



CLARIN-NL Work Plan 2015

1 Introduction

This document provides a description of the activities planned in CLARIN-NL for 2015. Actual budget amounts for the various activities can be found in the separate document with the budget details, and a separate document with the 2014 budget is available.

The current plan has been approved by the Board via e-mail on April 19, 2015.

2 Planned Activities

Originally, CLARIN-NL was targeted to finish by April 1, 2015. However, in February 2015, it turned out there was still a lot of money to spend, and a request was made to NWO for an extension of the project until the end of 2015. This request was awarded.

The main reasons why so much money remained are as follows:

1. A number of activities planned for 2014 were not started up because of a lack of human resources;
2. Very many subprojects spent less, sometimes significantly less, than originally budgeted;
3. A part of the costs made were booked on CLARIN-NL but concerned work for CLARIAH-SEED and were charged to CLARIAH-SEED

Since new activities had to start up as soon as possible, there was no opportunity to consult the whole field. Central principles for new activities in the plan for 2015 are:

- New activities must build upon results yielded in an earlier phase of CLARIN-NL or finalize unfinished tasks.
- New activities must take place in 2015 and yield finished results ultimately by Dec 31, 2015 (no extensions are possible)
- The CLARIN-NL EB must have confidence that persons to carry out the work are available.

Furthermore, general criteria of quality and the expected impact for the CLARIN infrastructure are taken into account.

The CLARIN-NL EB has, based on its knowledge and expertise, proposed some concrete ideas for subprojects and invited relevant researchers/software developers to



CLARIN-NL Work Plan 2015

make a proposal for these. In addition, a number of researchers have made a proposal by themselves. All these proposals have been evaluated by the EB.¹

2.1 Technical Implementation

- The Max Planck Institute has decided to reduce its activities for the TLA significantly. Though the data stored there will be safe, it is unclear how easy access to them will be in the future and to what extent access will be provided in a CLARIN-compatible manner. In addition, it is unclear whether new data can be stored there. With this, also the Netherlands loses an important centre for the storage of data in a CLARIN-compliant manner. Many so-called A-services are being or already have been discontinued by MPI and have to be transferred to other centres. Meertens Institute has indicated, in the framework of a continuation of the TLA activities without MPI, that it is prepared to ingest new data. However, in order to do this in an efficient manner, a data deposition service has to be set up. A pilot project for this will be carried out in CLARIN-NL. Project name: *MeertensDepositionService*. Involved party: MI (Marc Kemps-Snijders). 40k euro, Budget Id: 226.
- Daan Broeder (MI) continues his work for the CLARIN-NL portal, and for preparing the transition to CLARIAH until April 1, 2015. 7.5k euro, Budget Id: 227
- Improvements on the metadata search engine have been carried out. Involved party: MI (Matthijs Brouwer). 10k euro, Budget Id: 206.

2.2 Tools and Data

- PaQu, which was started up in 2014, originally consisted of two phases, of which only the first phase was financed. However, much of what was planned for phase 2 was already carried out in phase 1. What remains of phase 2 is (1) implementation of federated login, and (2) transfer to a CLARIN centre (in casu INL). These activities are now financed. Involved parties: RUG (Peter Kleiweg) and INL (10k euro, Budget Id: 220)
- For the PILNAR application federated login is being implemented. 5k euro, Budget Id: 222. Involved party: MI (Marc Kemps-Snijders).
- In various CLARIN-NL projects work has been carried out on the curation of dialect dictionaries (e.g. COAVA, Data Curation Service). A project has been defined (CARE: *Curatie en integratie Regionale woordenboeken*) to do this for the remaining (parts of) the dialect dictionaries. There is close collaboration with Flanders and we hope that financing and executing this

¹ The proposals for the projects *Subtiel* (Van den Heuvel) and *Gekaapte Brieven* (Pieters/Bennis) have been rejected after evaluation, because they cannot be seen as building on earlier activities or results of CLARIN-NL, and because the impact of these proposals for the CLARIN infrastructure was judged lower than that of other proposals that have been awarded



CLARIN-NL Work Plan 2015

- project can contribute to financing by Hercules for similar activities in Flanders (using the same procedures and formats). Involved party: RUN (Nicoline van der Sijs / Roeland van Hout); (53k euro, Budget Id: 209).
- A lot of work has been done in CLARIN-NL on providing better access to linguistically annotated corpora, *inter alia* in the projects OpenSONAR, CLARIN-NL LASSY, PaQU, AutoSearch, and (when cooperating with Flanders) GrETEL. The plan is to continue this in 2015 in three projects:
 - A pilot project for linking grammatical terms, examples and descriptions of constructions in TaalPortaal to query's in corpus query interfaces such as PaQu, OpenSONAR and GrETEL. Project name: *TaalPortaal links to Corpora* (TPC). Involved Parties: MI (Ton van der Wouden), INL, UU, RUG. 58k euro, Budget Id: 211.
 - Make a web application for the desktop corpus query tool *CorpusStudio* developed by Erwin Komen, which is already in used by renowned linguists, and which we are eager to incorporate in the CLARIN infrastructure. Project name: *CorpusStudio-Webversion*. Involved Party: MI (Erwin Komen). 37k euro, Budget Id: 210.
 - *OpenCGN*. A version of OpenSONAR, but for CGN and with extensions to deal with the audio files. Involved Party: RUN (Henk van den Heuvel). 40k euro, Budget Id: 224.
 - Co-ordinator Semantic Interoperability: Ineke Schuurman (UU) has continued her activities as coordinator semantic interoperability until April 1, 2015.² 3,4k euro, Budget Id: 205.

2.3 Education and Awareness

- The movies made earlier have turned out to be very successful for dissemination purposes. We plan to make 3 additional films, this time focused on the results achieved and how researcher scan use these. 15k euro, Budget Id: 223. Co-ordinator: Arjan van Hessen
- Budget has been reserved for supporting events, *inter alia* THATCAMP (held in January, 2k euro, budget id: 216), workshop SPARQL on STCN (Els Stronks, April 13, 2015; 500 euro, Budget Id: 218), and 5k euro for other events (Budget Id: 215).
- Website: We keep the website in the air in 2015 and as long after that as possible. We are negotiating the maintenance contract and hope to reduce the costs (14,5k euro, Budget Id: 214)
- Final Event. 30k euro was reserved for a final event to finish the CLARIN-NL project with. This event is planned for March 13, 2015. Expected costs are 15k.

² After that, she will also work on semantic interoperability in the context of CLARIAH, again as an employee of UU.



2.4 National Coordination Point

- The EB-members were appointed until April 1, 2015 (46k euro, Budget Id: 201, 202, 207, 208).
- The CLARIN-NL director will be appointed for 0.2 FTE for the remainder of 2015. (19k euro, Budget Id: 212)
- The project secretary's appointment runs until June 1, 2015 but its size decreases (from 0.4 to 0.2 fte).³ 6.5k euro, Budget Id: 212, 203, 204

2.5 EU-Line

The CLARIN EU fee was budgeted for 2015 and will be spent. 271k euro, Budget Id: 128.

3 2015 Overall Budget

Based on the data from the CLARIN-NL 2014 Annual Report, the amount that can be spent in 2015 freely (i.e. apart from commitments already made earlier) is 352,314.09 euros. The plan proposed here proposes activities for some 7k euros more, but we are confident that we will remain within budget since it is very likely that some budgeted items will cost less than planned.

Item	Amount
Commitments	€ 1.026.341,49
New activities	€ 359.588,73
Total	€ 1.385.930,22

The details for CLARIN-NL budget for 2015 can be found in Appendix A.

Appendix A Budget Details

Budget Details are available upon request.

³ Arwin van der Zwan has a corresponding appointment increasing in size at Huygens ING as project secretary for CLARIAH-CORE.