



CLARIN-NL Annual Report 2013

1 Introduction

This document provides a description of the activities carried out in CLARIN-NL for 2013. Explanation and Expansion of acronyms in this document can be found on the CLARIN-NL website¹ and in Appendix D.

2 Executive Summary

The major event in 2013 for CLARIN-NL was the *Interim Evaluation* by the International Advisory Panel (IAP). It provided a very positive evaluation of CLARIN-NL and several recommendations for further improving it. It based itself on various CLARIN-NL documents, discussions with CLARIN-NL board members and participants, in particular at the very successful IAP *Demonstrator Event* held in Utrecht on September 4, 2013, at which most results of CLARIN-NL so far were demonstrated.

Many running demonstrator and curation projects were successfully finished in 2013. The major projects for the construction of the infrastructure (IIP, S&D, TTNWW, CLAVAS)² were concluded. TTNWW resulted in a workflow system that enables a researcher to combine a wide variety of language and speech technology web services to enrich data for his or her research. A keynote speech on this system has been held at CLIN 2013 in Enschede. Other infrastructure projects, such as support for CLAM, FoLiA, CMDI, and ISOCAT are continuing, as are the Call 4 demonstrator and curation projects. Most CLARIN-centres initiated the formal procedure for certification as CLARIN centre, a procedure that is currently on-going.

Work was initiated on the CLARIN-NL portal. Until the portal is ready, [this page](#) lists and classifies the major results of CLARIN-NL so far, as well as the ones that are to be expected in 2014.

A wide range of events was organized by CLARIN-NL: various tutorials (both nationally and internationally), workshops, information sessions, etc. CLARIN-NL also supported various independent events financially, and it has supported various researchers to attend workshops and conferences. Many of these workshop and conference visits have led to presentations and publications. Newsflashes have been sent out, and CLARIN-NL has supported *eData & Research*, in which a wide variety of articles on or related to CLARIN have appeared. A number of new movies to show the results of various subprojects have been created. The website's underlying content management system has been upgraded to a state-of-the art version. Moreover, many CLARIN-participants wrote a [blog](#) about the progress of their projects and/or about their CLARIN-project presentations in and outside the Netherlands.

¹ <http://www.clarin.nl/node/8>

² For details on these projects, see infra.



The Netherlands was prominently present at the Annual CLARIN Conference, held this year in Prague, Czech Republic, and continues to play a leading role in the CLARIN ERIC.

As an aside, but highly relevant, a new CLARIAH proposal has been prepared and submitted and is currently being evaluated.

3 Planned Activities

Interim Evaluation

The *Interim Evaluation* was the major event for CLARIN-NL in 2013. For the Interim Evaluation the following activities were carried out:

- The CLARIN-NL EB supported by a number of experts performed a self-evaluation, which resulted in a report titled [CLARIN-NL Interim Evaluation: Self-Evaluation](#), which was accompanied by the [CLARIN-NL Fact Book](#).
- The Self-Evaluation and the Fact Book formed the basis for an evaluation of CLARIN-NL by the International Advisory Panel (IAP), written down in the *Interim Evaluation Report*.
- Draft versions of the Interim Evaluation Report and of the Response to the Interim Evaluation Report by the CLARIN-NL EB were discussed with the IAP on Sep 4, 2013. This discussion was combined with a successful event held in Utrecht in which most CLARIN-NL results were demonstrated to the IAP and others. After this event, the IAP finalized its [Interim Evaluation Report](#) and the EB its [Response to the Interim Evaluation Report](#), which were published on the CLARIN-NL website on September 17, 2013.

The CLARIN-NL Interim Evaluation was very positive. We quote from the Interim Evaluation Report:

The CLARIN-NL project has very clearly stated objectives and a very transparent and reasonable set of criteria for success or failure. It is clear that a lot of work has been achieved and that CLARIN-NL really succeeded to be an example for other European countries. CLARIN-NL sets an example to the whole CLARIN community not only in terms of the activities but also by the way that it is managed and administered. The CLARIN initiative and its management have been truly exemplary.

But in addition to these very positive comments, the IAP also made several recommendations to make CLARIN-NL even better. The Response to the Interim Evaluation addresses these recommendations and sketches the plan to accommodate them.

EU-Line (WBS 1.9)

As planned, the budget for this item was spent on the ERIC membership fee:

Alice Dijkstra is the representative for the Netherlands in the CLARIN ERIC General Assembly, with Jan Odijk as expert.

CLARIN-NL is very active with regard to the European enterprise to create a CLARIN infrastructure under the guidance of CLARIN ERIC, and in other CLARIN or infrastructure-related projects and initiatives. This can be illustrated with many examples:



- CLARIN ERIC is hosted by the Netherlands, more specifically Utrecht University
- The CLARIN ERIC director is Steven Krauwer.
- NL is represented in the National Coordinator's Forum, attends the monthly (virtual) meetings, and is active in it. (Jan Odijk/Daan Broeder)
- Active participation in CLARIN ERIC committees, e.g.³
 - Marc Kemps-Snijders, Member Standing Committee for CLARIN Technical Centres⁴
 - Daan Broeder member Centre Assessment Committee
 - Daan Broeder, member CLARIN ERIC standards committee
 - Menzo Windhouwer, member CLARIN ERIC standards committee
 - Eelco Ferwerda, member CLARIN ERIC Legal Issues Committee
 - Arjan van Hessen, Web Editorial Board
 - Arjan van Hessen, Arwin van der Zwan, Hetty Winkel in the CLARIN ERIC Office
- CLARIN-NL employees have organized LREC tutorials and workshops on CLARIN-related topics (metadata, semantic interoperability)
- Other tutorials and workshops that attract international researchers have been organized
- A large delegation from CLARIN-NL attended the second [CLARIN Annual Conference](#) in Prague, Czech Republic, and many of them had contributions in the form of presentations and/or showing demonstrators.
- Many CLARIN-NL people and/or organizations are involved in important European projects and initiatives that are directly relevant to infrastructures in general and the CLARIN infrastructure in particular, e.g.
 - [DASISH](#)⁵, which brings together [all 5 ESFRI research infrastructure initiatives](#)⁶ in the social sciences and humanities
 - [EUDAT](#)⁷, which brings together data management service providers and research communities who are directly involved with the design of data services. The consortium includes key representatives from research community infrastructures from diverse disciplines.
 - [RDA](#)⁸ (Research Data Alliance), which is an organisation that aims to accelerate and facilitate standardizing and harmonizing research data sharing and exchange.
 - [CLARA](#) is training a new generation of researchers who will be able to cooperate across national boundaries on the establishment of a common language resources infrastructure and its exploitation for the construction of the next generation of language models with wide theoretical and applied significance

³ This is the same as in 2012, with the exception of Remco van Veenendaal (member CLARIN ERIC Legal Issues Committee). Remco is still active, but his activities now fall under the responsibility of the Dutch Language Union.

⁴ Dieter Van Uytvanck, is Chair Standing Committee for CLARIN Technical Centres (on behalf of Germany)

⁵ <http://dasish.eu/>

⁶ <http://dasish.eu/links/>

⁷ <http://www.eudat.eu/>

⁸ <http://rd-alliance.org/>



NL-Line

Technical (WBS 1.1)

The project to create essential infrastructural functionality (IIP project) and the project to create search facilities in metadata and actual data in the infrastructure (S&D project) have run well and delivered intermediate status reports, and several initial versions of crucial components of the infrastructure.

Infrastructure Implementation Plan (IIP)

Participants in the IIP project are the centres that expressed the ambition and made the commitment to become a CLARIN type A/B centre in the Netherlands, i.e. MPI (Nijmegen), Meertens Institute (Amsterdam), INL (Leiden), and DANS (The Hague).

In the course of 2012, Huygens Institute for Dutch History (Huygens ING), which also has expressed the ambition to become a CLARIN Centre, submitted a plan for achieving this. The plan has been accepted, and has started in 2013.

As reported last year, a new centre type (B+), has been defined, In addition to the known types A and B. A B+-centre is a centre that explicitly also aims to host software and data management services (in addition to hosting data)) for the whole CLARIN community.

MPI has reconfirmed its ambition to become an A/B+ centre. DANS has made clear that it aims to be a B-centre (and has made a plan towards it). INL, HI, and MI aim to be B+-centres. All centres have worked towards certification as a CLARIN-centre. A dedicated workshop on this for whole CLARIN has been held in 2013, in Copenhagen.

Visibility of the metadata of many data and tools that reside at the centres via the Virtual Language Observatory is improving, but still many data and tools are lacking there.

The IIP project is responsible for contributing to the CMDI metadata infrastructure registries and tools such as the VLO metadata catalogue and the ARBIL metadata editor. It also works on building and maintaining several registries that are important for CLARIN such as ISOcat, RELcat and SCHEMAcat.

So far, only MPI has been certified as a CLARIN centre. In 2013 three additional centres (Meertens, DANS, and INL) decided to attempt certification as CLARIN centre. This process is currently running. Huygens has decided to postpone certification to 2014.

The IIP project ended in 2013, and the final reports are expected in early 2014.

Search & Develop (S&D)

The S&D project aims to create a network of centres in the Netherlands, focusing around text content search services. The S&D projects aims to create search facilities for search in metadata stored in a centralized database (itself filled and regularly updated via metadata harvesting), as well as federated search facilities in the actual data. A first prototype S&D federated search demonstrator



with limited functionality (term search only) has been created.⁹ At the Sofia CLARIN conference (Oct 2012), it became clear that there are different views (and implementations) on federated search (differences between Germany and the Netherlands). A dedicated CLARIN workshop on federated search was held in 2013, though the issue of these different views was not discussed there. The Netherlands, however, has decided to conform to the German approach.

The S&D project is otherwise essentially finished, though the final reports have not yet been submitted..

CLAVAS: A Vocabulary Service for CLARIN

It has become clear in the IIP project that ISOCAT cannot function as the only concept registry in the CLARIN infrastructure. For certain concepts ISOCAT is not particularly suited, and for other concepts independently maintained registries exist and duplication is to be avoided. For this reason, a separate project, CLAVAS, was started up to build an interface to a variety of concept and data category registries, so that CLARIN users can transparently use concepts and data categories from each of them. This project has been finished, but it still has to be formally closed. As reported last year, one part of the project, automatic normalization of organisation names, was not successful. It has been decided to have the relevant work done manually, and tests for this have been carried out by the Data Curation Service in 2013.

CLARIN Data Providers

The Netherlands has a range of centres that already or since recently have making data available in their core mission and that are of great importance to humanities researchers. These include the *Koninklijke Bibliotheek* (KB, National Library), *het Nederlands Instituut voor Beeld & Geluid* (NIBG, Netherlands Institute for Sound and Vision), *het Nationaal Archief* (NA, the National Archive), the *Digitale Bibliotheek van de Nederlandse Literatuur* (DBNL, Digital Library of Dutch Literature), the various university libraries, and also organizations such as Department of information Services of the *Tweede Kamer* (Parliament). CLARIN-NL wants to make sure that data from such organizations is available to humanities researchers in the CLARIN infrastructure in a CLARIN-compliant manner. For this reason, a special type of centre has been defined, a *CLARIN Data Provider* (also sometimes called a CLARIN D-Centre), and several of the organizations mentioned have been invited to make a proposal to become such a *CLARIN Data Provider*:

- KB has submitted a plan, which has been approved and is currently being executed. It was already clear last year that the originally targeted project end date (Feb 2013) was not feasible. KB has not been able to confirm a clear new target date, and the project has not been finished in 2013.
- NIBG executed its plan successfully already last year. The so-called Academia-collection is now accessible in a CLARIN-compatible manner, e.g. via the VLO. Access to the actual data does require a subscription to the Academia-service, which most (but not all) Dutch

⁹ <http://lux17.mpi.nl/ds/fedsearch/>



universities have. Most foreign universities and research institutes do not have such a subscription.

- DBNL has submitted a plan, which was approved. In the meantime, DBNL was transferred to the KB. DBNL has continued working on the project, and has finished it by the end of January 2014. It now has to submit the final reports, so that the results of the project can be evaluated.
- UBU has submitted a plan, which was approved as of March 19, 2013, and which has been executed and approved early in 2014.

Data Deposition Service

No revised plan was made for a Data Deposition Service (DDS). However, the EB still believes that DDS functionality is important and an essential ingredient for reducing costs for sustainability. It will have to be taken up in successors of CLARIN-NL (e.g. CLARIAH).

CMDI Metadata scheme for tools

The project to create a CMDI *metadata scheme for tools* has finished by January 2013, and it has been reported on at CLIN 2013. After evaluation, some modifications of the scheme are desirable, and full scale testing on all CLARIN-NL tools is required. A start with this has been made but it would be good to include this in the 2014 activities and budget..

Integration

An important activity started up in 2013 was an *integration activity* (the *CLAPOP project*) in which the various components created separately in a variety of projects (infrastructure functionality, search functionality, curated data, curated tools, demonstrators, web applications, etc. etc.) come together and are made accessible via a single portal to the CLARIN infrastructure. The CLARIN infrastructure and the portal towards it will have to grow incrementally and any new additions, modifications etc. will have to be accompanied by good PR activities. For the CLAPOP project, preparatory activities have been carried out to set up the portal (e.g. integration of federated login with Drupal), and to design the portal, and work is on-going on filling the portal with content aimed at and suited for the target audience: humanities researchers. A [simple web page with \(classified\) links](#)¹⁰ to results of CLARIN-NL has been made available, can already be used, and can form the basis for the content of the CLARIN-NL Portal. Similar activities are on-going at the ERIC level, and there is close collaboration between the ERIC and the national level. Discussions on the division of labour between the national and the ERIC level portals have started up and will continue in 2014.

Data, Tools and Services (WBS 1.2/1.3/1.4)

Calls

All projects from Call 1 and Call 2 have been finished.

¹⁰ <http://www.clarin.nl/node/404>



The CKCC project has finished and its results are currently being evaluated.

We targeted for all Call 3 projects to finish and be formally closed ultimately by mid-2013, but that turned out to be too optimistic. Most have finished the work but have either not yet been approved (EMIT-X, PILNAR, GrNe, BILAND), or have not submitted their final reports (DiscAn, Mindisc, FESLI, D-LUCEA, CORNETTO-LMF-RDF, Namespace). For others the financial reports are being processed (Multicon, VK, Polimedia).

All Call 4 projects have started, and are running. Some have requested an extension of the end date (see page 21 for details). The OpenSONAR project has adapted its proposal, made an extension, and strengthened its management. It has been approved and is running since September 2013.

All Call 4 projects will have to be finished before the end of 2014.

The Nederlab demonstrator project has been postponed to 2014, because work for CLARIAH by Nederlab had higher priority.

The LAISEANG project has started and has requested (and been given) an extension of the end date, so that it will end only in 2014.

All calls have been calls for data curation and demonstrator projects. In the approved projects data and tools developed and in use in the Netherlands will be curated and a series of demonstrators will be developed on the basis of existing technologies and applications to attempt to make them CLARIN-compatible. A major aim with these projects is to test the standards proposed within CLARIN against these data and tools, and to be able to provide evidence-based feedback on the proposed standards and thus to influence their determination in the CLARIN context.

In each project metadata in accordance with CMDI will have to be created and tested, making use of the experience gained and the tools developed in the Metadata project described above. Each demonstrator project will create a demonstrator web application, in which a clear separation between user interface modules and core modules is made. An API has to be defined and documented for the core module. The user interface will make calls to the core module via this API. This will make it possible later to easily turn the core module into a web service.

All projects have to deliver a document describing the requirements the infrastructure should meet and or desirable features it should offer, so that these can form input for the requirements and specifications of the CLARIN infrastructure.

In the first call, 17 project proposals were submitted. Of these, 11 were awarded funding by the Board on the basis of recommendations from the National Advisory Panel (NAP) and the Executive Board (EB). In addition, the CKCC project was awarded funding.

In the second call, 17 project proposals were submitted. Of these, 9 were awarded funding by the Board on the basis of recommendations from the National Advisory Panel (NAP) and the Executive Board (EB).

In the third call, Closed Call, five researchers were invited to submit a proposal. The disciplines covered were Philology/Classical Studies, Studies on selective language impairment, Religion Studies,

the study of discourse, and descriptive linguistics/language documentation. Four proposals were submitted, and have started early 2012.

No proposal was submitted for the discipline descriptive linguistics/language documentation, but the researcher in question was requested to still submit a proposal.

In the third call, Open Call, 22 project proposals were submitted. 2 proposals were rejected because of IPR-issues. Of the remaining 20 project proposals, 9 were awarded funding by the Board on the basis of recommendations from the International Advisory Panel (IAP), the National Advisory Panel (NAP) and the Executive Board (EB).

In the fourth call, Closed Call, 2 project proposals were submitted and awarded funding. In the Call 4 Open Call 18 project proposals were submitted. One of these was rejected by the EB because of non-compliance, and of the remaining 17, 9 could be awarded funding.

An overview of the projects is provided here. For each project, it consists of the project acronym and title, the project coordinator, a very short description of the project goal, a very short description of the scientific impact of the project, and its current status. The status describes the situation as of 31 December 2013, though where needed also its status as of March 2014 is specified. More detailed descriptions of the projects can be found on the CLARIN-NL website.¹¹

Call 1 Projects

ADEPT	Assaying Differences via Edit-Distance of Pronunciation Transcriptions
Project Coordinator	Dr. C. Gooskens (Groningen University)
Goal	The goal of the project is to provide a web application capable of measuring the differences in sets of phonetic (or phonemic) transcriptions via edit distance.
Budget	€ 57K
Centre	Meertens Instituut
Scientific impact	A user-friendly web application version of this existing tool will make it usable for a wide range of humanities researchers such as, phonologists, phoneticians, dialectologists and second-language learning specialists.
Status	The project has effectively finished and has submitted its final reports, which have been approved. [web service]

AAM-LR	Automatic Annotation of Multi-modal Language Resources
Project Coordinator	Prof.dr. L. Boves (Radboud University Nijmegen)
Goal	The AAM-LR project aims at building a demonstrator of a web service that will help field researchers to annotate audio- and video-recordings
Budget	€ 56K

¹¹ <http://www.clarin.nl/node/76>

Centre	MPI
Scientific impact	It will facilitate all field research in which digital audio or video recordings are used.
Status	The project has effectively finished and has submitted its final reports, which have been approved. [web application ¹² show case ¹³]

Adelheid	A Distributed Lemmatizer for Historical Dutch
Project Coordinator	Dr. H. van Halteren (Radboud University Nijmegen)
Goal	This project aims at providing a web-application with which an end user can have historical Dutch text tokenized, lemmatized and part-of-speech tagged.
Budget	€ 55K
Centre	MPI
Scientific impact	It will facilitate research on historical texts in fields such as historical linguistics, literature science and history.
Status	The project has finished and has submitted its final reports, which have been approved. The Second Call project INPOLDER makes use of the results of Adelheid. [web service]

DUELME-LMF	Converting DUELME into LMF format
Project Coordinator	Prof.dr. J.E.J.M. Odijk (Utrecht University)
Goal	The goal of the project is develop converters between DUELME and LMF format, and to create a curated DUELME resource fully compliant with standards supported by CLARIN.
Budget	€ 60K
Centre	INL
Scientific impact	The DUELME database makes it possible to address a variety of research questions related to Dutch multi-word expressions in research areas such as computational linguistics and natural language processing, theoretical linguistics and psycholinguistics.
Status	The project has effectively finished, and has submitted its final reports, which have been approved. [metadata ¹⁴]

¹² <http://lux17.mpi.nl/aamlr>. This is a temporary URL, which may disappear in the future

¹³ <http://www.clarin.nl/showcase/aam-lr/268>

¹⁴

http://oai.dev.clarin.inl.nl/oai_server.php?verb=GetRecord&identifier=hdl:10032/89511c6d452ec7110717d8b3fc359e4d&metadataPrefix=cmdi

INTER-VIEWS	Curation of Interview Data
Project Coordinator	Dr. H. van den Heuvel (Radboud University Nijmegen)
Goal	The INTER-VIEWS project will make a corpus of interview data available to the community of researchers in the humanities.
Budget	€ 58K
Centre	DANS
Scientific impact	This project will contribute to facilitating all research that makes use of interviews, and more specifically research into the Second World War.
Status	The project has effectively finished, and has submitted its final reports, which have been approved. [web service]

MIMORE	Microcomparative Morphosyntax Research Tool
Project Coordinator	Prof.dr. S. Barbiers (Meertens Institute and Utrecht University)
Goal	The demonstrator tool MIMORE will create a common search engine for the databases DynaSAND , DiDDD and MAND
Budget	€ 54K
Centre	Meertens Instituut
Scientific impact	The tool is especially relevant to research in variation linguistics. It will allow investigating theoretical questions concerning the language system and language variation, and the geographic distribution of (morpho-) syntactic variables.
Status	The project has effectively finished, and has submitted its final reports, which have been approved. [web service]

SignLinC	Linking lexical databases and annotated corpora of signed languages
Project Coordinator	Dr. O.A. Crasborn (Radboud University Nijmegen)
Goal	This project aims to link two independently evolved data sets for a signed language: the Corpus NGT and the lexical database of the Dutch Sign Centre.
Budget	€ 56K
Centre	MPI
Scientific impact	The linked databases and corpora will facilitate research in the area of sign language linguistics
Status	The project has effectively finished, and has submitted its final reports, which have been approved. [web service]

TICCLops	Text-Induced Corpus Clean-up online processing system
Project Coordinator	Dr. M. Reynaert (Tilburg University)
Goal	This demonstrator project will develop a tool that allows CLARIN users to submit their corpora for fully automatic spelling correction and normalization
Budget	€ 54K
Centre	INL
Scientific impact	The tool will facilitate all research that makes use of text corpora.
Status	The project has effectively finished, and has submitted its final reports, which have been approved. [web service]

TDS Curator	A web-services architecture to curate the Typological Database System
Project Coordinator	Dr. A. Dimitriadis (Utrecht University)
Goal	TDS Curator will make the Typological Database System (TDS) into a sustainable service that conforms to CLARIN infrastructural requirements.
Budget	€ 57K
Centre	DANS
Scientific impact	The (TDS) facilitates typological linguistic research by providing integrated access to multiple independently developed typological databases through a common web interface
Status	The project has finished and has submitted its final reports, which have been approved [web service]

TQE	Transcription Quality Evaluation
Project Coordinator	Dr. H. Strik (Radboud University Nijmegen)
Goal	The TQE project aims to create a completely automatic Transcription Quality Evaluation (TQE) tool.
Budget	€ 57K
Centre	MPI
Scientific impact	The tool will be useful for validating, obtaining, and selecting phone transcriptions, for detecting phone strings (e.g. words) with deviating pronunciation, and, in general, it can be usefully applied in all research - in various (sub-) fields of humanities and language and speech technology in which audio and phonetic transcriptions are involved.
Status	The project has effectively finished, and has submitted its final reports, which have been approved. [web service]

WFT-GTB	Integrating the Wurdboek fan 'e Fryske Taal into the Geïntegreerde Taalbank
Project Coordinator	Drs. H. Sijens (Fryske Akademy)
Goal	The project aims to carry out data curation of the <i>Wurdboek fan de Fryske Taal (WFT)</i> database and integrate the data in the <i>Geïntegreerde Taalbank (GTB)</i> and demonstrate it via the GTB dictionary web application.
Budget	€ 56K
Centre	INL
Scientific impact	Integrating the WFT dictionary in the GTB will facilitate studying formal, semantic and idiomatic aspects of Modern West Frisian by allowing extensive forms of free and structured (in particular thematic) search queries, and it will enable comparative studies with Dutch materials.
Status	The project has effectively finished and demonstrated its results at Euralex 2010 held in Leeuwarden. It has submitted its final reports, which have been approved. [web application]

CKCC project	CKCC project
Project Coordinator	Drs. G. Gerritsen (Huygens ING)
Goal	With respect to CLARIN-NL, the project increases its use of language technology and produces the results in CLARIN-compatible manner. The project runs until Nov 1, 2012.
Budget	€ 59K
Centre	DANS/Huygens ING
Scientific impact	The CKCC project (<i>Circulation of Knowledge and Learned Practices in the 17th-century Dutch Republic, 'Geleerdenbrievenproject'</i>) is an independently financed (NWO) project that is carried out by a consortium of Dutch institutes (Descartes, Historical Literature Amsterdam, Huygens, KB (National Library), DANS and VKS) and coordinated by Wijnand Mijnhardt (Utrecht University). It investigates, on the basis of a corpus of 20,000 letters of scientists from the 17th century in the Dutch Republic and using language technology, the research question of how knowledge circulated in the 17 th century. It was selected in the CLARIN-EU call for humanities and social sciences projects as the project proposal that “[would] best demonstrate the use of LRT and would show the potential of a research infrastructure in the humanities” (CLARIN Newsletter 6, p. 3). ¹⁵
Status	The project has finished and is currently being evaluated. [web application]

¹⁵ http://www.clarin.eu/files/cnl06_web_0.pdf

Call 2 Projects

Arthurian-Fiction	Arthurian Fiction in Medieval Europe
Project Coordinator	Dr. B. Besamusca (Utrecht University)
Goal	ArthurianFiction is a mixed demonstrator and data curation project that aims to carry out data curation for two databases with data for literary research in the area of European Arthurian fiction and to create a demonstrator that will allow users access to the data for searching and editing.
Budget	€ 47K
Centre	Huygens ING
Scientific impact	It will enable access to the Arthurian Fiction database to a wide range of humanities researchers such as literary historians in all European languages, cultural historians, medieval book specialists and art historians.
Status	The project is finished and has submitted its final reports, which have been approved. [web application]

C-DSD	Curating the Dutch Song Database
Project Coordinator	Dr. E. Stronks (Utrecht University)
Goal	C-DSD is a data curation project that aims to curate the Dutch Song Database (DSD), a database in the field of Literary Studies containing (meta-) data on 140.000 songs and their 15.000 sources from the Middle Ages to the present day.
Budget	€ 29K
Centre	Meertens Institute
Scientific impact	This project will create new opportunities for (international) collaboration with other projects with similar source materials.
Status	The project submitted its final reports, which have been approved. [metadata in the VLO]

COAVA	Cognition, Acquisition and Variation Tool
Project Coordinator	Dr. L. Cornips (Meertens Institute Amsterdam)
Goal	COAVA is a mixed demonstrator and data curation project proposal in which a tool is developed for easily exploring the linguistic characteristics of objects from two databases of language variation (historical dialectology) on the one hand and three first language acquisition databases on the other hand.
Budget	€ 66K
Centre	Meertens Institute
Scientific	It will enable interdisciplinary research into the relation between language

impact	acquisition and language variation.
Status	The project has finished and submitted its final reports, which have been approved. [web application]

INPOLDER	Integrated Parser and Lemmatizer Dutch in Retrospect
Project Coordinator	Prof.dr. A. van Kemenade (Radboud University Nijmegen)
Goal	INPOLDER is a mixed demonstrator and data curation project that aims to provide syntactically analysed corpus material for historical Dutch. It uses the Adelheid tagger for PoS-tagging and a variant of the Penn-Helsinki parser for historical texts for enriching the texts with parse trees.
Budget	€ 69K
Centre	Meertens Institute
Scientific impact	It will fill the gap in availability of syntactically analysed corpus material for historical Dutch, enabling historical syntactic, morpho-syntactic and lexical research as well as benefiting non-linguistic historical research.
Status	The project has finished and the final reports have been submitted and approved. [web application]

IPROSLA	Integrating and publishing resources on sign language acquisition
Project Coordinator	Prof.dr. P. Fikkert (Radboud University Nijmegen)
Goal	IPROSLA is a resource curation project that aims to integrate two different data sets on sign language acquisition by documenting the two data sets with CMDI, and archiving them at the MPI language archive. The data consist of a set of longitudinal data of deaf children from deaf and hearing parents collected at the UvA, and a new collection of longitudinal data collected at the RU from hearing children of deaf parents.
Budget	€ 67K
Centre	MPI
Scientific impact	By integrating two datasets in integrated archive and making them visible and accessible, it will enable better and more research on Sign Language.
Status	The project has finished and submitted its final reports, which have been approved. [metadata via the VLO]

NEHOL	Negerhollands Database
Project Coordinator	Prof.dr. P.C. Muysken (Radboud University Nijmegen)
Goal	NEHOL is a data curation project that aims to make available to the CLARIN community the data from the Dutch-lexifier Creole language Negerhollands, the

	now extinct Creole language of the Virgins Islands.
Budget	€ 43K
Centre	MPI
Scientific impact	It will enable research into Negerhollands and thereby into Creole languages and language creolization processes in general.
Status	The project has finished and submitted its final reports, which have been approved. [metadata via the VLO]

VU-DNC	VU Diachronic Newspaper Corpus
Project Coordinator	Prof.dr. W. Spooren (Free University Amsterdam)
Goal	VU-DNC is a data curation project that aims to make a unique diachronic corpus of Dutch newspaper articles from five major Dutch newspapers from 1950/1951 and 2002 (2 million words) available to humanities researchers, to extend the discourse annotation with additional lexico-grammatical features and to create a gold standard benchmark for OCR-post-correction tools, all of this in a fully CLARIN-compliant manner.
Budget	€ 59K
Centre	INL
Scientific impact	It will enable research into the recent history and development of the Dutch written language, as well as research into subjectivity and quotations and the lexico-grammatical features encoding them.
Status	The project has finished and submitted its final reports, which have been approved. [Data and its metadata]

WAHSP	Web-application for historical sentiment mining in public media
Project Coordinator	Prof.dr. T. Pieters (Utrecht University)
Goal	WAHSP is a demonstrator project that aims to provide advanced forms of text mining (more specifically, sentiment mining) in large historical datasets of newspapers and journals in the form of a CLARIN compliant web-application, addressing research questions of historians and policy researchers.
Budget	€ 120K
Centre	Huygens ING
Scientific impact	It will enable historical research by providing facilities for mining large historical data sets, where it focuses on the opinions/perceptions regarding the use and abuse of drugs between 1900 and 1945.
Status	The project has finished and submitted its final reports, which have been approved. [web application]. There was a problem with login, but it was agreed to solve this in combination with BILAND (see there).

WIP	War in Parliament
Project Coordinator	Dr. H. Piersma (NIOD)
Goal	WIP is a mixed demonstrator and data curation project that aims to make the Dutch Hansard database compliant with the CLARIN infrastructure and to provide an advanced search engine for it in order to make it possible to carry out historical and social science research, in particular with respect to the Second World War.
Budget	€ 118K
Centre	DANS
Scientific impact	It will enable historical research by providing tools for text mining large text corpora, applied in this project in particular for references to the Second World War in parliamentary debates.
Status	The project has finished and submitted its final reports, which have been approved. web application

Call 3 Projects (Closed Call)

FESLI	Functional Elements in Specific Language Impairment
Project Coordinator	Prof.dr. F. Weerman (University of Amsterdam)
Goal	FESLI (CLARIN-NL-11-006) is a mixed demonstrator and resource curation project in the Closed Call that aims to curate Dutch corpora from monolingual and bilingual children with Specific Language Impairment (SLI) and provide tools based on the CLARIN-NL project COAVA that allow quantitative and qualitative comparisons to be made esp. with regard to functional elements in language between the SLI data and CHILDES monolingual data.
Budget	€ 79K
Centre	Meertens Instituut
Scientific impact	The project will curate important sets of monolingual and multilingual SLI data, improve search facilities in the data and will thus greatly facilitate research into SLI.
Status	The project is running since May 1, 2012, and targeted to finish by April 30, 2013. It has received an extension until June 30, 2013. It is effectively finished, but the final reports have not been submitted yet. web application

PILNAR	Pilgrimage Narratives: creating a germ corpus for studying the profile of the modern pilgrim
Project Coordinator	Prof.dr. P. Post (Tilburg University)
Goal	PILNAR (CLARIN-NL-11-007) is a mixed demonstrator and resource curation project in the Closed Call that aims to curate a corpus of pilgrimage narratives and use the heuristic tool "Fields of the Sacred" as a search tool for disclosing the material.

Budget	€ 79K
Centre	Meertens Instituut
Scientific impact	It will make searching in pilgrimage narratives much easier and more powerful, this boosting research in the religion sciences.
Status	The project is running since April 1, 2012 and targeted to finish by April 1, 2013. It has received an extension until May 31 st , 2013. It has submitted its deliverables and final reports in December 2013, which are currently being evaluated. [web application]

GrNe	Grieks-Nederlands woordenboek (Greek-Dutch dictionary)
Project Coordinator	Prof.dr. I. Sluiter (Leiden University)
Goal	GrNe (CLARIN-NL-11-008) is a mixed demonstrator and resource curation project in the Closed Call that aims curate a part of the dictionary Ancient Greek-Dutch currently under construction and to create a tool for search functionality including an interface in which (inter alia) inflected forms can be input and the entries of their lemmata will be returned using an existing lemmatizer (called `parser` in the proposal, which appears to be a misnomer).
Budget	€ 55K
Centre	University Library of UvA
Scientific impact	It will make a part of the Classical Greek dictionary currently being produced available via a web interface, with increased search functionality offering search options both for lemmas and for inflected word forms, thus greatly facilitating research and education in classical philology.
Status	The project is running since April 1, 2012 and targeted to finish by April 1, 2013. It has received an extension until June 30, 2013. It has submitted its deliverables and final reports, which have been evaluated. The project is currently correcting some errors. [web application]

DiscAn	Towards a Discourse Annotation system for Dutch language corpora
Project Coordinator	Prof.dr. T.J.M. Sanders (Utrecht University)
Goal	DiscAn (CLARIN-NL-11-031) is a resource curation project in the Closed Call that aims curate a set of existing corpora with discourse annotations, to develop the foundations of a discourse annotation system, and to develop an interdisciplinary Dutch-Flemish discourse community of linguists from various subdisciplines.
Budget	€ 80K
Centre	MPI
Scientific impact	The project will bring together and harmonize a wide range of discourse annotated corpora that are currently located in various places, in multiple different format sand generally not easily accessible, thus greatly facilitating research into discourse properties of natural language and organizing a research community working on

	this topic.
Status	The project is running since April 1, 2012 and targeted to finish by April 1, 2013. It has been followed up by the Mindisc project (see there). The deliverables have been submitted, but the final reports are still to be submitted.

Call 3 Projects (Open Call)

BILAND	Towards a flexible and stable CLARIN-supported web-application for bilingual historical analysis of discourses in news media
Project Coordinator	Prof.dr. T. Pieters (Utrecht University)
Goal	BILAND (CLARIN-NL-11-017) is a demonstrator project in the Open Call that aims to enable advanced forms of discourse analysis in large historical datasets in the CLARIN infrastructure.
Budget	€ 120K
Centre	Huygens ING
Scientific impact	The project will greatly facilitate research in history and of advanced discourse analysis in historical datasets.
Status	The project is running since Feb 1, 2012 and originally targeted to finish by Feb 1, 2012. It has requested an extension until May 31 st , 2013, which has been given, and after that an additional extension until July 31 st , 2013 (also given). It has submitted its deliverables and final reports. Two problems arose when evaluating these: (1) Federated login was still not taken care of, despite the agreements made for WAHSP; (2) it turned out that a large part of the application runs on UvA servers, not on Huygens servers. We have requested the project to improve these issues and are working with Huygens ING towards a solution of this problem. [web application]

MultiCon	Multilayer Concordance Functions in ELAN and ANNEX
Project Coordinator	Dr. O.A. Crasborn (Radboud University Nijmegen)
Goal	MultiCon (CLARIN-NL-11-018) is a demonstrator project in the Open Call that aims to create collocation functionality for multimodal data in the tools ELAN and ANNEX by improving their query and concordance functionality and by providing visualization options for multilayer collocations.
Budget	€ 55K
Centre	MPI
Scientific impact	The project will offer new research opportunities in multimodality studies, both in linguistics and beyond.
Status	The project is running since April 1, 2012 and has finished on the target date of June 1, 2013. It has submitted its final reports.

D-LUCEA	Database of the Longitudinal Utrecht Collection of English Accents
Project Coordinator	Dr. H. Quené (Utrecht University)
Goal	D-LUCEA (CLARIN-NL-11-011) is a resource curation project in the Open Call that aims to curate a database of existing speech recordings of L1 and L2 speakers of English.
Budget	€ 67K
Centre	MPI
Scientific impact	This project will greatly facilitate linguistic and phonetic research into L1 and L2 acquisition.
Status	The project is running since April 1, 2012 and was targeted to finish by April 1, 2013. It is claimed to have finished the work, but has not yet submitted any deliverables or final reports.

EMIT-X	Early-Modern Image and Text eXchange
Project Coordinator	Dr. E. Stronks (Utrecht University)
Goal	EMIT-X (CLARIN-NL-11-009) is a resource curation project in the Open Call that aims to curate Emblems Project Utrecht (EPU) resources by converting them into a commonly agreed upon format (no name given) based on work by Rawles and Stäcker and making them available for OAI-PMH harvesting.
Budget	€ 20K
Centre	Huygens ING
Scientific impact	The project will greatly facilitate research in literary studies, esp. the study of emblems.
Status	The project is running since March 1, 2012 and originally targeted to finish by March 1, 2013. It has requested an extension until June 1, 2013, which has been given. The project has finished and submitted its deliverables and final reports. The metadata it produced were (and are) invisible in the VLO because of noncompliance. They are currently visible in a different way and are being evaluated.

NameScope	Mapping the Landscape of Names in Modern Dutch Literature
Project Coordinator	Prof.dr. K. van Dalen-Oskam (Huygens ING)
Goal	NameScope (CLARIN-NL-11-029) is a demonstrator project in the Open Call that aims to annotate a substantial amount of literary works with a rich tag set for named entities using a named entity recognizer and make them accessible and visualize them via a demonstrator web application.
Budget	€ 120K
Centre	INL

Scientific impact	The project will greatly enhance the research opportunities in literary studies esp. the study of named entities and their roles in literary works.
Status	The project is running since March 15, 2012 and targeted to finish by April 1, 2013. It has been given an extension until May 1, 2013. The project has finished its work, submitted its deliverables, and its final technical report. These have all been approved. [web application]

VK	Verrijkt Koninkrijk (Enriched Kingdom)
Project Coordinator	Dr. K. Ribbens (NIOD)
Goal	VK (CLARIN-NL-11-014) is a mixed demonstrator and resource curation project in the Open Call that aims to make a demonstrator to gather data on pillarization in Dr. Lou de Jong's VK and to enrich the resulting data and make them available to the CLARIN community.
Budget	€ 111K
Centre	DANS
Scientific impact	The project will greatly facilitate historical research, esp. on the Second World War, and on Lou de Jong's work on this topic, and more specifically on the representation of pillarization in this work.
Status	The project is running since Feb 1, 2012 and originally targeted to finish by Jan 31, 2013. It has requested an extension until March 31, 2013, which has been given. The project has finished, submitted its deliverables and final reports, which have been approved. [web application]

Cornetto-LMF-RDF	Curated Cornetto database in LMF and RDF
Project Coordinator	Prof.dr. P. Vossen (Free University Amsterdam)
Goal	Cornetto-LMF-RDF (CLARIN-NL-11-020) is a mixed demonstrator and resource curation project in the Open Call that aims to convert the Dutch semantic lexical database Cornetto to LMF (lexicon part) and RDF/SKOS (WordNet part).
Budget	€ 117K
Centre	INL
Scientific impact	This project will greatly enhance the research opportunities in study of semantics in particular and linguistics more generally.
Status	The project is running since May 1, 2012 and targeted to finish by May 1, 2013. The project has finished, submitted its deliverables and final technical report, which have been approved. [web application]

MIGMAP	Interactive migration maps for the 20th century
---------------	--

Project Coordinator	Dr. G. Bloothoof (Utrecht University and Meertens Institute Amsterdam)
Goal	MIGMAP (CLARIN-NL-11-012) is a demonstrator project in the Open Call that aims to develop a web application which visualizes migration by generation and gender based on the Meertens “Kaart” software module.
Budget	€ 54K
Centre	Meertens Institute
Scientific impact	The project will facilitate research into migration and linguistics, esp. language change.
Status	The project is running since March 1, 2012 and is targeted to finish by Dec 31, 2012. It has finished and its final reports and deliverables have been approved. [web application]

Polimedia	Interlinking multimedia for the analysis of media coverage of political debates
Project Coordinator	Prof.dr. H. Beunders (Erasmus University Rotterdam)
Goal	PoliMedia (CLARIN-NL-11-022) is a mixed demonstrator and resource curation project in the Open Call that aims to showcase the potential of cross-media analysis for research in the humanities by curating annotated data sets of different media types and developing a demonstrator that enables researchers to deploy such an interlinked collection for analysis of media coverage.
Budget	€ 111K
Centre	NIBG
Scientific impact	The project will enhance research opportunities for the cross-media analysis in historical studies.
Status	The project is running since April 1, 2012 and targeted to finish by April 1, 2013. It has finished, and submitted its deliverables and final reports, which have been approved. [web application]

Call 4 Projects

QuaMerdes	Quantitative Content Analysis of Media Researchers' Data
Project Coordinator	Dr. Jasmijn van Gorp (Utrecht University)
Goal	QuaMerdes is a demonstrator project from the closed call that will enable quantitative content analysis of television and printed media. It will expand the existing tool MerDES to support principles of quantitative content analysis and explore new data visualisation paradigms. QuaMerDES will be evaluated in a specific case study that enables media studies scholars to come to new insights about how representations of migrants on Dutch television are related to social, cultural and political values in Dutch society.

Budget	76k
Centre	NISV/KB
Scientific impact	It will enable quantitative content analysis of television and printed media.
Status	The project has started May 1, 2013, and is targeted to finish by May 1, 2014.

@PhilosTEI	<i>TICCLing Philosophy: a TEI corpus-building workflow towards a new computational methodology for philosophy</i>
Project Coordinator	Dr. Arianna Betti (Free university Amsterdam)
Goal	@PhilosTEI is a demonstrator project from the closed call in which an open source, web-based, user-friendly workflow is developed from textual digital images to TEI, based on an OCRopus / Tesseract web service and a multilingual version of the OCR-post correction web service TICCLops. The tool will be demonstrated on a multilingual, multi-script corpus of important 18th-20th-century European philosophical texts.
Budget	70k
Centre	HI
Scientific impact	It will contribute to increasing the digitally available empirical material, esp. for philosophy, but also for other disciplines
Status	The project has started on September 1, 2013 and is targeted to finish by September 1, 2014.

eBNM+	Linked Data on Middle Dutch Sources Kept Worldwide
Project Coordinator	Dr. André Bouwman (Leiden University Libraries)
Goal	<i>e-BNM+</i> is a mixed demonstrator and resource curation project from the open call with two goals: (1) to convert e-BNM collection of textual, codicological and historical information about thousands of Middle Dutch manuscripts kept worldwide into a flexible data structure that will turn e-BNM into a key open access resource to which many other resources can be easily linked; and (2) to create a web application for consultation, using faceted search, and collaborative editing.
Budget	68k
Centre	HI
Scientific impact	It will stimulate research that makes uses of Middle Dutch manuscripts.
Status	The project has started on May 1, 2013 and is targeted to finish by May 1, 2014.

VALID	<i>Vulnerability in Acquisition: Language Impairments in Dutch. Curating five valuable data sets</i>
--------------	---

Project Coordinator	Dr. Jetske Klatter (Radboud University Nijmegen)
Goal	VALID is a resource curation project from the open call that aims to curate five existing, digital data sets of language pathology data collected in the Netherlands, primarily on Dutch, in order to make them available for scientific research in CLARIN-compatible format, as a first, major step in the development of a VALID data archive.
Budget	64k
Centre	MPI
Scientific impact	It will enable better research on language pathology data, primarily for the Dutch language.
Status	The project has started on April 1, 2013 and is targeted to finish by April 1, 2014. It has requested (and been given) an extension until July 1, 2014.

COBWWWEB	Connections Between Women and Writings Within European Borders
Project Coordinator	Dr. Suzan van Dijk (Huygens ING)
Goal	COBWWWEB is a mixed demonstrator and resource curation project from the open call that aims to (1) connect the WomenWriters database to other national collections in women's literature; (2) build a research application for scholars; and (3) create a set of standards to exchange data based on CLARIN guidelines for shared metadata and service-based infrastructures, which will be implemented for a variety of databases by a range of international partners.
Budget	95k
Centre	HI
Scientific impact	It will enable better research on female writers and the reception of their works
Status	The project has started on April 1, 2013 and is targeted to finish by December 31, 2013. It has not finished yet.

OpenSoNaR	Online Personal Exploration and Navigation of SoNaR
Project Coordinator	Dr. Martin Reynaert (Tilburg University)
Goal	OpenSoNaR is a demonstrator project from the open call that aims to develop a tool for exploring the SoNaR-500 reference corpus by layman and specialist researcher alike, using an online frontend (WhiteLab) providing a range of interfaces that provide user-driven functionality, combined with the existing INL-developed back-end system (BlackLab).
Budget	120k
Centre	INL
Scientific	It will provide better empirical support for linguistic and literary research

impact	
Status	The project has started on October 1, 2013 and is targeted to finish on October 1, 2014.

SHEBANQ	System for HEBrew Text: ANnotation for Queries and Markup
Project Coordinator	Prof.dr. Wido van Peursen (Free University Amsterdam)
Goal	SHEBANQ is a mixed demonstrator and resource curation project from the open call that aims to curate the WIVU Hebrew Text Database and to develop a web application demonstrator that enables researchers to perform linguistic queries on the web resource and preserve significant results as annotations to this resource.
Budget	101k
Centre	DANS
Scientific impact	It will enable better research in the Hebrew bible texts, and in theology
Status	The project has started on May 1, 2013 and is targeted to finish on April 30, 2014.

DSS	Dutch Ships and Seamen
Project Coordinator	Prof.dr. Lex Heerma van Voss (Huygens ING)
Goal	DSS is a mixed demonstrator and resource curation project from the open call that aims to bring together the rich maritime historical data currently preserved in many different databases. To that end it will provide a tool chain and methodology for converting legacy datasets, and applies it to an initial set of selected datasets.
Budget	98k
Centre	HI
Scientific impact	It will enable better research into the maritime history of the Netherlands and its predecessors
Status	The project has started on April 1, 2013 and is targeted to finish on February 28, 2014. It has requested and received an extension until April 15, 2014.

EXILSEA	Exploiting ISOcat's Language Sections in ELAN and ANNEX
Project Coordinator	Dr. Onno Crasborn (Radboud University Nijmegen)
Goal	EXILSEA is a mixed demonstrator and resource curation project from the open call that aims to make annotated audio-visual resources more accessible to users of different languages by better exploiting the multilingual features of ISOCAT in the ELAN and ANNEX tools. In addition it will enhance the Corpus NGT, the world's first open access sign language corpus, by updating the existing IMDI metadata to CLARIN-standard CMDI descriptions using bilingual ISOcat categories.

Budget	65k
Centre	MPI
Scientific impact	It will enable multilingual support for research into Sign Language
Status	The project has started on March 1, 2013, and is targeted to finish on May 1, 2014. It has requested an extension till the end of 2014, which was awarded till Nov 1 st , 2014

ColTime	Collaboration on time-based resources
Project Coordinator	Dr. Onno Crasborn (Radboud University Nijmegen)
Goal	ColTime is a demonstrator project from the open call that aims to extend ELAN and ANNEX for the annotation and display of time-based resources such as audio and video with a referencing and note exchanging system.
Budget	68k
Centre	MPI
Scientific impact	It will enable better research on time-based resources
Status	The project has started on March 1, 2013, and is targeted to finish on May 1, 2014. It has requested an extension till the end of 2014, which was awarded till Nov 1 st , 2014

RemBench	A Digital Workbench for Rembrandt Research
Project Coordinator	Prof.dr. Volker Manuth (Radboud University Nijmegen)
Goal	RemBench is a demonstrator project from the open call that aims to demonstrate how (art) historians can benefit from linking eLaborate to resources created by museums, archives and libraries. The demonstrator will use eLaborate to link the Rembrandt Documents (RemDoc) database with related resources created by the Rijksbureau voor Kunsthistorische Documentatie (RKD), as well as with a standard university library catalogue.
Budget	121k
Centre	HI
Scientific impact	It will enable better research on the Rembrandt Documents
Status	The project has started on April 1, 2013 and is targeted to finish by March 31, 2014.



Other Projects

Cooperation between the Netherlands and Flanders CLARIN-NL cooperates with CLARIN Flanders in a cooperation project. The overall project consists of two parts:

- **Part I (1356k Euro, NL: 904k Euro, VL: 452k Euro)**
 - In this part a specific cooperation plan has been formulated in which the Netherlands and Flanders aim to adapt various existing components developed in CGN and especially STEVIN and incorporate them into a work flow system for web services that is currently being developed in CLARIN. The whole system will run on servers of recognized CLARIN-centres and aims to offer facilities to researchers from the humanities with little or no technical expertise. These facilities must 1) allow them to address their research questions in a better and easier manner, and 2) create opportunities for formulating completely new types of research questions
 - In the Netherlands this part consists on the one hand of the infrastructure implementation plan described above (section 3) in which the basis for the technical infrastructure will be designed and implemented (with a budget of 504k euro). On the other hand 400kEuro will be invested for the further realization of this part (in the TTNWW project). The formal decision on the exact usage of this amount is planned to be made by the CLARIN-NL board in early 2010. Flanders contributes in this part 452k Euro in accordance with the usual NL: 2/3-VL: 1/3 proportion (all in the TTNWW project).
- **Part II (920k Euro; NL: 680k Euro; VL: 340k Euro)**
 - The contribution from the Netherlands consists of 680k Euro spent on the first open call for demonstrator en resource curation projects.
 - In Flanders the research community represented in CLIF has formulated project proposals for a total amount of 340k Euro.

The project has started in March 2010, and was supposed to run for 2.5 years. It has requested and received an extension until Dec 31, 2012. The project has effectively finished, though there are still some software problems (memory leaks), esp. with regard to the speech modules, which are being investigated. A keynote speech on TTNWW has been held at CLIN 2013 (Jan 18, 2013) in Enschede. Final reports have not been submitted yet. [Link to workflow system](#)¹⁶.

ISOCAT Coordinator

The work by the ISOCAT Coordinator (Ineke Schuurman) started up in 2011 has continued in 2013. A separate report on the ISOCAT Coordinator activities in 2013 is available.

CLAM-support

In the context of the TTNWW project, Maarten van Gompel (initially Tilburg University, currently Radboud University Nijmegen) has developed CLAM, software to facilitate turning existing tools into RESTful web services. CLAM is successful and is in use by most participants in the TTNWW project. For this reason, systematic support, maintenance and extension of the CLAM-functionality are

¹⁶ <http://yago.meertens.knaw.nl/apache/TTNWW/>



required, which has been implemented in a project financed by CLARIN-NL to maintain and further extend CLAM, and to provide support to its users. (carried out by Maarten van Gompel since 2012, 1/10 appointment). This has been continued in 2013, extending the support also to FoLiA, and it will be continued in 2014. A separate report on the activities of 2013 is available.

Support for PID resolution

The EB has not received any proposal from the EPIC consortium, which aims to be a service provider for Persistent Identifiers for research data. The EB has received a proposal for such a service from *Beeld en Geluid*, which aimed to work with *Picturae*. The EB has decided that this proposal was not relevant for CLARIN-NL since all centres have arranged their own solution. The proposal will be resubmitted in the CLARIAH context.

Centres of Expertise (WBS 1.6)

The CLARIN-NL virtual Centre of Expertise on the curation of data, led by the Radboud University of Nijmegen but with close involvement of experts from the CLARIN Centres, has continued the CLARIN-NL Data Curation Service in 2013.

An actual overview of the data curation results and plans can be found on the [DCS project area](#)¹⁷ on the CLARIN-NL website. Here's a snapshot:

Completed:

- LESLLA (delivered in August 2013)
- IPNV Interviews with veterans (delivered in May 2013)
- Woordenboek Gelderse Dialecten, Rivierengebied (delivered in June 2013)
- Woordenboek Gelderse Dialecten, Veluwe part) (delivered in June 2013)
- Curation organisation names for OpenSkos (delivered in June 2013; finalised in August 2013)
- Six dialect dictionaries from Brabant (delivered in October 2013)
- WLD and WBD part III (delivered in December 2013)

In progress:

- DBD/TCULT (to be delivered in January 2014)
- Roots of Etnolect (collection Linda van Meel) (to be delivered in January 2014)
- Traces of Contact (collection Kofi Yakbo)

The DCS activities will not be continued in 2014.

As to the second Centre of Expertise, [Nederlab](#)¹⁸, on data and tools for the study of the Dutch language and culture across time, the Nederlab demonstrator project has been postponed to 2014

¹⁷ <http://www.clarin.nl/node/147>

¹⁸ <http://www.nederlab.nl/>



since higher priority had to be given to the CLARIAH Nederlab demonstrator, as was already reported in the CLARIN-NL Annual Report for 2012.

Metadata Harmonisation

Money has been reserved for a project on harmonizing all (separately and independently created) metadata records. By the end of 2013, a candidate to carry out this project was found, and the project will start by January 1, 2014.

Education and Awareness (WBS 1.7)

MinDisc A small project was carried out for the preparation and organisation of an international workshop on standards for discourse annotation of textual resources. It was closely related to and carried out by the participants of the DiscAn project. The workshop has been held, but the final technical and financial reports have not yet been received.

An overview is given of the various activities for stimulating education, training, and awareness carried out in 2012.

Conferences and workshops visited¹⁹

From	To	Event	City	Country	Who
12/04/2013	12/04/2013	CLARIN Literary Workshop	Amsterdam	NL	Many
11/26/2013	11/27/2013	Annual Meeting General Assembly	Wrocław	PL	AD, JO
10/28/2013	10/30/2013	2nd EUDAT Conference (October 28-30, Rome)	Rome	IT	DB
10/20/2013	10/22/2013	CLARIN ANNUAL Meeting	Prague	CS	Many
10/17/2013	10/18/2013	DARIAH/DASISH Access and Authentication Infrastructures Workshop	Köln	DE	DB
09/30/2013	09/30/2013	Workshop: CLARIN,	Rome	IT	IS

¹⁹ JO=Jan Odijk, DB=Daan Broeder, SB= Sjeff Barbiers, TP=Toine Pieters, AvdZ=Arwin van der Zwan, IS=Ineke Schuurman, MW=Menzo Windhouwer, AvH=Arjan van Hessen, AD = Alice Dijkstra



		Standards and the Text Encoding Initiative			
09/20/2013	09/20/2013	Digital Humanities Summerschool	Leuven	BE	JO, SB, TP
09/10/2013	09/10/2013	Open Research Data on the Horizon ... and beyond - RDA Europe Workshop prior to the RDA Plenary in Washington	Garching	DE	DB
07/18/2013	07/20/2013	Language Diversity Congress	Groningen	NL	JO, SB
05/17/2013	05/17/2013	OpenSKOS and controlled vocabularies meeting	Utrecht	NL	Hennie Brugman
04/25/2013	04/25/2013	CLARIN Centre Assessment Workshop	Copenhagen	DK	Centre representatives
04/24/2013	04/24/2013	CLARIN Federated Content Search workshop	Copenhagen	DK	JO
04/10/2013	04/10/2013	eHumanities Address Software and Tools	Amsterdam	NL	DB
03/21/2013	03/22/2013	NIAS DH Workshop	Wassenaar	NL	JO
02/28/2013	02/28/2013	SURF dag	Den Haag	NL	JO
01/23/2013	01/23/2013	The PoliMedia symposium: linking political debates and media	Amsterdam	NL	AvH
01/17/2013	01/18/2013	CLIN	Hengelo	NL	Many

At CLIN, Ineke Schuurman gave a keynote lecture on TTNWW.

At the Language Diversity Congress Jan Odijk gave a lecture on the relevance of CLARIN for comparative linguistics. At the Digital Humanities Summer School in Leuven, Jan Odijk gave a keynote



lecture, Toine Pieters gave a lecture on WAHSP/BILAND, and Sjef Barbiers gave a lecture and a hands-on session on MIMORE.

At the CLARIN Annual Meeting Sjef Barbiers demonstrated MIMORE, and Gerrit Bloothoofth demonstrated MIGMAP. There were posters by the Dutch CLARIN centres Huygens ING, MPI, Meertens, DANS, and INL.

Events organized by CLARIN-NL

From	To	Description	City	Ctry	Who
10/02/2013	10/02/2013	ISOCAT Workshop	Utrecht	NL	IS,MW
12/09/2013	12/09/2013	Semantic interoperability Workshop	Utrecht	NL	DB
10/14/2013	10/14/2013	CMDI Future Workshop	Utrecht	NL	DB
09/04/2013	09/04/2013	CLARIN-NL IAP Demonstration Event.	Utrecht	NL	JO, AvdZ
06/20/2013	06/20/2013	ISOCAT Tutorial	Utrecht	NL	IS,MW
06/04/2013	06/05/2013	CMDI Interoperability Workshop	Utrecht	NL	DB
02/21/2013	02/21/2013	Call 4 Kick-Off Meeting	Utrecht	NL	Many

Publications

Here's a (not fully complete) overview:²⁰

Odijk, J. (2013) 'Common Lab Research Infrastructure for the Arts and Humanities (CLARIAH), in *Uncharted Frontiers: the Netherlands' Roadmap for Large-Scale Research Facilities*, publication of the Dutch Ministry of Education, Culture and Science, p. 17.

Odijk, J. (2013). DUELME: Dutch Electronic Lexicon of multiword Expressions. In G. Francopoulou (Eds.), *LMF - Lexical Markup Framework* London, UK / Hoboken, US: ISTE / Wiley.

Sijs, Nicoline van der (2013) 'Open Dialectendata', NRC 2013-05-11. [[pdf](#)]

Van Uytvanck, D. (2013) 'ICT Tools enable eHumanities top-research', in *Uncharted Frontiers: the Netherlands' Roadmap for Large-Scale Research Facilities*, publication of the Dutch Ministry of Education, Culture and Science, p. 18-19.

Presentations

Here's a (not fully complete) overview:

Barbiers, S. (2013), MIMORE, presentation at the CLARIN Annual Conference, Prague, Oct 23, 2013, [[url](#)]

²⁰ For a more complete and up-to-date overview, see <http://www.clarin.nl/node/1431>



Bloothoof, G. (2013), MIGMAP, presentation at the CLARIN Annual Conference, Prague, Oct 22, 2013, [[url](#)]

Broeder, D. (2013). CMDI registry. poster at the CLARIN Annual Conference, Prague, Oct 23, 2013.

Filarski, G. (2013), eLaborate/ePistolarium, poster at the CLARIN Annual Conference, Prague, Oct 23, 2013.

Kemps-Snijders, M. (2013). CMDI Search, poster at the CLARIN Annual Conference, Prague, Oct 23, 2013.

Kemps-Snijders, M. (2013). TTNWW, poster at the CLARIN Annual Conference, Prague, Oct 23, 2013.

Kok, W. (2013), BlackLab: A corpus retrieval engine, poster at the CLARIN Annual Conference, Prague, Oct 23, 2013.

Odijk, J. (2013, Feb 21). *CLARIN-NL Call 4 Kick-Off Introduction*. Utrecht, CLARIN-NL Call 4 Kick-Off. [[ppt](#)]

Odijk, J. (2013), 'Example Queries for Federated Search', use case for the CLARIN Federated Content Search Workshop, Copenhagen, April 24, 2013. [[docx](#)]

Odijk, J. (2013), 'Example Queries for Federated Search', presentation held at the CLARIN Federated Content Search Workshop, Copenhagen, April 24, 2013 [[pptx](#)]

Odijk, J. (2013), "CLARIAH", presentation held at the EuroRisNet+ workshop *Research Infrastructures towards 2020*, Lisbon, July 5, 2013. [[pptx](#)] [[URL](#)]

Odijk, J. (2013), "Comparative Linguistic Research in the CLARIN Infrastructure", presentation held at the Language Diversity Congress, Groningen. 18 July 2013. [[pptx](#)]

Odijk, J. (2013) "Interim Evaluation", presentation held at the IAP Demonstrator Event, Utrecht September 4, 2013 [[pptx](#)]

Odijk, J. (2013) 'The CLARIN Infrastructure (NL Part)' presentation held at the IAP Demonstrator Event, Utrecht September 4, 2013 [[pptx](#)]

Odijk, J. (2013) 'The CLARIN Infrastructure (NL Part): Current State and Near Future', key note presentation held at the Digital Humanities Summer School, Leuven, September 20, 2013. [[pptx](#)]

Westerhout, E. & J. Odijk (2013), 'Metadata for tools: creating a CMDI profile for tools', presentation held at CLIN 2013, Enschede, Jan 18, 2013. [[pdf](#)]

Windhouwer, M. (2013), ISOCAT. poster at the CLARIN Annual Conference, Prague, Oct 23, 2013.

Support for other events

From	To	Event	City	Form of
------	----	-------	------	---------



				Support
12/04/2013	12/04/2013	CLARIN Literary Workshop	Amsterdam	financial
01/17/2013	01/18/2013	CLIN	Hengelo	financial

Support for travel of Dutch researchers to CLARIN-related events

Date	Description	City	Country	Who	Organization
------	-------------	------	---------	-----	--------------

CLARIN- NL Newsflashes

CLARIN- NL Newsflashes have been sent out to all registered persons on the following dates:

- [NewsFlash 28 \(16-07-2013\)](#)
- [NewsFlash 27 \(11-02-2013\)](#)

All Newsflashes can be accessed via <http://www.clarin.nl/node/82>.

CLARIN-EU Newsflashes

In 2013, no CLARIN (ERIC) newsflashes have appeared.

CLARIN-EU Newsletter

In 2013 no CLARIN (ERIC) Newsletters have appeared.

eData & Research

CLARIN-NL has continued its participation in the eData & Research Magazine in 2013. CLARIN-NL finds the existence of this magazine important, and aims to make itself more visible by actively participating in its board and editorial team and co-financing it. Several articles on CLARIN-NL have appeared in [eData&Research](#). On behalf of CLARIN-NL, Erica Renckens participates in the editorial team and Jan Odijk in the board. As of 2014, CLARIN-NL will, for financial reasons, not participate anymore, but it has been replaced by the CLARIAH-SEED project.

The articles that have been published in eData & Research in 2013 can be accessed via <http://www.clarin.nl/node/188>.

Movies and Screen Captures

Several projects have made movies or screen captures to illustrate their work or their demonstrators. See <http://www.clarin.nl/node/185> for a list of such videos.

Project	Link to Movie / Screen capture
GabMap	http://www.youtube.com/watch?v=anElatt5fFM&feature=plcp
Interviews / Veteran Tapes	http://www.youtube.com/watch?v=VpCXh_NVVso



WFT-GTB	http://www.youtube.com/watch?v=fevAPknCaDo&feature=player_embedded
---------	---

The EB had already commissioned the production of 6 short movies on CLARIN-NL projects (to Dyzlo). Below one can find the CLARIN-NL projects that have been selected for these movies. All movies are available.

All movies and screen captures can be reached via <http://www.clarin.nl/node/403>

Project	Link to Movie
MIMORE	http://www.youtube.com/watch?v=TBDGrqKYbIA
SignLinc	http://www.youtube.com/watch?v=je3ZKPw6w5s
Arthurian Fiction	http://www.youtube.com/watch?v=t0y-ecrVD5w
WFT-GTB	http://www.youtube.com/watch?v=aSWKJucbWY8
WIP	http://www.youtube.com/watch?v=tEqAIH2o2rc&feature=plcp
Search & Develop	http://www.youtube.com/watch?v=6CkCfaj7rao&feature=plcp

The movies have been made available in streaming mode via YouTube, have subtitles (at least in English) and the originals have been made available on solid state memory.

For 2013/2014 three new movies were commissioned, only one of which was finished and made available in 2013:

Project	Link to Movie
CKCC	http://www.youtube.com/watch?v=SI0lg7SCOcc
PILNAR	Expected mid 2014
Polimedia	Expected mid 2014

Website

The website was upgraded to Drupal 7 by an external company (*ClearWind*, now renamed to *Civinex*), which also takes care of the maintenance of the website.

The CLARIN-NL website is a project-internal website. We also aimed to add an external website, with nice lay-out and pictures, with information on the project targeted at people external to the project, and accompanied with regular PR activities (news messages, showcases, etc.). We have commissioned some work on this (design, editorial work), which has been carried out. This external part of the website will become part of the CLARIN-NL portal.

For the CLARIN-NL portal, preparatory work has been carried out and it is expected to go live Mid 2014.



Education

A special project for the selection of topics for educational modules was started up and finished. Currently, the selected educational modules are being implemented.

Various tutorials have been given on CMDI and ISOCAT (see overview events organized by CLARIN-NL). For contributions to seasonal schools and other tutorials, see above (section on conferences and workshops)

Helpdesk

The CLARIN-NL Helpdesk²¹ has run all year. The amount of requests is rather low, and therefore the Helpdesk was easily managed in 2013 by Jan Odijk and Daan Broeder.

A document explaining the concept "CLARIN-compatible" has been written²² In 2011 (originally specifically targeted at KNAW researchers). In 2013, a translation of the document in English has appeared.

Though it was planned to hold a meeting with the relevant researchers, inviting them to relevant tutorials, etc. to ensure that they indeed understand what it means to produce and deliver resources in a CLARIN-compatible manner, and actually apply this knowledge, this still has to be done and is an important ingredient of the plan for 2014.

National Coordination Point

The national coordination point has functioned well in 2013. Prof.dr. Jan Odijk (1FTE) is the Programme Director. Drs. Jolien Scholten was the project secretary (.5 FTE) until March 1, 2013, and she has been replaced since February 1, 2013 by Drs. Arwin van der Zwan.

The governance structure consists of an Executive Board (EB), Board, National Advisory Panel (NAP) and International Advisory Panel (IAP). Their tasks and responsibilities have been laid down in the consortium agreement.

Executive Board The Executive Board consists of 4 members:

- Prof.dr. Jan Odijk, programme director
- Prof.dr. Hans Bennis, chair and representative of the humanities researchers
- Dr. Arjan van Hessen, responsible for complementary activities such as PR, education, training and awareness creation.
- Ir. Daan Broeder, technical director

Apart for the programme director (1FTE), the members of the EB have a 1/5FTE appointment to carry out their tasks.

The executive board has held regular meetings (once every two weeks) to formulate, prepare and implement policy procedures and concrete activities.

²¹ <http://trac.clarin.nl/trac>

²² <http://trac.clarin.nl/trac/attachment/wiki/WikiStart/CLARIN%20compatible%20NL%20110805.pdf>



Board The Board consists of 8 members. A full list of its members can be found in Appendix A.

The current composition of the Board can also be found on the CLARIN-NL website (<http://www.clarin.nl/category/board/board>)

The Board has held 1 meeting in 2013 (March 30, 2013).

NAP The National Advisory Panel consists of 16 members, providing a good representation of the fields of the humanities, language and speech processing and infrastructure design and implementation. A full list of the members can be found in Appendix B

The current composition of the CLARIN-NL NAP can also be found on the CLARIN-NL website (<http://www.clarin.nl/category/board/national-advisory-panel>).

The NAP has held 1 meeting in 2013 (October 8, 2013).

IAP The International Advisory Panel consists of 11 members. A full list of the members can be found in Appendix C and on the CLARIN-NL website (<http://www.clarin.nl/category/board/international-advisory-panel>).

The IAP has held 1 meeting in 2013, as part of the CLARIN-NL Interim Evaluation.

Consortium Agreement The consortium agreement, which was approved by the Board on Nov 2, 2009 and which was initially signed by 21 partners, has remained unchanged. The consortium is open to new partners, and several partners have joined the consortium in 2012, bringing the total number of partners to 33 (see <http://www.clarin.nl/node/7>). The consortium agreement can be found on the CLARIN-NL website (login required): <http://www.clarin.nl/node/72>.

CLARIN-NL Long Term Working Plan The CLARIN-NL Long Term Working Plan was approved by the Board on August 19, 2009, and has remained unchanged. It can be found on the CLARIN-NL website: <http://www.clarin.nl/system/files/CLARIN-NL%20Multiyear%20Programme%20090409-2.pdf>

Working Plan 2013 The Working Plan for 2012 was approved by the Board. It can be found on the CLARIN-NL website (<http://www.clarin.nl/page/about/47>)

Overall, the governance structure as set up functions appropriately.

4 2012 Finances

A detailed report on the CLARIN-NL finances in 2013 can be found in the 2013 Financial Report.

5 Other Activities

As a response to the invitation of the *Commissie voor de update van de Nationale Roadmap voor Grootschalige Onderzoeksfaciliteiten* (Committee for the update of the National Roadmap for Large Scale Research Facilities), CLARIN-NL and the Dutch players in DARIAH have joined forces and formulated a new proposal for a large scale national infrastructural project called *Common Lab Research Infrastructure for the Arts and the Humanities (CLARIAH)*, uniting all humanities organizations in the Netherlands. The project has been put on the Roadmap and has received 1



CLARIN-NL Annual Report 2013

million euro as 'seed capital', to maintain the dynamics of the consortium and to prepare a new, improved, proposal. The CLARIAH-SEED project has developed plans and executed most of them in 2013, though the project continues until January 1, 2015. On October 1, 2013, a revised proposal has been submitted, which is currently being evaluated. Details about it can be found on the CLARIAH website (www.clariah.nl)



Appendix A Composition of the Board

Name	Affiliation
Prof.dr. Geert Booij (until Sep 1, 2012)	LUCL, Universiteit van Leiden
Prof.dr. Lou Boves	CLST, Radboud Universiteit Nijmegen
Dr. Peter Doorn	DANS
Prof.dr. Martin Everaert	UIL-OTS, Universiteit van Utrecht
Prof.dr. Jaap van den Herik	TiCC, Universiteit van Tilburg
Prof.dr. Aafke Hulk (until Sep 1, 2012)	ACLIC, Universiteit van Amsterdam
Prof.dr.ir John Nerbonne	CLCG, Universiteit van Groningen
Prof.dr. Els Stronks (as of Sep 1, 2012)	Universiteit van Utrecht
Ir. Peter Wittenburg (as of Sep 1, 2012)	Max Planck Instituut voor Psycholinguïstiek, Nijmegen

Appendix B Composition of the NAP

Name	Affiliation
Prof.dr. Willem Adelaar (until Sep 1, 2012)	LUCL, Universiteit van Leiden
Prof.dr. Sjef Barbiers	Meertens Instituut
Dr. Jeannine Beeken	INL
Prof.dr Antal van den Bosch	Radboud Universiteit Nijmegen
Dr. Gosse Bouma	Universiteit van Groningen
Drs. Hennie Brugman	Meertens Instituut
Prof.dr. Karina van Dalen-Oskam	Huygens Instituut voor Nederlandse Geschiedenis
Dr. Paul Doorenbosch	Koninklijke Bibliotheek
Dr. Willemijn Heeren	Universiteit van Twente
Ir. Marc Kemps-Snijders	Meertens Instituut
Prof.dr. Kees Mandemakers	Internationaal Instituut voor Sociale Geschiedenis
Dr. Nelleke Oostdijk	CLST, Radboud Universiteit Nijmegen
Prof.dr. Reinier Salverda	Fryske Akademy
Prof.dr. Ted Sanders (until Sep 1, 2012)	UiL-OTS, Universiteit Utrecht
Prof.dr. Piek Vossen	Vrije Universiteit Amsterdam
Prof.dr. Sally Wyatt (until Sep 1, 2012)	eHumanities (KNAW)
Dr. Joris van Zundert	Huygens Instituut voor Nederlandse Geschiedenis
Prof.dr. Arjen Versloot (as of Sep 1, 2012)	Universiteit van Amsterdam
Prof.dr. Toine Pieters (as of Sep 1, 2012)	Universiteit van Utrecht



Appendix C Composition of the IAP

Name	Affiliation	Country
Prof.dr Scott Farrar	University of Washington, Department of Linguistics	USA
Prof.dr David Hoover	New York University	USA
Prof.dr. Clifford Lynch	Coalition for Networked Information	USA
Prof.dr. Harold Short	King's College London, Centre for Computing in the Humanities	UK
Prof.dr. Benjamin T'sou	City University of Hong Kong	China
Prof.dr Hugo Van hamme	Katholieke Universiteit Leuven	Belgium
Prof.dr. Tamás Váradi	Linguistics Institute, Hungarian Academy of Sciences	Hungary
Prof.dr. Koenraad De Smedt	University of Bergen	Norway
Prof.dr. Erhard Hinrichs	University of Tübingen	Germany
Prof.dr. Jan Hajič	Charles University Prague	Czech republic

Appendix D Acronyms

Acronym / Term (URL)	Type	Expansion	Dutch expansion
@PhilosTEI	CLARIN-NL subproject	TICCLing Philosophy: a TEI corpus-building workflow towards a new computational methodology for philosophy	
AAI	software	Authentication and Authorisation Infrastructure	
AAM-LR	CLARIN-NL subproject	Automatic Annotation of Multi-modal Language Resources	
ACLCL	organisation	Amsterdam Center for Language and Communication	
Adelheid	CLARIN-NL subproject	A Distributed Lemmatizer for Historical Dutch	
ADEPT	CLARIN-NL subproject	Assaying Differences via Edit-Distance of Pronunciation Transcriptions	
AGORA	project		
Alfalab	project		
Anéla	organisation	Dutch Association for Applied Linguistics	
ANNEX	software	Annotation Explorer	
ANP	company	National Netherlands News Agency	Algemeen Nederlands Persbureau
ANP radio bulletin collection	data	ANP radio bulletin collection	
ANWB	organisation	Dutch automobile	Algemene Nederlandse



		association	Wielrijdersbond
API	term		
ARBIL	software		
Arthurian Fiction	CLARIN-NL subproject	Arthurian Fiction in Medieval Europe	
Asymenc	project	Asymmetrical Encounters: Digital Humanities Approaches to Reference Cultures in Europe, 1815–1992; HERA Project	
AV	term	Audio-Visual	
AVatech	project	Advancing Video Audio Technology in Humanities Research	
AXES	project	Access to AudioVisual Archives	
BG	organisation	Netherlands Institute for Sound and Vision	Instituut voor Beeld en Geluid
BIGGrid	project	Dutch eScience Grid	
BILAND	CLARIN-NL subproject	Towards a flexible and stable CLARIN-supported web-application for bilingual historical analysis of discourses in news media	
BioPort	data	Biographic Portal	Biografisch Portaal
Birth of TV	project	Birth Television Archive	
BlackLab	software		



BLARK	term	Basic Language Resource Kit	
BRIDGE	project	Building Rich Links to Enable Television History Research	
BSON	software	Binary JSON	
CATCH	project	NWO Programme Continuous Access To Cultural Heritage	
CATCHPlus	project		
CDH	organisation	Centre for Digital Humanities	
C-DSD	CLARIN-NL subproject	Curating the Dutch Song Database	
CELEX	data	Dutch Centre for Lexical Information	
Center for Language Studies	organisation	Centre for Language Studies	
CEPR	organisation	Center for Economic policy Research	
CESSDA	infrastructure	Council of European Social Science Data Archives	
CGN	data	Spoken Dutch Corpus	Corpus Gesproken Nederlands
CHAT	organisation	Centre for Humanities and Technology	
CHILDES	infrastructure	Child Language Data Exchange System	
CHIP	project	Project in the CATCH programme	
CHOICE	project	Project in the CATCH programme	



CineXPRES	project		
CIO	term	Chief Integration Officer	
CKCC	project	<i>Circulation of Knowledge and Learned Practices in the 17th-century Dutch Republic,</i>	Geleerdenbrievenproject
CLAM	software	Computational Linguistics Application Mediator	
CLAPOP	CLARIN-NL subproject	CLARIN-NL Portal Project	
CLARA	project	Common Language Resources and their Applications	
CLARIAH	infrastructure	Common Lab Infrastructure for the Humanities	
CLARIAH-SEED	infrastructure project	CLARIAH Seed Capital project	
CLARIAH-SEED	infrastructure project	CLARIAH Seed Capital project	
CLARIN	infrastructure	The CLARIN infrastructure	
CLARIN ERIC	organisation	CLARIN ERIC	
CLARIN-EU	infrastructure project	The European CLARIN enterprise	
CLARIN-NL	infrastructure project	The Netherlands national project for CLARIN	
CLARIN-PP	infrastructure project	The European CLARIN preparatory phase project	
CLARIN-prep	infrastructure project	The European CLARIN preparatory phase project	
CLAVAS	CLARIN-NL subproject	CLARIN Vocabulary	



		Service	
CLAVAS	software	CLARIN Vocabulary Service	
CLCG	organisation	Center for Language and Cognition Groningen	
CLICKNL	organisation	Creativity, Learning, Innovation, Co-creation, Knowledge	
CLIF	organisation	Computational Linguistics in Flanders	
CLIN	conference	Computational Linguistics in the Netherlands (local conference)	
Clio Infra	infrastructure		
CLIO-DAP	project		
CLS	organisation	Centre for Language Studies	
CLST	organisation	Centre for Language and Speech Technology	
CLUE	organisation	Research institute for the heritage and history of the Cultural Landscape and Urban Environment	
CMDI	software	Component MetaData Infrastructure	
COAVA	CLARIN-NL subproject	Cognition, Acquisition and Variation Tool	
COBWWEB	CLARIN-NL subproject	Connections Between Women and	



		Writings Within European Borders	
COGIS	organisation	Dutch expert centre on the (psycho)social effects of war, persecution, aggression and violence	Kennisinstituut sociale en psychische gevolgen van oorlog, vervolging en geweld
ColTime	CLARIN-NL subproject	Collaboration on time-based resources	
Comerda	software		
COMMIT	project		
COREX	software	Corpus Exploitation Tool	
Cornetto	data		
CORNETTO-LMF-RDF	CLARIN-NL subproject	Curated Cornetto database in LMF and RDF	
CPG	organisation	Centre for Parliamentary History	Centrum voor Parlementaire Geschiedenis
CTO	term	Chief Technical Officer	
Culturele Dynamiek	grant programme	NWO grant programme: Cultural Dynamics	
D3	software	Data Driven Documents	
DAM-LR	project	Distributed Access Management for Language Resources	
DANS	organisation	Data Archiving and Networking Services	
DARIAH	infrastructure	Digital Research Infrastructure for the Arts and Humanities	
DARIAH-EU	infrastructure project	The European DARIAH	



		enterprise	
DARIAH-NL	infrastructure project	DARIAH in the Netherlands	
DARIAH-PP	infrastructure project	the DARIAH Preparatory Phase Project	
DASISH	infrastructure project	Data Service Infrastructure for the Social Sciences and the Humanities	
DBD	data	Dutch Bilingual Database	
DBNL	organisation	Digital Library for Dutch Literature	
D-COI	data		
DCR	data	Data Category Registry	
DCS	CLARIN-NL subproject	Data Curation Service	
DDS	software	Data Deposition Service	
DEN	organisation	Digital Heritage Netherlands	Kenniscentrum Digitaal Erfgoed
DH	term	Digital Humanities	
DiDDD	data	Diversity in Dutch DP Design	
DiscAn	CLARIN-NL subproject	Towards a Discourse Annotation system for Dutch language corpora	
DLU	organisation	Dutch Language Union	
D-LUCEA	CLARIN-NL subproject	Database of the Longitudinal Utrecht Collection of English Accents	
DOBES	project	Documentation of Endangered Languages	
DSD	data	Dutch Song Database	Liederbank



DSS	CLARIN-NL subproject	Dutch Ships and Seamen	
DUELME	data	Dutch Electronic Lexicon of Multiword Expressions	
DUELME-LMF	CLARIN-NL subproject	Converting DUELME into LMF format	
Dynamics of Memory	grant programme	NWO grant programme: Dynamics of Memory	
DynaSAND	software	Dynamic Syntactic Atlas of the Dutch Dialects	Dynamische Syntactische Atlas van de Nederlandse Dialecten
EAGLES	project	Expert Advisory Group on Language Engineering Standards	
EB	term	Executive Board	
e-BNM	data	Bibliotheca Neerlandica Manuscripta database	
eBNM+	CLARIN-NL subproject	Linked Data on Middle Dutch Sources Kept Worldwide	
EC	organisation	European Commission	
EDBO	data	Early Dutch Books Online	
EDBO	data	Early Dutch Books Online	
eHg	organisation	eHumanities Group	
EHPS-Net	network	European Historical Population Samples Network	
EHRI	infrastructure project	European Holocaust Research	



		Infrastructure	
eHumanities Group	organisation	KNAW eHumanities Group	
EIP	instrument	European Innovation Partnerships	
ELAN	software	EUDICO Linguistic Annotator (Annotation Tool)	
ElasticSearch	software	open source, distributed real- time search and analytics engine for the cloud	
EMIT-X	CLARIN-NL subproject	Early-Modern Image and Text eXchange	
EPIC	consortium	European Persistent Identifier Consortium	
ePistolarium	software		
ERC	organisation	European Research Council	
ERIC	organisation	European Research Infrastructure Consortium	
ESF EURYI	grant programme	European Science Foundation European Young Investigator Awards	
ESFRI	organisation	European Strategy Forum on Research Infrastructures	
EU	organisation	European Union	
EUDAT	project	European Data Infrastructure	
EUR	organisation	Erasmus University	



		Rotterdam	
European Film Gateway	organisation		
Europeana	data		
EURYI	grant programme	European Science Foundation European Young Investigator Awards	
EUScreen	project		
EXILSEA	CLARIN-NL subproject	Exploiting ISOcat's Language Sections in ELAN and ANNEX	
eXist	software	Native XML database	
EYE	organisation	EYE Film Museum Amsterdam	
FA	organisation	Fryske Akademy	
FESLI	CLARIN-NL subproject	Functional Elements in Specific Language Impairment	
FIDLR	project	Framework for the Interoperability of Dutch Language Resources	
FIDLR-Start	project	Project to prepare the FIDLR project proposal	
FlareNet	project	Fostering Language Resources Network, European Network project.	
FlockDB	software	Distributed Fault-Tolerant Graph Database	
FP7	programme	Framework Programme 7	



		(European Union)	
Frontwise	company		
FTE	term	Full-Time Equivalent	
GIA	organisation	Groningen Institute of Archaeology	
GIS	term	Geographic Information System	
Global labour	project	Global Collaboratory on the History of Labour Relations 1500-2000	
Gridline	company		
GrNe	CLARIN-NL subproject	Greek-Dutch dictionary	Grieks-Nederlands woordenboek
GTB	software	Integrated Language Database	Geïntegreerde Taalbank
GTRP	project	Goeman, Taeldeman, van Reenen Project	
HERA	organisation	Humanities in the European Research Area	
HI	organisation	Huygens ING Institute	
HI-POD	project	Historical Patterns of Development	
HLT	term	Human Language Technologies	
HLZ	project	HSN LINKS Zeeland	
HSN	infrastructure	Historical Sample of the Netherlands	
HSS	term	Humanities and Social Sciences	
Huygens ING	organisation	Huygens ING Institute	
HvA	organisation	Hogeschool van Amsterdam	



IAP	term	International Advisory Panel	
IARPA SESAME Project	project	Video SEArch with Speed and Accuracy for Multimedia Events	
IBM	company	International Business Machines	
IBM NL	company	IBM Netherlands	
ICON UU	organisation	Institute for Cultural Inquiry Utrecht University	
ICORDI	project		
ICPSR	organisation	Interuniversity Consortium for Political and Social Research	
ICT	term	Information and Communication Technology	
ICT Roadmap for the Top Sectors	roadmap		
ID	term	Identifier	
IIAV	organisation	International Information Centre and Archive for the Women's Movement	Internationaal Informatiecentrum en Archief voor de Vrouwenbeweging
IIP	CLARIN-NL subproject	Infrastructure Implementation Project	
IISG	organisation	International Institute for Social History	Internationaal Instituut voor Sociale Geschiedenis
IISH	organisation	International Institute for Social History	Internationaal Instituut voor Sociale Geschiedenis
Images for the Future	project		Beelden voor de Toekomst
IMDI	software	ISLE MetaData Initiative	
IMIX	project	NWO programme for Interactive	



		Multimodal Information eXtraction	
iMmovator	network	Cross media Network	
INL	organisation	Institute for Dutch Lexicology	Instituut voor Nederlandse Lexicologie
INPOLDER	CLARIN-NL subproject	Integrated Parser and Lemmatizer Dutch in Retrospect	
Institute for Historical Literature and Cultural Studies	organisation		
INTERA	project	Integrated European language data Repository Area	
Interop	project	NSF Community-Based Data Interoperability Networks Programme	
INTER-VIEWS	CLARIN-NL subproject	Curation of Interview Data	
IOP	programme	Innovation-targeted Research Programmes	Innovatiegerichte Onderzoeksprogramma's
IPNV	data	Interview Project Nederlandse Veteranen	Interview Project Dutch Veterans
IPR	term	Intellectual Property Rights	
IPROSLA	CLARIN-NL subproject	Integrating and publishing resources on sign language acquisition	
ISLA	organisation	Intelligent Systems Lab Amsterdam	
ISLE	project	International Standards for Language Engineering	

ISO	organisation	International Standards Organisation	
ISO TC37/SC4	standard	ISO Subcommittee focusing on Language Resource Management	
ISOcat	data	Project to implement the ISO DCR standard	
ISOCAT	software	ISO Data Category Registry	
JSON	software	JavaScript Object Notation	
KB	organisation	National Library of the Netherlands	
KB Newspaper Archive	data	National Library Newspaper Archive	
KDC	organisation	Catholic Documentation Centre	Katholiek Documentatie Centrum
KITLV	organisation	Royal Netherlands Institute of Southeast Asian and Caribbean Studies	Koninklijk Instituut voor Taal-, Land- en Volkenkunde
KNAW	organisation	Royal Netherlands Academy of Arts and Sciences	
KNAW Over Grenzen	KNAW grant programme	KNAW Crossing Boundaries	KNAW Over Grenzen
Knowledge Concepts	company		
LAISEANG	CLARIN-NL subproject	Language Archive of Insular South East Asia and West New Guinea	



LDE CGHD	organisation	Leiden Delft Erasmus Center for Global Heritage and Development	
LESLLA	data	Low Educated Second Language and Literacy Acquisition	
Leverhulme Trust	grant programme		
LINKS	project		
LinkTV	project		
LIRICS	project	Linguistic Infrastructure for Interoperable Resources and Systems	
LMF	standard	Lexical Markup Framework	
LMF	standard	Lexical Markup Framework	
Lorentz	organisation	Lorentz Center	
LOT	organisation	Netherlands Graduate School of Linguistics	Landelijke Onderzoeksschool Taalkunde
LREC	conference	Language Resources and Evaluation Conference	
LRT	term	Language Resources and Technologies	
LT	term	Language Technology	
LUCAS	organisation	Leiden University Center for the Arts in Society	
LUCL	organisation	Leiden University Centre for Linguistics	
Maddison Project	project		
MAND	data	Morphological	



		Atlas of the Dutch Dialects	
MD	CLARIN-NL subproject	Metadata Project	
MD4T	CLARIN-NL subproject	Metadata for Tools Project	
Meertens	organisation	Meertens Institute	
Meertens Institute	organisation	Meertens Institute	
Merdes	software	Media Researchers Data Exploration Suite	
META-NET	project		
META-SHARE	sharing facility		
METS	standard	Metadata Encoding and Transmission Standard	
MI	organisation	Meertens Institute	
MIGMAP	CLARIN-NL subproject	Interactive migration maps for the 20th century	
MIMORE	CLARIN-NL subproject	Microcomparative Morphosyntax Research Tool	
Mindisc	CLARIN-NL subproject	Minimal Set for Discourse Annotation	
MPG	organisation	Max-Planck-Gesellschaft	
MPI	organisation	Max Planck Institute for Psycholinguistics	
MT	term	Management Team	
MU	organisation	Maastricht University	
Multicon	CLARIN-NL subproject	Multilayer Concordance Functions in ELAN and ANNEX	



MuSeUM	project	Project in the CATCH programme	
NA	organisation	National Archive	Nationaal Archief
Namespace	CLARIN-NL subproject	Mapping the Landscape of Names in Modern Dutch Literature	
NAP	committee	National Advisory Panel	
Nederlab	project		
NeDiMAH	network	Network for Digital Methods in the Arts and Humanities	
NEHOL	CLARIN-NL subproject	Negerhollands Database	
NGT	data	Dutch Sign Language Corpus	Corpus van de Nederlandse Gebarentaal
NIAS	organisation	Netherlands Institute for Advanced Study in the Humanities and Social Sciences	
NIBG	organisation	Netherlands Institute for Sound and Vision	Nederlands Instituut voor Beeld en Geluid
NIOD	organisation	Netherlands Institute for War Documentation	Nederlands Instituut voor Oorlogsdocumentatie
NISV	organisation	Netherlands Institute for Sound and Vision	Instituut voor Beeld en Geluid
NL	country	the Netherlands	
NLeSc	organisation	Netherlands eScience Centre	
NLP	term	Natural Language Processing	
Notas	organisation	Dutch Organisation for Language and	Nederlandse Organisatie voor Taal- en Spraaktechnologie



		Speech Technology	
Noterik	company	Noterik BV	
NSF	organisation	National Science Foundation	
NWO	organisation	Netherlands Organisation for Scientific Research	
NWO Creative Industry	grant programme	NWO grant programme: Creative Industry	
NWO Open Competition	grant programme	NWO grant programme: Open Competition	NWO Open Competitie
NWO-Groot	grant programme	NWO Infrastructure Investment Fund	
OAI	organisation	Open Archives Initiative	
OAI/PMH	standard	Protocol for Metadata Harvesting	
OAuth2	software	Oauth 2.0	
OCR	term	Optical Character Recognition	
OCROPUS	software	OCROPUS	
OHT	project	Oral History Today	
OLAC	standard	Open Languages Archive Community	
OpenSONAR	project		
OpenSONAR	CLARIN-NL subproject	Online Personal Exploration and Navigation of SoNaR	
OWL	standard	OWL Web Ontology Language	
PC	organisation	Programme Committee	

PCM	company	large publisher in the Netherlands (Perscombinatie Meulenhoff & Co)	
Picturae	company		
PID	term	Persistent Identifier	
PILNAR	CLARIN-NL subproject	Pilgrimage Narratives: creating a germ corpus for studying the profile of the modern pilgrim	
Pionier Grant	grant programme	specific type of grant awarded by NWO in the past	
PMH	standard	Protocol for Metadata Harvesting	
Polimedia	CLARIN-NL subproject	Interlinking multimedia for the analysis of media coverage of political debates	
POS	term	Part of Speech	
Posthumus Institute	organisation	National Graduate School the Posthumus Institute	Landelijke Onderzoeksschool het Posthumus Instituut
PR	term	Public Relations	
QuaMerdes	CLARIN-NL subproject	Quantitative Content Analysis of Media Researchers' Data	
R&D	term	Research and Development	
RCE	organisation	Cultural Heritage Agency	
RDA	project	Research Data Alliance	
RDF	standard	Resource	



		Description Framework	
RDF/SKOS	standard	RDF / Simple Knowledge Organisation System	
Readspeaker	company		
Redis	software	Open Source Key-Value Store	
RELcat	software	Relation Registry	
RemBench	CLARIN-NL subproject	A Digital Workbench for Rembrandt Research	
REST	standard	Representational State Transfer	
RKD	organisation	Netherlands Institute for Art History	
RMES	organisation	Research School for Media Studies	
RU	organisation	Radboud University Nijmegen	
RUG	organisation	Groningen University	Rijksuniversiteit Groningen
RUN	organisation	Radboud University Nijmegen	
S&D	CLARIN-NL subproject	Search & Develop	
SAMPA	standard	Speech Assessment Methods Phonetic Alphabet	
SAND	data	Syntactic Atlas of the Dutch Dialects	Syntactische Atlas van de Nederlandse Dialecten
SARA	organisation	Foundation Academic Computing Centre Amsterdam	
SCHEMAcat	software	Schema Registry	



SEALINCMedia	project	Socially-enriched access to linked cultural media	
Seecr	company	Software Craftmanship	
SHEBANQ	CLARIN-NL subproject	System for HEBrew Text: ANnotation for Queries and Markup	
SignLinC	CLARIN-NL subproject	Linking lexical databases and annotated corpora of signed languages	
SKOS	standard	Simple Knowledge Organization System	
SLI	term	Specific Language Impairment	
SME	term	Small and Medium Enterprises	
SO	term	Support Office	
SOA	standard	Service Oriented Architecture	
SONAR	data		
SPARQL	software	Query Language for RDF	
SPEX	organisation	Speech Processing Expertise Centre (RU, Nijmegen)	
SPIN	grant programme	NWO Programme in the 90s	
SPINlab	organisation	Spatial Information laboratory VU	
Spinoza Grant	grant	specific type of grant awarded by NWO	
Spinque	company		



SQL	standard	Structured Query Language	
STEVIN	project	Dutch-Flemish programme for realizing the BLARK and HLT research for Dutch	
STITCH	project	Project in the CATCH programme	
SURF	organisation	SURF	
SURFnet	organisation	University Network Services Organisation	
SURFsara	organisation		
SWOT	term	Strength Weakness Opportunity Threat Analysis	
Target Holding	company		
TB	term	Technical Board	
TBDL	organisation	Tilburg Big Data Lab	Tilburg Big Data Lab
TCULT	project	Languages and Cultures in the Utrecht districts Lombok and Transvaal	Talen en Culturen in het Utrechts Lombok en Transvaal
TDS	software	Typological Database System	
TDS Curator	CLARIN-NL subproject	A web-services architecture to curate the Typological Database System	
Teezir	company		
TEI	standard	Text Encoding Initiative	
Telecats	company		
TERENA	project	Trans-European Research and Education Networking	



		Association	
Thaesis	company		
TiCC	organisation	Tilburg Centre for Cognition and Communication	
TICCLops	CLARIN-NL subproject	Text-Induced Corpus Clean-up online processing system	
Tilburg University School of Humanities	organisation		
TLA	organisation	The Language Archive	
TQE	CLARIN-NL subproject	Transcription Quality Evaluation	
TransAtlantis	project	NWO Horizon Programme about the influence of the US on public debates in the Netherlands	
TROVA	software		
TROVe	software	TransMedia Observatory	
TTNWW	CLARIN-NL subproject	LRT Tools for Dutch as web services in Work flows	TST Tools voor het Nederlands als Web services in Work flows
TU	organisation	Twente University CTIT	
Tweede Kamer BIP	organisation	Dutch Parliament Office for Automation and Projects	Tweede Kamer der Staten Generaal Bureau voor Informatisering en Projecten
Tweede Kamer DIV	organisation	Dutch Parliament Information Service	Tweede Kamer der Staten Generaal Dienst InformatieVoorziening
UBU	organisation	Utrecht University Library	



UC Davis	organisation	University California Davis	
UIL-OTS	organisation	Utrecht Institute of Linguistics OTS	
UK	country	United Kingdom	
UM	organisation	University of Maastricht	
URL	term	Universal Resource Locator	
USA	country	United States of America	
UT	organisation	University of Twente	
UU	organisation	Utrecht University	
UvA	organisation	University of Amsterdam	
VALID	CLARIN-NL subproject	Vulnerability in Acquisition: Language Impairments in Dutch. Curating five valuable data sets	
Vancis	company		
VCC	organisation	Virtual Competency Centre	
VerteldVerleden	project	Open Platform for Oral History	
VICI	grant programme	specific type of grant awarded by NWO	
Video Active	project		
VIDI	grant programme	specific type of grant awarded by NWO	
VK	CLARIN-NL subproject	Enriched Kingdom	Verrijkt Koninkrijk
VKS	organisation	Virtual Knowledge Studio	
VL	region	Flanders	Vlaanderen
VLO	software	Virtual Language Observatory	



VRE	software	Virtual Research Environment	
VSNU	organisation	Association of Collaborating Dutch Universities	Vereniging van Universiteiten
VU	organisation	VU University of Amsterdam	
VU-DNC	CLARIN-NL subproject	VU Diachronic Newspaper Corpus	
VVV	organisation	Touristic Office	
WAHSP	CLARIN-NL subproject	Web-application for historical sentiment mining in public media	
WAYF	standard	Where Are You From	
WBD	data	Brabant Dialects Dictionary	Woordenboek van de Brabantse Dialecten
WBS	term	administrative code for Utrecht University	
WFT	data	Dictionary of the Frisian Language	Wurdboek fan 'e Fryske Taal
WFT-GTB	CLARIN-NL subproject	Integrating the Wurdboek fan 'e Fryske Taal into the Geïntegreerde Taalbank	
WIP	CLARIN-NL subproject	War in Parliament	
WIVU database	data	Hebrew Bible Database constructed by the Werkgroep Informatica (WI) of the Free University (VU)	
WLD	data	Limburg Dialects Dictionary	Woordenboek van de Limburgse Dialecten
WMTC	organisation	Netherlands Research School of Science, Technology and	



		Modern Culture	
WP	term	Work Package	
WTMC	organisation	Netherlands Graduate Research School of Science, Technology and Modern Culture	
Xpath	standard	XML Path Language	
Xquery	standard	W3C XML Query	
xTAS	software	Text Analysis in a Timely Manner	