

# Core Trust Seal, a CLARIN perspective

---

Daan Broeder – CLARIN ERIC

Building trustworthy repositories – DARIAH Virtual Event 2020

October 14 2020

# CLARIN - Common Language Resources and Technology Infrastructure

CLARIN



- 🌀 Create and maintain an infrastructure to support the **sharing, use and sustainability of language data and tools for research** in the humanities and social sciences.
- 🌀 Main vision to make digital language resources and tools from all over Europe and beyond accessible through a single sign-on online environment for the support of researchers in the humanities and social sciences
- 🌀 CLARIN is a mature infrastructure coordinated by the CLARIN ERIC established in 2012

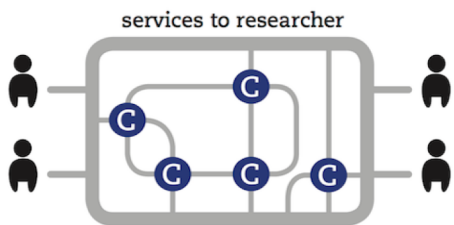
# CLARIN centers



- 🌀 CLARIN's distributed infrastructure is made out of **centres**
- 🌀 With a backbone provided by CLARIN-B centres
  - 🌀 **technical centres, providing resources and/or services**
  - 🌀 often a university or an academic institute, offer the scientific community access to resources and services on a sustainable basis.
- 🌀 **Strict criteria to become a CLARIN B-Centre**
  - 🌀 B-centre requirements target a stable technical and institutional foundation
  - 🌀 CLARIN Centre Assessment Committee checks during assessment procedure
  - 🌀 technical coordination, including discussing the centre requirements is done by CLARIN Centre Committee
- 🌀 Other types are CLARIN-C and -K centres

# CLARIN Centres

- 🌀 24 certified B centres
- 🌀 20 C-centres
- 🌀 17 K-centres



# Why certification?

- 🌀 Improve quality of the center contributions to CLARIN infrastructure
  - 🌀 of management & technical staff, of implementations: repository, access systems, contacts & communication channels, continuity
- 🌀 Decrease interoperability problems
- 🌀 Prove the contributions from national projects to EU infrastructure
- 🌀 Making internal processes explicit also useful for centres themselves
- 🌀 Improved service quality also for local users
  
- 🌀 Already CLARIN first centre assessment procedures since 2013 including a requirement for external assessment via DSA

# CLARIN B-centre certification

- 🌀 General requirements: centre mission, domain, funding, explicit policies
- 🌀 Technical requirements: data management
- 🌀 Many general requirements can be provided by external assessment as CTS
- 🌀 CTS is not a perfect match for CLARIN requirements but definitely worthwhile
  - 🌀 Outsourcing of testing procedure
  - 🌀 High level of prestige that CTS offers to centres and infrastructure.
  - 🌀 CLARIN funds the CTS certification fee and synchronizes its certification cycle with that of CTS (3yrs)
- 🌀 CLARIN-B centre assessment procedure also maps important CLARIN requirements also specifically to similar question in the CTS question list

# CLARIN-B centre requirements I

(2019-09-30)

- 🌀 Proper and clearly specified repository system (evident from a CTS certification)
- 🌀 Centre repository context, mission and scope should be made explicit (CTS R0, R1)
- 🌀 Recommendation letter from national CLARIN coordinator
- 🌀 Reference to CLARIN e.g. website, logo and link to national CLARIN project
- 🌀 Description of funding situation and expectations also wrt future centre activities (CTS R3 is part of this)
- 🌀 Explicit statements wrt offered CLARIN compliant resources and services
- 🌀 Explicit statements about their data & services target audience, IPR aspects and licenses management (CTS: R2)
- 🌀 Explicit statements wrt privacy policy



# CLARIN-B centre requirements II

(2019-09-30)

- 🌀 Registration CLARIN Centre Registry
- 🌀 Human access to the resources & services
- 🌀 Correct technical implementations wrt data management for interoperability between CLARIN centres and central services
  - 🌀 OAI PMH endpoint
  - 🌀 CMDI for LR & LT metadata, including specific content negation for MD
  - 🌀 Use of PIDs for resource metadata and resources
  - 🌀 Membership CLARIN SPF and national IDF (AAI)
  - 🌀 Implementation Geant DP-CoC (AAI)
- 🌀 Security Guidelines:
  - 🌀 Accepted server certificates



# CLARIN and CTS, observations from the CLARIN CAC

- 🌀 CTS very much about **LTP**, CLARIN is very much about **Access**
  - 🌀 CTS is vague about domain and center contributions, CLARIN needs explicit statements on resource contributions
  - 🌀 CTS considered overkill for ao. process/management/role descriptions
  - 🌀 Currently CTS takes 6M, seems overly long
  - 🌀 CTS terminology considered difficult (OAIS)
- 🌀 Conclusion CTS is very useful but suggestion to enable partial CTS, adding the “Access” topic and certification of services,