

CLARIN Standards Committee Meeting

Pisa, 8 October 2018

Participants present ('+')

+	Piotr Bański	Germany	
+	Karlheinz Mörrh	Austria	
+	Ineke Schuurman	Belgium	
	Claus Povlsen	Denmark	
+	Neeme Kahusk	Estonia	
+	Jussi Piitulainen	Finland	
	Francesca Frontini	France	
+	Hanna Hedeland	Germany	
+	Andreas Witt	Germany	
	Maria Gavrilidou	Greece	
+ -	Penny Labropoulou	Greece (apologies but I need to be in CLIC)	
+ +	Fahad Khan	Italy	
	Peteris Paikens	Latvia	
+	Daan Broeder	Netherlands	
+	Jan Odijk	Netherlands	
+	Menzo Windhouwer	Netherlands	
	Koenraad De Smedt	Norway	
+	Tomaž Erjavec	Slovenia	
+	Christian Thomas	Germany	(Susanne Haaf standing in)
+			

Agenda

1. Opening, potential round of introductions
2. Review of last year's conclusions to see to what extent they have been followed
3. Points requested by the Secretariat (reporting for the GA):
 - 3.1 A roundup of major activities, decisions and achievements in the past year
 - 3.2 New/updated information on coordination with other institutions/research infrastructures/projects
4. Review of the proposed standards categorization from the white paper on unified CLARIN recommendations for the use of standards (circulated in May 2018)
5. A glance at Hanna's and Piotr's Bazaar presentation -- we count on the Members' active support
6. Plans for 2019 (individual plans and CSC plans to be submitted in a report to the GA)
7. Other issues

- 7.1 The CSC's stance on CLARIN's membership in the TEI consortium
- 7.2 Jan Odijk's statement
- 8. Conclusion of the meeting

Links to share

1. <https://office.clarin.eu/v/CE-2017-1109-Committees-reports-2017.pdf> (last year's CSC report)
2. <https://docs.google.com/document/d/1IQMCowMgKLAXrK-ZoDrM7kgjXgiHBcASKWfVMO2s3qY/edit> (fields of expertise)
3. <https://docs.google.com/document/d/1uFw8VSohE3H24LE-mkUTbsLayTWQr-NHkHVCvekVyG4/edit> (proposal for namespaces)
4. <https://standards.clarin.eu/sis/> (Standards Information System in the clarin.eu domain)
5. <http://www.meta-share.org/ontologies/omtd-share-ontology.owl/documentation/index-en.html> (OpenMinTeD)
6. <https://textencodinginitiativeconsortium.wildapricot.org/Join/benefits> (TEI membership benefits)
7. Standards whitepaper:
 - a. Documentation formats
https://docs.google.com/document/d/1PARY_j7nAfpd0Zkl_6D2471Pg36WIMhS2OUzs4Hsv_M/edit?usp=sharing
 - b. Text content formats
https://docs.google.com/document/d/1Biaqz0rIEfbHZnlAkH6jk91W_dR044_KbTF0m5YxJeU/edit?usp=sharing
 - c. Metadata formats (table mostly blank)
https://docs.google.com/document/d/18Ra_rNWWMAVyql19GhC3H23uPConXtleG7yoWncsXpg/edit?usp=sharing

Notes / minutes

Agenda point 2

...

Points discussed in another committee (brought up here by Jan Odijk) as being goals for CSC, copied for last year or so, to be made more specific/realistic:

1. *Start migration from the current lists of standards to a new system, eventually the Standards Information System*
2. *Make an overview, by means of questionnaire, of areas of*

specialization of CLARIN researchers, so as to provide help to CLARIN researchers in matters requiring expertise in standards

3. Implement a new facility for regulating the assignment of XML namespaces based on the clarin.eu domain and, if accepted, process the individual requests

4. Liaise with the Centre Committee on the implementation of the above points

5. Continue the individual members' involvement in standards bodies and report on that to the BoD and the NCF

6. Continue work on interoperability, e.g. , the unification of MIME - type use across centres

7. Reply to individual standards - related queries from CLARIN researchers

CQP mentioned (by Jan Odijk) as highly relevant especially for federated search (but missing from some list of standards)

Making overviews of standards belongs to standards committee rather than interoperability committee (back and forth between Jan and Daan) [suppose Piotr's white paper on text standards is a good start]

(Agenda point 7.2) Interoperability committee -- two tasks:

- Prepare website containing recommendations for developers, centres, national coordinators on interoperability issues
- First special focus on parliamentary data (there will be a workshop; goal is to develop a proposed format for parliamentary texts which should be applied by projects working in this field)

[Was there a more or less specific date when feedback from standards committee would be desired? This November?]

Expertise questionnaire (issue 2 of “last year’s list”)

TODO: Each of us fill out these items below within 2 weeks

1. Name
2. Name of the CLARIN Consortium / Consortia
3. Institutional affiliation(s)
4. Contact e-mail address
5. Primary field(s) of expertise
6. Secondary field(s) of expertise
7. Permission to be contacted by the CSC if a need arises (??)
8. Comments ...?

Spreadsheet to follow

List of subject areas (for fields 5 and 6 in the questionnaire):

<https://erc.europa.eu/sites/default/files/document/file/erc%20peer%20review%20evaluation%20panels.pdf>

→ Issue 3 of “last year’s list” (XML-namespaces) was done last year.

Agenda point 4

Re “white paper” on text formats, distinction between “text as a format” and “text as a genre” needs made [I thought the paper is clearly about the latter but maybe not, then]

Piotr will set up the paper so that others can comment in Google docs (when logged in). Others please comment.

Discussion about calling some formats “deprecated” - not considered a successful label? What do we want to say instead? (“current”, “outdated” are ok; “emerging” must not be “deprecated” - maybe “under assessment”?) (There was a separate category for “discouraged” so maybe there is no need to deprecate things in addition to that.)

→ categories should reflect the steps of the standards lifecycle

We should focus on getting this document into a public form. Do the work. Somehow. And find out which centers support different formats and what tooling they can provide. Maybe the latter task is for interoperability committee?

Further workflow:

- People comment on document
- Regular VCs to discuss comments → have regular meetings to keep the work going

Agenda point 5

Piotr and Hanna will take part in the CLARIN Bazaar with the topic of asking B centre representatives about what standards they use/accept/deliver → people should be naming the standards or contact persons to address with an e-mail asking for information on this; Standards Committee members are asked to support this initiative by coming by the stand themselves or sending a delegate

Insecurity if the outcome will be an issue for the Standards or the Interoperability Committee.