1 Summary

In the following we will report on the twelve B centre evaluation round which was carried out Dec 2018 – April 2019 (6 centres). This evaluation round uses the *CLARIN B Centre Checklist*[1] version 6.

For each centre, the assessment committee will recommend if the centre can be granted/re-granted the label as a CLARIN B-Centre, optional give a general comment, and perhaps list remarks or suggestions for improvements.

All centres have the option to clarify the points mentioned below by short statements to the Chair of the CLARIN Assessment Committee and to the chair of the SCCTC in case that the remarks or corresponding evaluation results are not agreed on.

For this round two centres applied for re-granting:
- LINDAT/CLARIN
- The CLARIN Centre at the University of Copenhagen

Four centres applied for assessment:
- CEDIFOR - Centre for the Digital Foundation of Research in the Humanities, Social, and Educational Sciences
- DCH - Language Archive Cologne, Data Centre for the Humanities
- Eurac Research CLARIN Centre
- PORTULAN CLARIN Research Infrastructure for the Science and Technology of Language

For the following centres **CAC suggest certifying the centres as CLARIN B-Centres** when they have got the CTS (Core Trust Seal) and considered the remarks and needed clarification specified below for each centre:
- CEDIFOR - Centre for the Digital Foundation of Research in the Humanities, Social, and Educational Sciences
- DCH - Language Archive Cologne, Data Centre for the Humanities
- Eurac Research CLARIN Centre
- PORTULAN CLARIN Research Infrastructure for the Science and Technology of Language
- LINDAT/CLARIN
- The CLARIN Centre at the University of Copenhagen

The evaluation of B-Centres will be valid for three years from the date the CTS is achieved, as all centres are in the assessment process of the CTS.

---

[1] Clarin B Centre Checklist 2018-02-07

However, reassessment might be needed before the three years has passed if changes in the requirements in the meanwhile are agreed on within the SCCTC.

2   General Points

The reviewers in the CAC would like to draw attention to the following general issues.

Checks to be ignored when all resources are public available

The reviewers suggest that SCCTC decides which parts of the checklist that can be ignored if a centre only offers open resources and are planning to keep it that way.

If centres are only providing public resources and plan to keep it that way, we suggest that they are not obliged to join the SPF and to implement Authentication/Authorisation login.

Test of access from IdPs

The reviewers cannot always test SP access from both national and non-national IdPs for a centre, but as minimum the test include existence of CLARIN IdP login and the availability of a discovery service for national and non-national IdPs. The assessment committee also wants to thank the AAI taskforce for helping with testing authentication. **However, some centres are not implementing authentication as they only offer publicly available resources.**

Automated checks

Earlier, the SCCTC committee has agreed to aim for more automatic checks of the compliance in the future. CAC has in this round used the metadata curation module[2]. The curation module has been a very useful tool in this round - giving a very useful overview and reporting option for errors. The Centre Registry[3] is still used to give a clear overview of the Service Provider registration with the identity federations.

Update of information material about B-centres and checklist needed

The reviewers see a need for more clear information about the CLARIN concept "B-centre".

Quality of metadata

The reviewers do not make statements about the quality of the metadata or the content of the resources. Only random samples of resources were investigated, and only for the validation of the records, existence of CMDI metadata and PID use.

3   Centre Evaluation

3.1   LINDAT/CLARIN

It is recommended that LINDAT/CLARIN is prolonged as a CLARIN B centre, when it fulfils the following criteria:
- The CoreTrustSeal is re-granted. Please inform the CLARIN Office when you get the CoreTrustSeal.

---

[2] Curation Module, available at https://clarin.oeaw.ac.at/curate/#!Collections
[3] Centre Registry, available at https://centres.clarin.eu/

Other remarks:
- Some problems were noticed logging in using with two identities sharing the same email address. This might give a problem in some usage scenarios and complicates testing. However, this is beyond the current formal requirements and should be discussed by the CLARIN centre committee and/or assessment committee. It should not influence the certification.

3.2 The CLARIN Centre at the University of Copenhagen

It is recommended that The CLARIN Centre at the University of Copenhagen B is re-granted the label as CLARIN B centre, when it fulfils the following criteria:
- The CoreTrustSeal is granted. Please inform the CLARIN Office when you get the CoreTrustSeal.
- The requirement for a SAML Attribute debug script should be met, which is now disabled for security reasons.

Other remarks:
- Some problems were noticed logging in using with two identities sharing the same email address. This might give a problem in some usage scenarios and complicates testing. However, this is beyond the current formal requirements and should be discussed by the CLARIN centre committee and/or assessment committee. It should not influence the certification.

3.3 CEDIFOR - Centre for the Digital Foundation of Research in the Humanities, Social, and Educational Sciences

It is recommended that CEDIFOR is granted the label as CLARIN B centre, when it fulfils the following criteria:
- The CoreTrustSeal is granted. Please inform the CLARIN Office when you get the CoreTrustSeal.
- Although metadata assessment is not carried out by the assessment committee, a description to each resource should be added for these resources: https://vlo.clarin.eu/record?11&count=4&docId=21.11105_47_0000-000B-D4D9-1&fq=collection:CEDIFOR.Corpus&fqType=collection:or&index=2 and https://vlo.clarin.eu/record?12&count=4&docId=21.11105_47_0000-000B-D24B-4&fq=collection:CEDIFOR.Corpus&fqType=collection:or&index=3
- The centre arranges a meeting with the national SCCTC member to have a dialogue about fulfilling the criteria mentioned in this section.

Other remarks:
- All resources provided are publicly available. Since the resource centre is only offering resources without access restrictions, the centre is not maintaining a shibboleth endpoint and therefore has no SAML metadata.

3.4 DCH - Language Archive Cologne, Data Centre for the Humanities

It is recommended that DCH – Language Archive Cologne, Data Centre for the Humanities at https://lac.uni-koeln.de/ is granted the label as CLARIN B centre, when it fulfils the following criterion:
- The CoreTrustSeal is granted. Please inform the CLARIN Office when you get the CoreTrustSeal.

Other remarks:
- There is an issue with the CLARIN IdP implementation that is currently worked on, test will be performed after implementation[4]. Currently, login via SPF can be done via https://ka3.uni-koeln.de/account/index.

3.5   Eurac Research CLARIN Centre

It is recommended that Eurac Research CLARIN centre repository is granted the label as CLARIN B centre, when it fulfils the following criteria:
- The CoreTrustSeal is granted. Please inform the CLARIN Office when you get the CoreTrustSeal.
- When more than the current four resources are available, to demonstrate that the number of resources grows[5].

3.6   PORTULAN CLARIN Research Infrastructure for the Science and Technology of Language

It is recommended that PORTULAN CLARIN repository is granted the label as CLARIN B centre, when it fulfils the following criteria:
- The CoreTrustSeal is granted. Please inform the CLARIN Office when you get the CoreTrustSeal.
- When ready to remove the "beta"-label from the frontpage of the repository.

Other remarks:
- All resources provided are publicly available. Since the resource centre is only offering resources without access restrictions, the centre is not maintaining a shibboleth endpoint and therefore has no SAML metadata.

---

[4] The centre and CLARIN agrees that: The implementation of the SP in Cologne implements the SAML2.0 standard very strictly and has revealed an issue in the implementation of the CLARIN IdP. The situation has been discussed with the CLARIN SP group and they are currently preparing an updated version of the CLARIN IdP.

[5] "The ERCC is currently still building up its collection. It is expected that the size of the repository will keep growing steadily within the next few months when we will slowly add all corpora that exist at the Institute for Applied Linguistics." (in mail to Assessment Committee Feb 25, 2019)