language Resources and Tools” (<http://clarin-lt.lt/?p=571>)

The Netherlands

CLARIAH mid term review has been submitted. CLARIAH+ follow up proposal is being prepared. Meertens is preparing DSA submission. Daan Broeder will replace Marc Kemps-Snijders as SCCTC representative.

Norway

No written report

Poland

- Workshops
- D-Space as user provider for internal web applications
- NextCloud integration with clarin-pl application including d-space

Portugal

- Meeting with the Portuguese RCTSai federation.

Slovenia

- Established the CLARIN.SI organisation on GitHub, <https://github.com/clarinsi>
- Registered first service for Slovene, Croatian and Serbian in WebLicht

Sweden

No written report

United Kingdom
No written report

6 Any other business

None

Next meeting: A doodle will be sent out in the course of May 2017 (we will take into account the NCF meetings).