

Title	B Centre Assessment Report June 2018 – round 11
Version	2.1
Author(s)	CAC and Lene Offersgaard
Date	2018-10-03
Status	Approved by SCCTC.
Distribution	Assessed centres, CAC, SCCTC
ID	CE-2018-1247



1 Summary

In the following we will report on the eleventh B centre evaluation round which was carried out March 2018 – June 2018 (9 centres), and March 2018- September 2018 (3 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, optionally 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 10 centres applied for re-certification:

- [ASV Leipzig](#) - Leipzig Corpora Collection
- BAS - [Bayerisches Archiv für Sprachsignale](#)
- BBAW - [Berlin-Brandenburg Academy of Sciences and Humanities](#)
- [CLARIN-PL Language Technology Centre](#)
- EKUT - [Eberhard Karls Universität Tübingen](#)
- HZSK - [Hamburger Zentrum für Sprachkorpora](#)
- IDS - [Institut für Deutsche Sprache](#)
- IMS - [Institut für Maschinelle Sprachverarbeitung](#)
- MPI-PL - [MPI for Psycholinguistics](#)
- UDS - [Universität des Saarlandes](#)

And two centres have not been granted the B-Centre label before:

- CEDIFOR - [Centre for the Digital Foundation of Research in the Humanities, Social, and Educational Sciences](#)
- DCH - [Language Archive Cologne](#), Data Centre for the Humanities

By July 20th 2018 the status was (version 1.1 of this report):

- that the assessment of 3 centres were still in process. The delay of the assessment was caused by lack of time from the Chair of the Assessment Committee and questions to be clarified by the centres.
- that CAC suggests to re-certify 9 centres as CLARIN B-Centres, when they have got the CTS (Core Trust Seal) and considered the minor remarks and needed clarification below.

These nine centres are:

- [ASV Leipzig](#) - Leipzig Corpora Collection
- BAS - Bayerisches Archiv für Sprachsignale

¹ [Clarín B Centre Checklist](#) 2018-02-07

- BBAW - [Berlin-Brandenburg Academy of Sciences and Humanities](#)
- EKUT - [Eberhard Karls Universität Tübingen](#)
- HZSK - [Hamburger Zentrum für Sprachkorpora](#)
- IDS - [Institut für Deutsche Sprache](#)
- IMS - [Institut für Maschinelle Sprachverarbeitung](#)
- MPI-PL - [MPI for Psycholinguistics](#)
- UDS - [Universität des Saarlandes](#)

By October 3rd 2018 the assessment committee the assessment committee also suggests to re-certify the following centre as CLARIN B-Centre, when it has achieved the CTS (Core Trust Seal):

- [CLARIN-PL Language Technology Centre](#)

The evaluation of B-Centres will be valid for three years. However, reassessment can be needed before the three years has passed if changes in the requirements in the meanwhile are agreed on within the SCCTC.

For the remaining two centres in this round, the assessment committee currently suggests continuing the dialogue with the centres in the coming months and to give the SCCTC recommendations later. More details will be included in the information send directly to the centres.

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 is planning to keep it that way.

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”.

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 includes 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

The centre committee has agreed to aim for more automatic checks of the compliance in the future, but this has still not been integrated due to limitations in time. Hopefully, the CAC will soon be able to integrate the metadata curation module², but in this round it was only used manually. The Centre Registry³ still gives a clear overview of the Service Provider registration with the identity federations.

Quality of metadata

² Curation Module, available at <https://clarin.oeaw.ac.at/curate/#!Collections>

³ Centre Registry, available at <https://centres.clarin.eu/>

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 existence of CMDI metadata and PID use.

User help page describing options for sign-up

Some centres guide users about their options and benefits of logging in. We suggest that all centres create a help text/page that informs users of the options for logging in and what resources they might get access to – sometimes depending on the chosen option for login. An example can be seen at: <https://lindat.mff.cuni.cz/en/how-do-i-sign-up>.

3 Centre Evaluation

3.1 ASV - Leipzig Corpora Collection

It is recommended that ASV 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.
- Has considered adding CLARIN logo on <https://wortschatz.uni-leipzig.de/en> or on <http://corpora.informatik.uni-leipzig.de/en> to improve visibility of the connection to CLARIN.
- From <http://clarin.informatik.uni-leipzig.de/repo/> please add links to <https://wortschatz.uni-leipzig.de/en> or <http://corpora.informatik.uni-leipzig.de>
- Using the curation service (https://clarin.oew.ac.at/curate/#!ResultView/collection//CLARIN_OAIProvider_Institute_of_Computer_Science_NLP_Group_University_of_Leipzig): The curation service reports a long list of not-valid records. Please clarify!
- We note that https://vlo.clarin.eu/record?2&docId=hdl_58_11022_47_0000-0000-20DF-1_64_type_61_dataprovider_38_id_61_2&tab=details and https://vlo.clarin.eu/record?6&docId=hdl_58_11022_47_0000-0000-20DE-2_64_type_61_dataprovider_38_id_61_8&tab=details has no license specified.

Other remarks:

- All resources provided are publicly available. Since the resource centre is only offering resources without access restrictions, in coordination with the assessment committee the centre is no longer maintaining a shibboleth endpoint and therefore no SAML metadata.

3.2 BAS - Bayerisches Archiv für Sprachsignale

It is recommended that BAS - Bayerisches Archiv für Sprachsignale 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.

Other remarks:

For item 8: Please consider updating the optional endpoint for FCS to latest version.

3.3 BBAW - Berlin-Brandenburg Academy of Sciences and Humanities

It is recommended that BBAW - [Berlin-Brandenburg Academy of Sciences and Humanities](#) 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.

3.4 EKUT - Eberhard Karls Universität Tübingen

It is recommended that EKUT - Eberhard Karls Universität Tübingen repository is re-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.
- Please update certificate for <https://weblicht.sfs.uni-tuebingen.de/Tundra/>,

Other remarks:

- Please clarify options to enable https for <http://www.sfs.uni-tuebingen.de/en/ascl/resources/corpora.html> and www.sfs.uni-tuebingen.de/GermaNet/

3.5 HZSK - Hamburger Zentrum für Sprachkorpora

It is recommended that HZSK - Hamburger Zentrum für Sprachkorpora repository is re-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.
- Please clarify if the link to LINDAT is supposed to be an alternative for shib_test.pl or the implemented solution for federated login.
- In some cases tool metadata does not seem to contain a PID: "ANNIS at HZSK", please clarify.
- Please clarify that content negotiation from terminal can be done.

Other remarks:

- The curation module https://curate.acdh.oeaw.ac.at/#!/ResultView/collection//HZSK_Repository evaluate the resources provided from HZSK, please check the invalid records.

3.6 IDS - Institut für Deutsche Sprache

It is recommended that IDS - Institut für Deutsche Sprache repository is re-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.
- For item 1.b: Please consider adding a CLARIN logo with a link to clarin.eu to <https://repos.ids-mannheim.de/> and <https://clarin.ids-mannheim.de/>.

3.7 IMS - Institut für Maschinelle Sprachverarbeitung

It is recommended that IMS - [Institut für Maschinelle Sprachverarbeitung](#) repository is re-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.
- Using the CLARIN curation module (https://clarin.oeaw.ac.at/curate/#!/ResultView/collection//IMS_Repository) one item is labeled: Invalid record: oai_ims_uni_stuttgart_de_clarind_ims_38. Please clarify if this item can be updated.

3.8 MPI-PL - MPI for Psycholinguistics

It is recommended that MPI-PL - MPI for Psycholinguistics repository is re-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:

- For item 1.b: Please consider adding a CLARIN logo with a link to clarin.eu to <https://archive.mpi.nl/>

3.9 UDS - Universität des Saarlandes

It is recommended that UDS - Universität des Saarlandes repository 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:

- All resources provided are publicly available. Since the resource centre is only offering resources without access restrictions, in coordination with the assessment committee the centre is not maintaining a shibboleth endpoint and therefore no SAML metadata.
- The center statement points explicitly also to <http://corpora.clarin-d.uni-saarland.de/cqpweb/> which has a non-federated login. The cqpweb server is third-party software (free software developed outside the CLARIN context) that does not provide a Shibboleth login option. Ideally, CAC would suggest that federated identity management is added to cqpweb, but it accepts that the third-party product will not be changed.

3.10 CLARIN-PL Language Technology Centre

It is recommended that CLARIN-PL Language Technology Centre repository 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.

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

DCH - Language Archive Cologne, Data Centre for the Humanities has informed the assessment committee that some of the issues could not be resolved until the end of September. We suggest that the assessment committee continue the dialogue in the coming months and give the SCCTC recommendations later. More details will be sent directly to the centre.

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

We suggest continuing the dialogue with CEDIFOR in the coming months and give the SCCTC recommendations later. More details will be sent directly to the centre.