

<b>Title</b>	CLARIN Committees reports 2017
<b>Version</b>	1.0
<b>Author(s)</b>	TDB
<b>Date</b>	TDB October 2017
<b>Status</b>	Draft
<b>Distribution</b>	BoD, NCF, GA
<b>ID</b>	CE-2017-1109



1. Committee name: CLARIN Standards Committee
2. Chair (and reporter, if different): Piotr Bański
3. Web-Link: <https://www.clarin.eu/governance/standards-committee>
4. Members (Name Surname (Affiliation, Country)): see WebLink
- 5. Goals/mission (from bylaws)**

The main responsibility of the Standards Committee is to advise the Board of Directors on the adoption of standards to be supported by CLARIN ERIC. Its main tasks are:

- to collect, consolidate and prepare for publication in a single place the findings and recommendations related to standards emerging from the CLARIN preparatory phase project
- to maintain the set of standards supported by CLARIN and adapt them to new developments within or outside CLARIN
- to develop and implement procedures for the discussion and adoption of new recommendations for standards
- to ensure harmonisation of standards between CLARIN ERIC and related initiatives, such as (but not restricted to) the META project
- to ensure communication with international standards bodies such as (but not restricted to) ISO
- to publish and promote the standards supported by CLARIN
- to advise the Board of Directors in all matters related to standards

## 6. Major activities, decisions and achievements

A major achievement for the 2016-2017 cycle was establishing a liaison between CLARIN-ERIC and ISO TC37SC4 "Language Resource Management": the ground was laid by Andreas Witt at the ISO Conference in Copenhagen in 2016, documentation was prepared by Piotr Bański, and the formalities were taken over by Sebastian Drude and the BoD. The liaison is in effect since the ISO Conference in Vienna (June 2017), with Andreas Witt as the Liaison Officer.

GitHub space was created and prepared for the CSC (<https://github.com/clarin-eric/standards>). The code for the CLARIN Standards Information System has been cleaned and deposited there. The accompanying documentation has been formatted, reviewed and modified (<https://github.com/clarin-eric/standards/wiki/CLARIN-Standards-Information-System>).

Much bottom-up work on surveying the standards that are actually accepted, used and exported by CLARIN centres has been done by Hanna Hedeland (member since October 2017). This work is the foundation for creating a *unified* report/guidance on standards (this overwhelming item as whole was put on the agenda already in 2016 in Aix; much of work done in the 2016-2017 cycle has addressed this endeavour, bit by bit).

Claus Povlsen has co-authored the CLARIN+ report “Adoption and Implementation of Standards” (D5.3-7).

## **7. Coordination with other institutions/research infrastructures/projects**

DARIAH:

- Charly Mörth is deputy head of the DARIAH-EU National Coordinators Committee and member of DARIAH’s Senior Management Team

ISO TC37 SC4:

- Daan Broeder is a member of Working Group 1 “Basic descriptors and mechanisms for language resources” and a project leader of the published standard “Component metadata infrastructure - Part 1: General model (CMDI-1)”
- Maria Gavrilidou is a member of WG 1 and a co-leader of the project “Component metadata infrastructure (CMDI) – Part 3: META-SHARE metadata model”
- Andreas Witt is the convener of Working Group 6 “Linguistic Annotation”, a co-leader of the project “Corpus Query *Lingua Franca*” in WG 6, and a member of the other working groups
- Piotr Bański is a member of WG 6 and WG 4 “Lexical resources”, and a co-leader of the project “Corpus Query *Lingua Franca*” in WG 6
- Fahad Khan (member since October 2017) is a project co-leader in WG 4

Parthenos:

- Piotr Bański is a member of the “Textual processing and edition” expert group in Work Package 4 “Standardization”

TEI:

- Piotr Bański and Andreas Witt are convenors of the Special Interest Group “TEI for Linguists” (LingSIG)

## **8. Outstanding themes, plans for 2018**

The major initiatives and actions planned for 2018 are as follows:

- move the CLARIN Standards Information System from its original location in IDS Mannheim to a clarin.eu subdomain
- effect a warning on all (or as many as possible) current “lists of standards” that may be understood upon first contact as containing the current CLARIN standards recommendations as outdated, and add a redirect to the system under development

(initially a copy of the working stylesheet(s), eventually the Standards Information System)

- propose to the BoD a short questionnaire on areas of specialization of CLARIN researchers (ideally filled out as part of the user profile in the CLARIN account), so that the CSC can turn to individual CLARIN researchers for help in matters requiring deep expertise in solving standards-related issues
- propose to the BoD a new facility for regulating the assignment of XML namespaces based on the “clarin.eu” domain and, if accepted, process the individual requests
- liaise with the Centres Committee on the implementation of the above points
- continue the individual members’ involvement in standards bodies and report on that to the BoD and the NCF
- continue work on the running topics of the CSC (interoperability in CLARIN: general wide-scope issues and low-level technical issues such as the unification of MIME-type use across centres; reacting to individual standards-related queries from CLARIN researchers).