Title	Minutes SCCTC 15 March 2017
Version	1
Author(s)	LST
Date	2017-03-16
Status	Draft
Distribution	Centre Committee
ID	CE-2017-0994

Participants

Matej Durco (AT), Jan Hajic (CZ), Dirk Goldhahn (DE), Lene Offersgaard (DK), Krista Liin (EE), Martin Mathiessen (FI), Harris Papageorgiou (GR), Marc Kemps-Snijders (NL), Marcin Pol (PL), João Silva (PT), ILC4CLARIN (IT), Caspar Jordan (SE), Cyprian Laskowski (SI), Martin Wynne (UK), Dieter Van Uytvanck (CLARIN ERIC, Chair), Linda Stokman (CLARIN ERIC, minutes)

Excused: Kiril Simov (BG),Andrius Utka (LT), Roberts Darģis (LV), Freddy Wetjen (NO)

0   Action points

#	Action	By whom	By when
1	Check if the A-centre logo has been communicated to CZ and double check the procedure in terms of communicating on getting such a status	Linda / Dieter	April 2017
2	Add link that leads to a page of certified centres/ relevant page to the centre logo. Will be on the styleguide and will be published soon ones the comm.officer has looked at it as well.	CLARIN Office	April 2017
3	Set up small author task force to summarize proposals made for archiving software (agenda point 4)	Lene	Next SCCTC meeting

1   Agenda

1. Agenda
2. Approval minutes last meeting & action point status (CE-2017-0959)
3. Centre assessments: short status update
4. Discussion item: how to archive software? (Lene)
5. Reports from the task forces: (taskforce coordinators)
   - CMDI
   - Metadata curation
   - Federated Content Search
   - Persistent Identifiers
   - AAI
6. Status update per country/member (all) -> **please add a snippet for your country to the reporting google doc**)
7. Any other business

2   Approval minutes and action points

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

3   Centre assessments: short status update

CMU has applied for re-assessment (9th round).

For the 8th round, two B-Centre assessments need to be done, people are currently preparing to assess these centres. Hopefully the report is ready by the next SCCTC meeting.

The deadline for the next A-Sercvice assessment is not yet decided, this date will be communicated in the next SCCTC meeting.

4   Discussion item: how to archive software? (Lene)

It is agreed to set up a small author task force to summarize proposals made in the meeting today. We will approve or further discuss them at the next SCCTC meeting. Coordinator of the task force will be Lene. *See action point 3*.

5   Reports from the task forces

CMDI:
- No report

Metadata curation:
- No report

Federated Content Search:

- Doing final review and last additions of FCS 2.0 specification before SCCTC can have it for decision (next meeting).
- Now working on endpoint tester test cases for coverage of FCS 2.0 features when source is available
- Further endpoint migration support
- Documenting Graphical Query builder
- Outreach at different workshops and conferences, currently at DHN 2017.

Persistent Identifiers:
- No report

AAI:
- Nothing to report

6 Status report of the countries

These reports are a *public* extract of the common NCF and SCCTC monthly report for February 2016, which is also available as *internal* document CE-2017-0981.

Austria
- New (DH) projects starting (go!digital 2.0 call) (= new persons being hired)
- Davor Ostojic leaving us of 31.3. => set back for our contributions to VLO, Curation Module (until I find replacement)
- Progress in Parthenos
  - WP5 - mapping of CMDI to CIDOC-PE (see also material on the mapping in https://github.com/acdh-oeaw/parthenos_mapping)
  - WP6 - exploring integration of services into the d4science infrastructure
  Via a first use case NERLiX (NER & Entity Linking) - description
  Discussion on CLARIN concerns in Parthenos yesterday (14.3.)
  Will approach CLARIN secondment partners in Parthenos
- CLARIN Centre-B - Aiming for next re-assessment round (28.4.)

Bulgaria
No written report

Czech Republic
- CLARIN-Plus Workshop on Repositories hosted in Prague
- Additional evaluation documents, asked for by the Ministry for the 2017 LINDAT/CLARIN evaluation
- Intensive work on Universal Dependencies v2.0 release (planned for March) and the UD Shared Task - data preparation, baselines, evaluation data translation and annotation

Denmark
- Fixing issues with SPF/Wayf-login
- Preparing CMDI metadata for lexica resources, updating current MD with PID etc.

Dutch Language Union
No written report

Estonia
- Organizing the workflow on how to handle resources that are being developed (and will take time before they are properly described), but already need some kind of archival.
- Working on getting more direct references to the ResourceProxy fields for our resources so that data links would show up better in VLO.
- Adding services (tools) to our workflow system Keeleliin (Language Line) and introducing it to users.
- Writing/correcting documents (such as the risk analysis plan) necessary for the DSA.
- Plan to participate in 9th assessment round (April 28)

Finland

Done

- Funding secured for 3 years.
- Defining usage of PIDs for updates, accruing corpora (unfortunately only in Finnish yet)
  Key points: PID usage, what is a version, list of preferred DataCite relations

In Progress

- Moving to Github (will be in progress for a while)
- B2SAFE-pilot: Kielipankki willing to replicate to it
    - Seems to be difficult, very slow progress
- Workflow manager "Mylly" in beta-testing, for details see https://www.kielipankki.fi/support/mylly/
- Kielipankki prepares to store ("confidential") sensitive data, possibly we certify according to ISO 27001.

Update on sensitive data

The data is sensitive in nature, eg. videotaped conversations of regular health checks of young children. The goal is to securely distribute the unmodified data for research purposes.

At this point it is unclear though, whether the data is officially classified, options are "Confidential" (leaking the data would inflict "damage" to the stakeholders) or "Restricted" (a leak would "cause harm"). If it is not yet classified, it is also unclear who has the authority to classify it or if it even needs to be classified in this case.

We seem to move in uncharted territory. "Confidential" classified data cannot be distributed via the open internet without making sure all involved systems are certified to deal with confidential data. At this point, not even login via home organisations is certified, let alone the laptops of the researchers.

Conclusions drawn so far:

- Storing and processing the data mentioned above will require high security standards, ISO 27001 for Kielipankki will likely be a prerequisite.
- Strong authentication of potential users will be needed (2-factor)
- We will need to help researchers harden their laptops.

Germany

- Ongoing/revived discussion on details of usage of MIME types (e.g. in WebLicht), e.g.:
    - Preferred solution: standard MIME types + format variants
    - Which format variants are needed?
    - Which additional parameters to use (tokenized, charset…)
- More updates after developer meeting in Saarbrücken (28.03.)
- Currently starting to deploy Kontext (NoSketch frontend with FCS support) at Leipzig centre (https://github.com/czcorpus/kontext)

Greece

No written report

Hungary

No written report

Italy

- Not much to say. Still moving into the last dspace (not yet accomplished)
- Working on Weblicht services based on available NLP tools
- Constantly update our github
- Contacted eudat people (cineca in Italy) to start a B2SAFE procedure for data

Latvia

No written report

Lithuania

No written report

The Netherlands

No written report

Norway

- The Clarino C center of the National Library in Norway has been up and running for some. We provide the handle PID service to the Clarino society and provide access to the norwegian language resources. The C center has been operating flawlessly since the last meeting.
- We are currently upgrading our comedy (metadata editor) due to instabilities. Will be finished later today.
- We are awaiting financial support from the national research council to start the development of the long-term storage service. An answer will be given to us medio 2017.

Poland

- Not much to report. We do d-space update to newest version.
- We did renew expired Certificates.
- Workshops in Lodz at February 3-4, 2017 - Digital tools for analysis of the language in the humanities and social sciences

Portugal

- Nothing to report

Slovenia

- Reorganised our dspace-related git repositories, working model and servers to be in line with the new clarin-dspace collaborative model, and began working on shared issues

Sweden

- SND has sent an application for B centre status and hopes to be considered in assessment round 9.

United Kingdom

- Installation of CLARIN DSpace is underway at the Oxford Text Archive, in collaboration with LINDAT
- A application for B-Centre status is planned after the re-launch of the repository on the DSpace platform

6   Any other business

DVU: The next online SCCTC meeting will be in April and the physical one will be held on 16 and 17 May 2017. The SCCTC meeting will take place in the morning of 16th Tuesday. Venue details and the exact agenda, will follow, Ideas for agenda points or sessions that can be held during these days are welcome. A questionnaire will be distributed shortly.

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