

Call for Participation in *CLARIN Workshop (type I) -- Towards Interoperability of Lexico-Semantic Resources*

The goal of the workshop is to improve the interoperability, usability and ease of access of CLARIN lexico-semantic resources. To achieve this goal, we will work on the design of CLARIN open virtual lexical platform facilitating better access to lexical resources and research web applications by Humanities & Social Sciences (H&SS) users. The workshop will have a discussion and group-work based character. Most of the time is reserved for chaired discussions in small topic-related subgroups, concluded with reports presented to the general forum by their chairs during the summing-up session.

Researchers interested in the participation in the workshop should send a letter to the following address: clarin-pl@pwr.edu.pl -- including a description of interest:

1. Short description of the background of the applicant, especially related to the workshop goal and topics. If an applicant comes from an institution which is a member of CLARIN, this should be clearly expressed.
2. Description of the motivation for the participation in the workshop.

The description should be approximately 500 words.

The deadline for submission is December 5th 2016. Decisions concerning acceptance will be sent by December 8th 2016.

The workshop is sponsored by CLARIN. Accommodation during the workshop (including one night in a university hostel) is covered. Partial reimbursement of travel costs up to 200 Euros will be offered to all participants.

The maximum capacity is limited to 25 persons who will be selected by PC on the basis of the delivered descriptions of interest. Priority will be given to participants from the CLARIN members (www.clarin.eu).

Workshop Short Description

Name workshop	CLARIN Workshop (type I) -- Towards Interoperability of Lexico-Semantic Resources
Type	I: Thematic workshop (no follow-up implementation project envisaged)
Programming Committee	Maciej Piasecki, Wrocław University of Science and Technology, maciej.piasecki@pwr.edu.pl Erhard Hinrichs, Eberhard-Karls-Universität Tübingen Kadri Vider, University of Tartu, CELR
Location	CELR, University of Tartu, Tartu, Estonia
Dates	31 st January (12 a.m. opening) - 1 st February 2017 (4 p.m. closing)
Summary	The goal of the workshop is to improve the interoperability, usability and ease of access of CLARIN lexico-semantic resources in order to build a CLARIN open virtual lexical platform facilitating better access to lexical resources by Humanities & Social Sciences (H&SS) users and research web applications. The topics will include: <ul style="list-style-type: none">• the choice of a few common formats for CLARIN lexico-semantic

	<p>resources (LSR);</p> <ul style="list-style-type: none"> • the design of a system of web services for accessing CLARIN LRSs, • the common virtual platform for linking, displaying and browsing LRSs; • gathering requirements from H&SS related to LRSs and the platform. <p>The outcome will be an integrated report with an action plan that will be offered to the wider CLARIN community for further discussion, testing and refinement. The workshop is planned for 3 invited experts and up to 25 participants in total.</p>
--	--

Full Description

Workshop goal

The main goal of the workshop is to initiate the works on the improvement of interoperability, usability and ease of access of CLARIN LRSs for the needs of a their better visibility for H&SS users and their enhanced utilization in research applications. The key idea is to coordinate development of a system of web services for accessing LRSs and a common virtual lexical platform built on the top of them. The platform is intended to be an open generic solution that will allow for effective linking, displaying and browsing of the rich variety of data included in CLARIN LRSs. One of the functions of the platform will be a kind of federated search for LRSs. The platform will be an open solution with potentially many installations and many web application based on them.

The main topics will include:

1. the choice of a few common formats for CLARIN lexico-semantic resources (LSR);
2. the design of a system of web services for accessing CLARIN LRSs,
3. the common virtual platform for linking, displaying and browsing LRSs;
4. gathering requirements from H&SS related to LRSs and the platform.

The outcome of the workshop will be an integrated report with an action plan for each of the key topics discussed in the working groups. It will be next offered to the wider CLARIN community for further discussion, testing and refinement. The first, limited set of standards and requirements will be selected for further work. As a follow up we will also prepare a joint proposal of CLARIN type II workshop devoted to the construction of a first set of web services for accessing selected LRSs and a first prototype of the platform. The platform development is planned to be a collaborative work, but CLARIN-PL is prepared for serve as the main developer due to a similar goal in its action plan.

Link to CLARIN strategic priorities

The topics of the workshop are closely related to several CLARIN strategic pillars (cf. CE-2015-0657), namely:

- *integration of data* – the common format and strategy of integration will form the basis for the interoperable multifaceted resources and will thus enable “the development of tools that allow for mapping between different standards”
- *integration of services* – the common platform and system of web services will simplify providing multilingual web services for semantic annotation, exploration, exploitation, enhancement, analysis, manipulation and visualization of language data; support for mono- and multilingual processing workflows,
- *ease of access* – the common platform and system of web services will simplify access to different types of lexico-semantic resources. They will will enhance multilingual browsing and visualization services.
- *crossing borders* – the common format, platform and strategy of integration will facilitate crossing borders between languages, countries, and infrastructures; “international collaboration both at the RI and at the thematic level”.

Profile of proposing team

Maciej Piasecki is an Associated Professor at Wrocław University of Science and Technology and a coordinator of the G4.19 Computational Linguistics and Language Technology Research Group. He holds PhD in Computer Science for work on Natural Language Processing. Maciej has been the leader of the Polish wordnet project since its beginning in 2005 till now. Maciej’s main research topics are: extraction of the lexico-semantic knowledge from corpora, semi-automatic wordnet expansion, Distributional Semantics and relational lexical semantics. Maciej has been also working on morpho-syntactic processing of Polish (co-author of the first Polish tagger), Information Extraction, formal semantics and Machine Translation.

Erhard Hinrichs is a Full Professor (Ordinarius) for General and Computational Linguistics, Eberhard-Karls-Universität Tübingen, Seminar für Sprachwissenschaft, Tübingen, Germany, 1991 – present. He obtained Ph.D., Linguistics from The Ohio State University, Columbus, Ohio, USA. August 1985. He has been the leader of the GermaNet project since its beginning till now. Erhard Hinrichs is CLARIN National Coordinator of Germany.

Kadri Vider is a researcher of Language Technology at Institute of Computer Sciences in University of Tartu. She managed studies of Estonian Wordnet till 2007 when she started to work in Estonian Ministry of Education and Research arranging Estonian membership in CLARIN infrastructure also. Since coming back into University of Tartu in 2011 she manages Center of Estonian Language Resources - a research infrastructure consortium executing obligations of membership of Estonia in CLARIN ERIC. Kadri Vider is CLARIN National Coordinator of Estonia.

Indication of profile(s) and number of envisaged participants

We would like to bring together key wordnet developers from CLARIN, researches with experience in wordnet applications, as well as selected experts from the area of linking wordnets and lexical semantics technology from outside CLARIN. The external experts are meant to provide reference points facilitating the discussions during the workshop.

Participants from CLARIN will be selected on the basis of descriptions of interest collected in an open call. However, we will work on obtaining good coverage with respect to CLARIN wordnets and the workshop topics. We aim at 2 invited plenary talks. The rest of places will be open to moderated discussions and collaboration in working groups. We expect up to 25 participants for 2 days.

Workshop date(s), location and timetable

The workshop will be held on 31th Jan.

planned for the second half of January 2017 and will be hosted by the Institute of Computer Science, University of Tartu, Estonia. Its exact date will be set on the basis of a questionnaire distributed among the key participants. The workshop will last for 2 days starting with lunch at 12 a.m. on 1st day and finishing late afternoon 4 p.m. on 2nd day.

Provisional agenda

The core of the workshop will be chaired parallel working group meetings, each devoted to one of the key topics. They will start with short presentations of (selected) speakers, addressing questions included in the CfP. Presentations will be followed by extensive discussions. Their results will be consolidated into reports and action plans. The working group meetings will be preceded by 1-2 invited plenary talks focused on the experience related to the workshop topics. The workshop will close with the plenary session during which chairs of all the working groups will present their reports and action plans that will be subject to a general discussion. It will be next offered to the wider CLARIN community for further discussion, testing and refinement.

Program Committee:

Maciej Piasecki
Erhard Hinrichs
Kadri Vider

Invited Experts (confirmed so far):

Francis Bond - author of the Open Multilingual WordNet
Christiane Fellbaum - Princeton WordNet