

Title Minutes SCTCC April 2016
Version 1
Author(s) LS
Date 2016-05-10
Status Final
Distribution Centre Committee
ID CE-2016-0768



Participants

Dana Dannélls (SE), Andrius Utka (LT), Freddy Wetjen (NO), Francesca Frontini (IT), Krista Liin (EE), Lene Offersgaard (DK), Marcin Pol (PL), Martin Mathiessen (FI), Matej Ďurčo (AT), Tomáš Erjavec (SI), Xaris Papageorgiou (GR), Pavel Straňák (CZ), Dieter Van Uytvanck (CLARIN ERIC, Chair), Linda Stokman (CLARIN ERIC, minutes)

Excused: Dirk Goldhahn (DE), Marc Kemps-Snijders (NL) Kiril Simov (BG),

0 Action points

Action	By whom	By when
Update FAQ of Persistent Identifiers	DvU, MP	May 2016

1 Agenda

The following agenda was accepted:

1. Agenda
2. Approval minutes last meeting & action point status ([CE-2016-0741](#))
3. Status of the A-service assessment pilot
4. Renewed DSA procedure ([background](#)) – effect on centre assessment timing
5. Reports from the task forces: (task force coordinators)
 1. [CMDI](#)
 2. [Metadata curation](#)
 3. [Federated Content Search](#)
 4. [Persistent Identifiers](#)
 5. [AAI](#)
6. Status update per country/member (all)
7. Any other business

2 Approval minutes and action points

The minutes of the February 2016 SCCTC meeting have been approved.

3 Status of the A-service assessment pilot

Lene (DK): Decided to put the B-centre assessment in the same distribute round to the assessment committee later this week.

Dieter: The LINDAT application is well received and put up for assessment by assessment committee.

4 Renewed DSA procedure ([background](#)) – effect on centre assessment timing

The DSA procedure is slightly going to change, see the background link.

Lene (DK) reports: Re-assessment of B-centre is currently not possible, waiting for DSA to release new questions to be able to apply for re-assessment.

Dieter: We need to find out if there is a difference between the new and re-assessments. Dieter will try to clarify with the DSA board.

Dieter suggests to prolong B-centres for now, until the new DSA-procedure is in place. Since there were no objections to this suggestion it has now been taken up as policy.

5 Reports from the task forces

Short status update from other taskforces:

- **CMDI:**

Work on the new 1.2 spec: <https://trac.clarin.eu/wiki/CMDI%201.2/Specification>

- **Metadata curation:**

Deliverable D2.1 submitted

Technical description of the [Curation Module](#)

Code on github: <https://github.com/acdh-oeaw/clarin-curation-module>

A simple web application for human interaction: <http://clarin.oeaw.ac.at/curate>

(allows to evaluate profile or instance (by URL))

Further work planned / on-going:

- Setup a continuous pipeline (integrate with VLO-harvesting/importing)
- Setup the feedback loop with the data providers
- Provide HTML version of the reports (currently only XML)
- General user-experience of the web app

- **Federated Content Search:**

No report

- **Persistent Identifiers:**

TODO: Update on previously reported issues:

2 issues have been reported:

- Freddy Wetjen (Norway): a problem with a specific version of the EPIC API that returns HTML instead of JSON. Will be investigated in-depth with the help of Willem Elbers (willem@clarin.eu).
 - Status update: GWDG has confirmed this issue was still present at the time of the issue report (18.2.2016). A bug fix was available already since mid 2015, however a dependency on handle system version 8 prevented deployment of this fix, according to GWDG. GWDG has updated the involved system(s), including this fix, [on 17.3.2016](#). I still have to confirm with the user if the issue is actually resolved now.
- Tomaž Erjavec (Slovenia): problem with resolving of the handles with prefix 1839. A first analysis has shown there is a synchronisation issue with the handle server at the MPI and the GWDG. This has been reported to the support team of the GWDG. Here too Willem will follow-up.

- The synchronisation issue related to the 1839 handle are already existing for quite some time. Given the (human) resource limitations a the MPI it is not foreseen that the replication will be solved anytime soon. Currently 50% of the resolution requests are handled by GWDG (round-robin) and might result in errors caused by the synchronization issue. This is not ideal and the current proposed solution which is under discussion is to stop mirroring the 1839 prefix to GWDG.
- Dirk Goldhahn reported an issue (18.3.2016) with handle resolution alternating between an ok response and a HTTP 400 Bad Request response from the GWDG EPIC api.
 - This issue surfaced after the aforementioned system update at GWDG happened and it was suggested to file a bug report with GWDG. A follow up is needed to check the current status of this issue.
- **AAI:**
 - VC meeting 22.3.16, relatively low turnout
 - <http://clarin.eu/category/clarin-member> requested and received for eduGAIN via DFN-AAI (still needs documentation)
 - DFN-AAI exports CLARIN SPs to eduGAIN by default.
 - Promoted [Attribute Checker](#)
 - TNC16: We likely won't attend
 - Co-operation with Dariah AAI intensifying (Their "taskforce" is now forming.)

8 Status report of the countries

Austria

- Davor Ostojic (Davor.Ostojic@oeaw.ac.at) will take over as representative in SCCTC for Austria.
- We organize a workshop on Tools and Services as part of CLARIN-PLUS "Facilitating the Creation of National Consortia" Series on 2.-3. June.
- Otherwise concentrating on the rework of data management and repositories.

Bulgaria

No report

Czech Republic

- working on CLARIN Plus tasks
 - Robust AAI foundation
 - Human Interface Guidelines (+ pilot implementation)
- [updated the repository system](#) to DSpace 5.5 + other improvements
- Working with Spanish and Italian centres to install and setup our repository system
- working on new/improved web applications and services
 - parsing pipeline for all languages in Universal Dependencies
 - displaying syntactic trees in corpus search tool Kontext
- Preparing K-centre on Treebanking documentation/help/tutorial for PID assignment to dynamic resources (e.g. grammar/treebanks under constant development) with colleagues from Bergen
- preparing new projects that would rely on CLARIN services

Denmark

- Reassessment of B-centre, but waiting for DSA to release new questions to be able to reassess the DSA.
- Busy with internal administration to prepare for Advisory Board Meeting in the Danish digital humanities “umbrella” project DIGHUMLAB.
- Planning national workshops for users.
- Still working on “easier deposit” tutorial.
- Mitch involved in CMDI testing.
- Added CLARIN-DK toolbox to the Language Resources Switchboard.

Dutch Language Union

No report

Estonia

- Experienced some trouble with our servers, had to reorganize them.
- Preparing services for users: adding resources and documentation (corpus query interface Korp and workflow web application Keeleliin).
- Small steps in providing the documents for DSA assessment, checking the new DSA requirements/guidelines

Finland

- Negotiations about funding for the next 5 years underway
- Adaption of a user friendly analysis tool “Mylly” (“The mill”) (originally for bio informatics as “Chipster”)
- REMS Pilot with Clarino still in preparation.
- Metdatacuration: We are cleaning up our Metashare instance and make relations between Metadata, versions, PIDs clearer.

Germany

- CLARIN-D just had it’s quarterly developer meeting in Saarbrücken
- CLARIN-D is preparing for the Forum CA3 2016 in Hamburg. This will be a (German speaking) conference marking the transition from the implementation phase to the use phase for CLARIN-D. The infrastructure and its tools for accessing, analysing and archiving language resource will be presented. <http://www.clarin-d.de/en/current-issues/forum-ca3-2016>
- First centres working on adopting their endpoints to FCS 2.0

Greece

CLARIN-EL is working on preparation issues in order to apply for Type-B certification according to DSA procedures and requirements. In addition, discussion and negotiation procedures regarding funding are in progress.

Italy

- ILC-CLARIN has overcome a blocking issue with IDP registration: IDEM director has accepted to bypass central CNR and we have the approval for ILC-CLARIN Service Provider. Also, a new ad hoc server has been bought, the Dspace is being populated, and ILC-CLARIN will soon qualify as a C-center and be available for harvesting.
- Meanwhile, CLARIN-IT will be soon constitute a consortium. Among the Interested partners, Andrea Abel from EURAC Bozen says that they are ready to become a center and will hire a person to this purpose (more info on this next time).
- We also helped another partner-to-be, University of Siena, to become part of the IDEM identity federation, and thus to be able to use their institutional identifiers to access CLARIN resources and services

Lithuania

- We continue to develop CLARIN-LT repository. It seems that we are stuck with some technical problems. We hope to solve them soon;
- We are finalizing the Lithuanian corpus search interface, which will be soon seen online;
- The Lithuanian agency MOSTA has started expert evaluation of Lithuanian research infrastructures. We have submitted the detailed report on CLARIN-LT activities. It looks like our funding after this year will depend on this evaluation.

The Netherlands

No report

Norway

- We continue to enhance our national repository with new metadata and resources for the norwegian language bank. The technology is up and running. We need to focus somewhat more on the user interface for search in the database.
- The OAI interfaces are up and running. Still, there are only resources without usage limitations (open).
- We are in the launching phase of a service for national PID handling base on EPIC handles.

Poland

- We prepare CLARIN-PL conference titled research infrastructure for social sciences and humanities which will be 25 – 27 April 2016 in Wroclaw.
- Our focus is on next year's finances. At April 18 we (Maciej) will present CLARIN-PL 2.0 project to the Ministry Commission. The application passed the formal evaluation.
- We finishing implementation all web services from our partners to our server.

Portugal

No report

Slovenia

Jozef Mišutka is visiting week April 11-15 in order to upgrade our LINDAT repository to the latest version and look into various small problems we have with the repository.

We also plan to (test) install their KonText web concordancer.

We also have now Nagios for monitoring the repository and other CLARIN.SI services.

We started looking at Grid processing for automatic annotation of large corpora.

Sweden

Språkbanken (the Swedish Language Bank), one of Swe-Clarín's consortium centers, has submitted a certificate request for becoming a B-centre.

United Kingdom

No report

6 Any other business

- Matej (AT) asked if the invitation to the CLARIN PLUS Workshop in Vienna is well received and if confirmations to attend can be send to AT shortly.
- Dieter: The updated agenda is now available and registration has opened for the Centre meeting in May Utrecht, see: <https://www.clarin.eu/event/2016/centre-meeting>

Lunch (10 and 11 May) and dinner (10 May) for this Centre meeting will be provided by CLARIN ERIC. Participants are expected to arrange and pay their own travel and accommodation.