

**Title** Minutes joint Centre and Assessment Committee October 2014  
**Version** 1  
**Author(s)** Dieter Van Uytvanck  
**Date** 2015-01-07  
**Status** Final  
**Distribution** Centre and Assessment committee  
**ID** CE-2014-0457



## 1 Action points

#	Action	By whom	By when
5	Draft metadata curation service specification	MD	Dec 8
9	Distribute updated Central Services list (CE-2014-0398)	DvU	Dec 8
10	Distribute updated A-service criteria (CE-2014-0401)	DvU	Dec 8

## 2 Attendants

Lene Offersgaard (DK+ assessment committee), Martin Mathiessen (FI, observer), Jens Edlund (SE, observer), Marcin Pol (PL), Krista Liin (EE), Mihkel-Mikelis Putrinš (EE, observer), Tomas Krilavicius (LT + assessment committee), Remco van Veenendaal (DLU), Marc Kemps-Snijders (NL), Matej Durco (AT), Jan Hajič (CZ), Bart Jongejan (DK, observer), Freddy Wetjen (NO), Menzo Windhouwer (NL, observer), Twan Goosen (NL, observer), Daan Broeder (assessment committee)

**Dieter:** (opens 14:37) Welcome to the Joint SCCTC and Assessment meeting

**Dieter:** Are there objections to the presence of observers? No objections to observers. Welcome to the new members Sweden and Lithuania.

## 3 Previous minutes

Minutes (CE-2014-0415) are approved.

Agenda point 8 will be revisited.

Some points are overdue, deadlines should have been reset.

## 4 Assessment state

New round: Polish Centre.

CMU TalkBank has indicated that they have their OAI-PMH endpoint ready (with CMDI metadata). Progress made, closer to finalization of assessment procedure.

**Krista:** November will see a restart of DSA assessment of Estonian Centre.

**Marc:** AAI status Huygens: cannot report information about status.

## 5 Draft criteria for A services

**Dieter:** first things first. The criteria for A services ...

**Dieter:** (walks through draft criteria) 1. The need to be B Centre.

**Matej:** this should be uncoupled.

**Daan:** difference with E Centre?

**Dieter:** E Centre has no formal relationship with CLARIN but provides computing resources.

Discussion: does offering A service entail being a B Centre?

**Marc:** drop notion of 'A Centre', define SLAs.

**Matej:** it's nice to have badge 'A Centre', it gives the criteria a label.

**Marc:** limit assessment to specific services, not to Centres.

**Dieter:** SLAs are key. A service without a consortium could be called an E Service.

**Dieter:** One could want to simplify the whole labeling scheme perhaps.

**Marc:** (repeats point about SLAs & specific services). Make quality of services (uptime etc.) explicit, that's all. SLA defines what criteria are there.

**Lene:** (missed).

**Matej:** 'A' would mean level of quality Marc, but I thought they would also indicate importance within infrastructure.

**Daan:** one could call Centres with A services A Centres, straightforward.

**Dieter:** terminology matter.

**Marc:** SLAs define everything.

**Daan:** SLAs are a challenging demand.

**Menzo:** we need a contact point, so A services would then need to be limited to B Centres.

**Dieter:** A services are addressed through ERIC.

**Dieter** question: who uses computer Centres?

**Marc:** We use EPIC handles, from SARA computing Centre on best-effort basis.

**Dieter:** should SARA be listed as E centre.

**Marc:** would ERIC be an A Centre if it is responsible for A service?

**Dieter,** no, A Centres ...

**Daan:** PID handles are not specific to CLARIN, why approach it as A service? Different from CMDI services.

**Dieter:** A services are essential to infra, but not all essentials are A services. Let's start to define A services w.r.t. CLARIN community. Proposed definition: an A Centre provides an A Service to the CLARIN community.

**Marc:** we can call it what we want, but we cannot give guarantees about OpenSKOS f.i. as an 'A service'.

**Daan:** funding is strongly related to explicit SLA for service?

**Dieter:** goal of A service is to gradually upscale availability services over time.

**Marc:** the A service definition for me is an High Availability Service.

**Dieter:** I agree but we should go through list of current services as it needs to be determined whether they are in this category.

**Dieter** criterion 2?

**Marc:** 2 is okay, but what about need of upscaling in terms of higher load (ERIC expansion?)?

**Remco:** We should not make remarks about upscaling, it is logical that we need to cope with that situation.

**Dieter:** 3a.: is it possible to define a description?

**Daan:** may descriptions change?

**Dieter:** no, we do not want to allow limiting descriptions during the course.

**Dieter:** point b. Not controversial.

**Marc:** we should also take into account procedural aspects of changes to SLAs/change management.

**Dieter:** yearly revision.

**Dieter:** c. the proposed guaranteed uptime proposed is not ambitious (97.5%).

**Daan:** penalties if promise broken?

**Dieter:** We'll include that under 4.

**Daan:** what about best-offer for non-crucial services.

**Dieter:** focus on absolute minimum level of service in terms of update percentage.

**Remco:** what about (response) times for issue resolution?

**Dieter:** That's a criterion that is hard to measure f.i. with Nagios. Also number of incidents can be too high.

**Marc:** Centres would need to run system administrator shifts, that's costly. Also international/time zone problems can be challenging for SLAs.

**Martin:** in SLAs one always defines rules and exceptions such as only on weekdays, 9-17 h.

**Dieter:** let's work out the details later, based on best practices.

**Dieter:** criterion 4 is important. We could monitor with Nagios, on time intervals such as every 30 mins. One hand: warnings, the secondly: monitor performance.

**Dieter:** criterion 5: we already got feedback. We have committed to OSS. This criterion fits in the philosophy of CLARIN, also it ensures sustainability/maintainability.

**Daan:** why open source for very general services such as PID?

**Dieter:** we can calculate metrics over the source code, and we can control the availability and not be reliant on provider of software or on our capacity to reimplement the open protocol.

**Daan:** I don't see the need, it could be a very widespread software/protocol with lots of competitive implementations.

**Marc:** agree.

**Dieter:** it's consistent with our policy to demand open source implementations and given the current situation with open source implementations already there is no justification or value of being more permissive.

**Daan:** we could leave this kind of decision with the CAC.

**Krista:** it's related to point 7.

**Marc:** there are many ways to fulfill the criterion.

**Krista:** yes but OSS is advantageous in this regard.

**Marc:** the relation between point 7 and OSS is weak.

**Dieter:** point 6 and 7 are relevant. As for point 6, the documentation should cover more than just how to deploy a VM.

**Marc:** we need a definition of 'migration'.

**Dieter:** we found with CLARIN-D that software is not always very migratable.

**Mihkel-Mikelis:** we should test in other Centre the migratability, plus this would help cooperation/knowledge exchange.

**Dieter:** On to the 8<sup>th</sup> criterion, the CLARIN branding

**Krista:** The A service label could be put next to CLARIN logo.

Criterion cancelled. Lots of detail questions about this.

**Daan:** A service should be registered at Centre Registry.

**Dieter:** point added.

**Bart:** we should also cover information security requirements.

**Marc:** what is the time frame for the final version to come into actual effect?

**Dieter:** at least half a year because the testing of services at other Centres takes time, for instance.

**Dieter:** I would hope can finish the criteria at least before the end of the year.

**Remco:** specify under 3. The bandwidth requirements.

**Dieter:** noted (in template).

## 6 List of central services (potential A-services) ([CE-2014-0398](#))

**Dieter:** we need to connect the Centre Registry to the Nagios monitoring. We need developers. CLARIN ERIC will hire developers next year. But Centres should also look at the list of work to be done and contribute to development.

**Dieter:** let's go through the list of 30 potential A services. VLO?

**Marc:** we should consider usage statistics to decide what has A status.

**Daan:** VLO is not essential.

**Dieter:** is VCR A service?

**Oliver:** I cannot say anything about that right now.

**Dieter:** FCS? That service is too beta still.

**Dieter:** CLARIN IdP and Discovery Service? I believe it is critical.

**Marc:** impact if CLARIN IdP is down?

**Dieter:** homeless IdP users cannot access protected resources. Of Discovery Service,

**Marc:** CLARIN should give us SLA for CLARIN IdP because a lot more than CLARIN resource access depends on it.

**Martin:** (agrees.)

**Dieter:** DS is currently replicated RZG/MPI.

**Dieter:** monitoring is important for accountability.

**Mihkel-Mikelis:** why not use third-party service for monitoring?

**Marc:** the monitoring is low impact, just for internal workflow.

**Mihkel-Mikelis:** no it is essential for all Centres offering services.

**Dieter:** As for the Component Registry, backend is important, frontend not.

**Matej/Dieter:** this is all just a wishlist at this point, do note that.

**Dieter:** **Marc,** is OpenSKOS on offer as A service?

**Marc:** all depends on actual SLA up for negotiation.

**Dieter:** Centre Registry is important, many things dependent on it.

**Dieter:** Trac/SVN? Seems not really critical.

**Menzo:** downtime of Maven also impacts compilation of CLARIN software (services) in the event of failure of such services.

**Dieter:** WebLicht source code is not available, we should inform. Some users may be affected from unavailability.

**Oliver/Dieter:** not really critical.

**Dieter:** WWW, is somewhat important and we do have an SLA. Piwik: not critical. EPIC PID service: PID resolution is an A service, but the rest (PID management) is not.

**Jan:** the current list of services is not a good reflection of 'criticality'. Some may not be A service, but they have heightened level of importance. The value of a service to the community should be distinguished from 'technical (time) criticality'. We are willing to promise A level SLA for other services than those decided on this list.

**Dieter:** the agenda point about supported metadata formats is now postponed to a next meeting due to lack of time.

## 7 Taskforces

**Matej** (About metadata curation taskforce). I have written my findings on the MDCuration Trac page. Anecdotal checks have been done to some success. Automation is still crucial. Controlled vocabulary is the way to go to map various spellings of facet values to canonical one. Value normalization is the main focus of me – this bugs the most people.

**Martin:** (About AAI taskforce). CLARIN's DP CoC requirement has given more credibility to DP CoC. Access and attribute release problems are known to IdPs. We should induce

pressure from users to pressure IdPs into release of attributes to SPF SPs. The UFAL SP Validator is now being used ... Draft. Metadata workflow. SPF should be more reliability. Publishing a corpus electronically has low Return on Investment for researchers.

**Dieter:** SP and DiscoJuice configuration should be taken up with SP operators (check usage of current standards/SAML metadata URLs).

**Dieter:** we should make sure all consortia are taking part in FCS taskforce.

**Menzo:** (about the CMDI taskforce). V 1.2 Toolkit is done, but we need more test cases. Formal spec in natural language is currently being worked on. Of concern is the limited practical participation in the taskforce's effort (taskforce members).

**Dieter:** reconsider taskforce memberships and call for active participants.

**Marc:** does limited participation threaten to stall CMDI 1.2 progress?

**Dieter:** we need more developers/development in general, this is an important issue.

**Marc:** why not approach consortia first instead of Centres?

**Dieter:** we will, during coming National Coordinators Forum.

**Dieter:** PID Taskforce: DataCite DOIs are CLARIN compliant but they come with advantages and disadvantages as compared to Handles.

**Daan:** what about data granularity? What constitutes a 'data set'?

**Dieter:** unsolved challenge.

## 8 The update per member.

**Lene:** our focus is user involvement. We need to promote and interact with researchers. For instance we now stress that we interact with them not solve their problems for them. Bearded Finn: Finland hopes to become member of CLARIN ERIC.

**Freddy:** we did work on CMDI profiles. We have written a paper on our new repository structure. We have launched it for CLARINO. We have not yet completed process to become B Centre, because of a lack of human resources. We hope we will make the next assessment round. We are making a lot of progress.

**Jens:** we are new members. Contractual issues with the research funding agency are still being sorted out.

**Marcin:** we are preparing for CMDI 1.2. We want to use DSpace as our repository system.

**Krista:** in ten days we will start another .... We want to automate handle assignment and also make it user friendly. We are participating in Digital Humanities event.

**Tomas:** we will Tuesday be formally accepted. In about January we may start progress to create a Centre.

**Remco:** we missed the fourth assessment round, we work towards implementing DSpace as repository.

**Marc:** Huygens still in assessment round. Maybe CLARIAH will have effects.

**Jan:** we like that other Centres start to implement DSpace. There is a competitor at play, so we will have a difficult situation to handle with the ministry. We have agreement with USC(?) to take our oral history of Holocaust (Spielberg) as resource into CLARIN. A popular national language instruction material resource could be a K(?) service, but there are complexities. It will have more impact than niche academic resources, which may be a factor to consider.

**Dieter:** as for the CLARIN-D situation, we have a focus on preparation for re-assessment (fifth round). We evaluate effort needed, seems manageable. We work on the VLO and metadata curation issues.