CLARIN-*NL* Yearly Report 2011

1 Introduction

This document provides a description of the activities carried out in CLARIN-NL for 2011.

2 Planned Activities

EU-Line

It was decided to await the developments around the CLARIN ERIC (European Research Infrastructure Consortium) before undertaking any activities for the EU-Line. In the course of 2011, there were many developments, and decisions have been taken for the activities and budget for the EU-Line:

- The CLARIN ERIC will come into existence in February 2012

- A large part of the budget for the EU-Line will be reserved for the membership fees the Netherlands is obliged to the CLARIN ERIC as its host, to cover membership for a period of five years. This was explicitly agreed upon between the CLARIN-NL board, het Ministerie van Onderwijs, Cultuur en Wetenschap (OCW – the Ministry of Education, Culture and Science), and het Nederlands Instituut voor Wetenschappelijk Onderzoek (NWO – the Dutch Organization for Scientific Research).

- A small amount of the budget was used for smoothing the transition from the European CLARIN-prep project (which ended in June 2011) to the start of the ERIC (Feb 2012). This involves mainly salaries for the CLARIN-prep coordinator (Steven Krauwer) and some budget for accompanying activities (e.g. supporting SDH 2011 Copenhagen)

NL-Line

Technical

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 centers that expressed the ambition and made the commitment to become a CLARIN type A/B centre in the Netherlands, i.e. Max Planck Institute for Psycholinguistics (MPI (Nijmegen)), Meertens Instituut (Amsterdam), Instituut voor Nederlandse Lexicologie (INL – Institute for Dutch Lexicology (Leiden)) and Data Archiving and Networked Services (DANS (The Hague)).

In the course of 2010, Huygens Institute for Dutch History (Huygens ING) has also expressed the ambition to become a CLARIN Centre, and they have been invited to submit a plan to do so. The plan

has been submitted, has been evaluated by the executive board (EB), and returned to Huygens ING for adaptation.

By the end of 2011, however, it became clear that the candidate centers, esp., DANS and Huygens ING, want to reconsider their role as a CLARIN centre. This issue will be discussed in early 2012 in the CLARIN Centre Meeting.

Search & Develop (S&D)

The S&D project aims to create a network of centers in the Netherlands, focusing around search services. The S&D project 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.[1] Researchers have been requested to submit user scenarios, but the few that have been submitted[2] are too complex for the initial prototypes and will be reconsidered in a later phase of the project. The S&D project is running according to plan, and an intermediate progress report is available.[3]

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.

CLARIN Data Providers

The Netherlands has a range of centers 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 – Royal Library), *Nederlands Instituut voor Beeld & Geluid* (NIBG – Dutch Institute for Sound and Vision), *Nationaal Archief* (NA – the National Archive), *Digitale Bibliotheek van de Nederlandse Literatuur* (DBNL – Digital Library of Dutch Literature), the various university libraries, and also organizations such as the Department of information Services of the *Tweede Kamer.* 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*. Several of the organizations mentioned have been invited to make a proposal to become such a *CLARIN Data Provider.* KB and NIBG have been invited and they have submitted proposals which have been positively evaluated. DBNL and NA have been invited and they have submitted proposals that require still adaptation after the evaluation by the EB. The University Library of Utrecht has been invited, but no proposal has been received yet.

---

[1] http://lux17.mpi.nl/ds/fedsearch/

[2] http://www.clarin.nl/node/218

[3] http://www.clarin.nl/node/110 (access required)

Data Deposition Service

In the work plan for 2011, it was planned to set up a *Data Deposition Service* (DDS). A plan for this DDS was made, and submitted to the Board, but it was rejected by the Board. The EB aims to submit a revised plan, in part based on input from DANS, in 2012.

CMDI Metadata scheme for tools

It was planned to set up a small project to create a CMDI *metadata scheme for tools*. Though the plan for this has been elaborated, actual activities for making the scheme have not yet started in 2011. This activity will be shifted to 2012.

Data, Tools and Services

Calls Some projects from Call 1 (from 2009) have not been formally closed yet. Most projects from Call 2 (from 2010) are still running. In 2011 Call 3 has been organized, which consisted of an Open Call and a Closed Call. The Open Call was open to everybody who is eligible for funding, but for the closed call only researchers who were explicitly invited to submit a proposal could do so. The researchers invited for the Closed Call were selected by the EB from specific disciplines which were also selected by the EB. The selection of disciplines was done after an analysis of the discipline coverage, and was intended to ensure participation of disciplines that were not represented or underrepresented in CLARIN-NL. The researchers were selected on the basis of quality and availability with them of relevant data and / or tools.

All three calls are 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 IMDI or (if available 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 though revisions are required for two proposals, it is expected that all four projects will start early 2012.

No proposal was submitted for the discipline descriptive linguistics/language documentation, but the researcher in question has been requested to still submit a proposal in early 2012.

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).

A detailed report on the procedure followed for the third open call, together with reports on the evaluation by the IAP and NAP, has been made and was approved by the Board on December 12th, 2011.

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, the budget, the centre the project works with, a very short description of the scientific impact of the project, and its current status. More detailed descriptions of the projects can be found on the CLARIN-NL website.[4]

Call 1 Projects
Call 1 project information on website

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

---

[4] http://www.clarin.nl/node/281

	field researchers to annotate audio- and video-recordings
Budget	€ 56K
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, has submitted its final technical report (which it is revising) and still has to submit the financial report.

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 is finalizing and is about to submit the final reports. 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.

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 Instituut 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. Comments

have been sent on the final reports and the status of the software documentation. [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 is delayed and is targeted to finish by Feb 1, 2012. [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 *Wurdboek fan de Fryske Taal (WFT)* database and integrate the data in the *Geïntegreerde Taalbank (GTB)* 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 but esp. the status of the data categories in ISOcat has to be clarified.

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 (*Circulation of Knowledge and Learned Practices in the 17th-century Dutch Republic*, 'Geleerdenbrievenproject') is an independently financed (NWO) project that is carried out by a consortium of Dutch institutes (Descartes Centre, Historical Literature Amsterdam, Huygens, KB (National Library), DANS and Virtual Knowledge Studio (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 17th 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 [Language Resources and Technology] and would show the potential of a research infrastructure in the humanities" (CLARIN Newsletter 6, p. 3).[5]
Status	The project is running according to schedule

Call 2 Projects
Call 2 project information on website

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	It will enable access to the Arthurian Fiction database to a wide range of

---

[5] http://www.clarin.eu/files/cnl06_web_0.pdf

impact	humanities researchers such as literary historians in all European languages, cultural historians, medieval book specialists and art historians.
Status	The project is running and has requested a project extension until March 1, 2012, which has been approved.

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 Instituut
Scientific impact	This project will create new opportunities for (international) collaboration with other projects with similar source materials.
Status	The project has received an extension until Nov 30, 2011. The project has submitted their final reports and the EB is evaluating them.

COAVA	Cognition, Acquisition and Variation Tool
Project Coordinator	Dr. L. Cornips (Meertens Instituut)
Goal	COAVA is a mixed demonstrator and data curation project 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 Instituut
Scientific impact	It will enable interdisciplinary research into the relation between language acquisition and language variation.
Status	The project is running according to schedule.

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 analyzed 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 Instituut
Scientific	It will fill the gap in availability of syntactically analyzed corpus material for

impact	historical Dutch, enabling historical syntactic, morpho-syntactic and lexical research as well as benefiting non-linguistic historical research.
Status	The project is running according to schedule, but a request for extension may be expected (Currently targeted end date is Feb 29, 2012).

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 is running according to schedule.

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 is running and has requested an extension until September 1, 2012, which has been approved.

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 MW) 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 is running and has, due to pregnancy leave of one of the researchers, requested an extension until April 1, 2012, which has been approved.

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 is running and has requested an extension until June 1, 2012, which has been approved.

WIP	War in Parliament
Project Coordinator	Dr. H. Piersma (NIOD Instituut voor oorlogs-, holocaust- en genocidestudies (NIOD) – NIOD Institute for war, holocaust en genocide studies)
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 is running and has requested an extension until July 31, 2012, which has been approved.

Call 3 Projects (Closed Call)
Call 3 project information on website

FESLI	Functional Elements in Specific Language Impairment

Project Coordinator	Prof.dr. F. Weerman (University of Amsterdam)
Goal	FESLI 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 targeted to start early 2012.

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 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 targeted to start early 2012.

GrNe	Grieks-Nederlands woordenboek (Greek-Dutch dictionary)
Project Coordinator	Prof.dr. I. Sluiter (Leiden University)
Goal	GrNe is a mixed demonstrator and resource curation project in the Closed Call that aims to 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 targeted to start early 2012.

DiscAn	**Towards a Discourse Annotation system for Dutch language corpora**
Project Coordinator	Prof.dr. T.J.M. Sanders (Utrecht University)
Goal	DiscAn is a resource curation project in the Closed Call that aims to 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 sub disciplines.
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 formats and 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 targeted to start early 2012.

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 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 targeted to start early 2012.

MultiCon	**Multilayer Concordance Functions in ELAN and ANNEX**
Project Coordinator	Dr. O.A. Crasborn (Radboud University Nijmegen)
Goal	MultiCon 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	Status The project is targeted to start early 2012.

D-LUCEA	Database of the Longitudinal Utrecht Collection of English Accents
Project Coordinator	Dr. H. Quené (Utrecht University)
Goal	D-LUCEA 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 targeted to start early 2012.

EMIT-X	Early-Modern Image and Text eXchange
Project Coordinator	Dr. E. Stronks (Utrecht University)
Goal	EMIT-X 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 targeted to start early 2012.

Namescape	Mapping the Landscape of Names in Modern Dutch Literature
Project Coordinator	Prof.dr. K. van Dalen-Oskam (Huygens ING)
Goal	Namescape 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 targeted to start early 2012.

VK	Verrijkt Koninkrijk (Enriched Kingdom)
Project Coordinator	Dr. K. Ribbens (NIOD)
Goal	VK 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 targeted to start early 2012.

Cornetto-LMF-RDF	Curated Cornetto database in LMF and RDF
Project Coordinator	Prof.dr. P. Vossen (Free University Amsterdam)
Goal	Cornetto-LMF-RDF 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 targeted to start early 2012.

MIGMAP	Interactive migration maps for the 20th century
Project Coordinator	Dr. G. Bloothooft (Utrecht University and Meertens Instituut)
Goal	MIGMAP 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 Instituut
Scientific impact	The project will facilitate research into migration and linguistics, esp. language change.
Status	The project is targeted to start early 2012.

PoliMedia	Interlinking multimedia for the analysis of media coverage of political debates
Project Coordinator	Prof.dr. H. Beunders (Erasmus University Rotterdam)
Goal	PoliMedia 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 targeted to start early 2012.

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-centers 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 2) in which the basis for the technical infrastructure will be designed and implemented (with a budget of € 504K). On the other hand € 400K will be invested for the further realization of this part (in the TTNWW project). The formal decision on the exact usage of this amount has been made by the CLARIN-NL board in early 2010. Flanders contributes in this part € 452K in accordance with the usual NL: 2/3-VL: 1/3 proportion (all in the TTNWW project).

- **Part II (€ 920K; NL: € 680K; VL: € 340K)**
  - The contribution from the Netherlands consists of € 680K 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.

The project has started in March 2010, will run for 2.5 years, and is running according to schedule.

ISOCAT Coordinator

It was observed that work on ISOcat was carried out in an isolated manner in the various subprojects, that assistance was needed, and that for many problems a uniform approach must be selected for all subprojects. For this reason, an ISOcat Coordinator has been appointed since MARCH 1, 2011. The function has been filled by Ineke Schuurman. A separate report on her activities in 2011 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 required. Therefore, Maarten van Gompel is now also active in a project financed by CLARIN-NL to maintain and further extend CLAM, and to provide support to its users. (1/5 appointment, initially for one year).

Centers of Expertise

The plans for Centers of Expertise have been filled with a single virtual Centre of Expertise on the curation of data, led by the Radboud University Nijmegen but with close involvement of experts from the CLARIN Centers. A plan for this Data Curation Service (DCS) was developed by Nelleke Oostdijk (Radboud University Nijmegen), and it was approved by the Board in June 2011. The DCS has started preparatory activities in the fall of 2011, but will start with actual data curation as of 2012. The initial data set to be curated is the Dutch Bilingual Database (DBD). While curating these data, the DCS will investigate, plan and budget for additional candidates to be curated. The plan and budget for each selected resource will be submitted to the EB for approval before the data curation activities start.

Education and Awareness

An overview is given of the various activities for stimulating education, training, and awareness carried out in 2011.

Conferences and workshops visited

From	To	Description	City	Ctry	Who
17/01/2011	18/01/2011	CMDI workshop	Nijmegen	NL	DB
21/01/2011	21/01/2011	Oratie Paul Wouters	Leiden	NL	JO
11/02/2011	11/02/2011	CLIN	Gent	BE	JO
20/02/2011	20/02/2011	DARIAH workshop	London	GB	DB
08/03/2011	08/03/2011	Virtual Organization Workshop	Mannheim	DE	DB
28/03/2011	01/01/2011	eHumanities Symposium	Wassenaar	NL	JO
30/03/2011	30/03/2011	Diachroon Corpus	Den Haag	NL	JO
01/04/2011	01/04/2011	TTNWW workshop	Tilburg	NL	DB
07/04/2011	07/04/2011	SURFnet event about data management	Utrecht	NL	DB
12/04/2011	13/04/2011	EPIC User Forum	Amsterdam	NL	DB
09/05/2011	09/05/2011	EU workshop Federated Search	Nijmegen	NL	DB

18/05/2011	18/05/2011	SURFshare Open OnderzoeksData Dag	Amsterdam	NL	JO
26/05/2011	27/05/2011	FlareNet workshop	Venice	IT	DB
08/06/2011	08/06/2011	Dealing with data	Den Haag	NL	JO
09/06/2011	10/06/2011	Federated Identity Management Workshop	Genève	CH	DB
02/09/2011	03/09/2011	Peter Wittenburg International workshop	Kleve	DE	JO, DB
07/09/2011	09/09/2011	Panel Digging into Data	London	GB	HB
12/09/2011	13/09/2011	IRISC workshop	Helsinki	FI	DB
29/09/2011	29/09/2011	Alfalab Symposium	Amsterdam	NL	JO
10/10/2011	10/10/2011	Creatieve industrie en wetenschappelijk onderzoek	Amsterdam	NL	JO
19/10/2011	19/10/2011	Workshop 'Verrijkte publicaties'	Amsterdam	NL	AvH
01/11/2011	01/11/2011	CATCH meeting at HITIME	Amsterdam	NL	JO
05/11/2011	19/11/2011	IJCNLP Chiang Mai	Chiang Mai	TH	JO
17/11/2011	18/11/2011	SDH 2011	Copenhagen	DK	AvH
09/12/2011	09/12/2011	SARA 40	Amsterdam	NL	JO

Events organized by CLARIN-NL

From	To	Description	City	Ctry	Who
09/02/2011	09/02/2011	CLARIN-NL Call 2 Kick-off	Utrecht	NL	EB
10/05/2011	10/05/2011	ISOcat Workshop	Utrecht	NL	JO, AvH
25/08/2011	25/08/2011	CLARIN-NL Call 3 Info Session	Utrecht	NL	EB
24/10/2011	24/10/2011	ISOcat Workshop, part 2	Utrecht	NL	

Publications

Cornips, L., Kemps-Snijders, M., Snijders, M., Swanenberg, J. and Vriend, F. de (2011), "Bridging the Gap between First Language Acquisition and Historical Dialectology with the Help of Digital Humanities". In: Proceedings of the 2nd Conference on Supporting Digital Humanities (SDH 2011).

Nerbonne, J., Colen, R., Gooskens, C., Kleiweg, P and Leinonen, T. (2011), "Gabmap — A Web Application for Dialectology". In: Nerbonne, J., Grondelaers, S., Speelman, D. and Perea, M. (eds), Dialectologia, Special Issue II (Spec. Iss. Production, Perception and Attitude), pp. 65-89.

Odijk, J. (2011), "CLARIN Developments in the Countries: the Netherlands", CLARIN NewsFlash March 2011, p. 2.

Odijk, J. (2011). "CLARIN-NL Story: How to successfully convince your government", CLARIN Newsletter 13 (June 2011), p. 5.

Odijk, J. and Hessen, A. van (2011), "Sharing Resources in CLARIN-NL", Proceedings of the Language Resources, Technology and Services in the Sharing Paradigm workshop at IJCNLP 2011, pp. 98-106 (November 2011), Chiang Mai, Thailand.

Presentations by the CLARIN-NL EB members

Bennis, H. (2011), "Populating the infrastructure: the case of the Netherlands", presentation held at the CLARIN Coordinators Meeting, Budapest, June 29 and 30, 2011.

Broeder, D. (2011), "CMDI - CLARIN Component Metadata Infrastructure", presentation held at the CMDI Tutorial, Nijmegen, January 17, 2011.

Broeder, D. (2011), "Een kijkje in de CLARIN keuken", presentation held at SURFnet event about data management, Utrecht, April 7, 2011.

Broeder, D. (2011), "Introduction, Federated Search as a EU CLARIN demonstrator", presentation held at the EU workshop Federated Search, Nijmegen, May 9, 2011.

Broeder, D. (2011), "NP 24622 CMDI-1 Metadata Component Framework - New Standardization Work within TC37/SC4", presentation held at the ISO TC37/SC4 meeting, Seoul, June 13, 2011.

Broeder, D. (2011), "Populating the infrastructure using standards", presentation held at the CLARIN Coordinators Meeting, Budapest, June 29 and 30, 2011.

Odijk, J. (2011), "Introduction to Language and Speech Technology", presentations held at the LOT Winterschool, Amsterdam, January, 2011 (Introduction, Grammar and chart, Parsing, Machine learning, Automatic Speech Recognition, Machine Translation).

Odijk, J. (2011), "Call 2 Kick-off - Introduction", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Odijk, J. (2011), "Linguistic Research in the CLARIN Infrastructure", presentation for the KNAW eHumanities Workshop, NIAS, Wassenaar, March 29, 2011.

Odijk, J. (2011), "CLARIN-NL Call 3" and "Requirements and desiderata", presentations held at the Call 3 information session, Utrecht, August 25, 2011 .

Odijk, J. and Hessen, A. van (2011), "Sharing Resources in CLARIN-NL", Presentation held at the Language Resources, Technology and Services in the Sharing Paradigm workshop at IJCNLP 2011, Chiang Mai, Thailand, November 12, 2011.

Presentations by others

Barbiers, S. (2011), "MIMORE", presentation held at the Call 3 information session, Utrecht, August 25, 2011.

Berg, M. van den and Sluijs, R. van (2011), "NEHOL - NEgerHOLlands Database", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Bruin, M. de (2011), "C-DSD - Curating the Dutch Song Database", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Caspel, P. van (2011), "CLARIN-NL HELPDESK", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Cornips, L., Kemps-Snijders, M., Snijders, M., Swanenberg, J. and Vriend, F. de (2011), "Bridging the gap between first language acquisition and historical linguistics with the help of digital humanities", presentation held at Supporting Digital Humanities, University of Copenhagen, November 17, 2011.

Crasborn, O. (2011), "Support for sign language resources in the Netherlands and beyond in CLARINNL", presentation of SignLinC and IPROSLA (Call 2) held at the CLARIN-D kick-off meeting, Leipzig, July 22, 2011.

Dalen-Oskam, K. van and Besamusca, B. (2011), "Arthurian Fiction in Medieval Europe: Narratives and Manuscripts", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Dalen-Oskam, K. van (2011), "ArthurianFiction", presentation held at the Call 3 information session, Utrecht, August 25, 2011.

Duin, P. (2011), "CMDI Component Registry", presentation held at the CMDI Tutorial, Nijmegen, January 17, 2011.

Duin, P. (2011), "CMDI Virtual Language Observatory Faceted Browsing", presentation held at the CMDI Tutorial, Nijmegen, January 17, 2011.

Ďurčo, M. (2011), "CMDRSB", presentation held at the CMDI Tutorial, Nijmegen, January 17, 2011.

Fikkert, P. (2011), "Integrating and publishing resources on sign language acquisition (IPROSLA), presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Gerritsen, G.H. (2011), "Geleerdenbrievencorpus en analysetools", presentation held at the workshop 'Diachroon Corpus', Den Haag, March 30, 2011.

Halteren, H. van (2011), "Processing 14th century Dutch text through Clarin, step 1: Adelheid", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Halteren, H. van (2011), "Adelheid: Tagging and Lemmatizing Historical Dutch Texts through the Clarin Infrastructure", presentation held at Galatea II, Antwerp, February 18, 2011.

Halteren, H. van (2011), "Dealing with Medieval Text", presentation held for the course "Language Engineering Applications" at KU Leuven, Leuven, February 21, 2011.

Kemenade, A. van and Postma, G. (2011), "International Parser and Lemmatizer of Dutch in Retrospect (INPOLDER)", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Kemps-Snijders, M. (2011), "The Dutch Search and Development project and the Meertens search engine", presentation held at the EU workshop Federated Search, Nijmegen, May 9, 2011.

Kemps-Snijders, M. (2011), "Search and Develop", presentation held at the Call 3 information session, Utrecht, August 25, 2011.

Kleiweg, P. and Leinonen, T. (2011), "Gabmap - Doing dialect analysis on the web", poster at ICLAVE 6, Freiburg, June 30, 2011.

Leinonen, T. (2011), "Dialektgeografiska kartor i Gabmap", tutorial at the workshop vid Svenska litteratursällskapet i Finland r.f., Helsinki, November 23 and 25, 2011.

Nerbonne, J. (2011), 'Gabmap: A Web Application for Analyzing Linguistic Variation', Tutorial given at Digital Humanities 2011, Stanford, California, USA, May 19, 2011.

Nerbonne, J., Wieling, M. and Heeringa, W. (2011), "Gabmap - A web application for measuring and visualizing distances between language varieties", workshop at the international conference Methods in Dialectology XIV, London, Ontario, August 2-6, 2011.

Nerbonne, J. (2011), "Gabmap, een webapplicatie voor dialectologie", presentation held at the Call 3 information session, Utrecht, August 25, 2011.

Nerbonne, J. (2011), "Gabmap", tutorial in connection with the "Workshop on comparing approaches to measuring linguistic differences", Gothenburg, Sweden, October 24-26, 2011.

Piersma, H. (2011), "War in parliament", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Ravenek, W. and Van den Heuvel, C.M.J.M. (2011), "Circulation of knowledge: Integrated access to texts and metadata", presentation held at the International Workshop Representing the Republic of Letters, Huygens Institute for the History of the Netherlands (in cooperation with Descartes Centre UU), Den Haag, June 30 – July 1, 2011.

Ravenek, W. (2011), "Supporting the Exploration of a Corpus of 17th-Century Scholarly Correspondences by Topic Modeling", presentation held at Supporting Digital Humanities 2011, Copenhagen, November 18, 2011.

Schuurman, I. and Windhouwer, M. (2011), "ISOcat introduction", presentation held at the ISOcat Workshop, Utrecht, May 10, 2011.

Schuurman, I. and Windhouwer, M. (2011), "Content of the Data Category Registry", presentation held at the ISOcat Workshop, Utrecht, May 10, 2011.

Schuurman, I. and Windhouwer, M. (2011), "Data Category specifications", presentation held at the ISOcat Workshop, Utrecht, May 10, 2011.

Schuurman, I. and Windhouwer, M. (2011), "RELcat and friends", presentation held at the ISOcat Workshop, Utrecht, May 10, 2011.

Schuurman, I. and Windhouwer, M. (2011), "ISOcat: known issues", presentation held at the ISOcat Workshop, Utrecht, May 10, 2011.

Schuurman, I. (2011), "CLARIN-NL: Dealing with ISOcat", presentation held at the CLARIN Coordinators Meeting, Budapest, June 29 and 30, 2011.

Schuurman, I. and Windhouwer, M. (2011), "CLARIN-NL ISOcat workshop 2011, part 2", presentation held at the ISOcat workshop 2011, part 2, Utrecht, October 24, 2011.

Schuurman, I. and Windhouwer, M. (2011, "Bulk loading ISOcat data categories with the Data Category Interchange Format", presentation held at the ISOcat workshop 2011, part 2, Utrecht, October 24, 2011.

Snelders, S. and Pieters, T. (2011), "WAHSP: Towards a flexible and stable CLARIN-supported web-application for historical sentiment mining in public media", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Snelders, S. and Pieters, T. (2011), "WAHSP: Towards a flexible and stable CLARIN-supported open-source web-application for historical sentiment mining in public media", presentation held at the Call 3 information session, Utrecht, August 25, 2011.

Stehouwer, H. (2011), "Implementing SRU/CQL on Trova", presentation held at the EU workshop Federated Search, Nijmegen, May 9, 2011.

Strik, H. (2011), "TQE: Transcription Quality Evaluation", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Trippel, T. (2011), "CMDI use in the NaLiDa project", presentation held at the CMDI Tutorial, Nijmegen, January 17, 2011.

Van den Heuvel, C.M.J.M. (2011), "Themes and Debates Emerging from 17th Century Correspondences of Dutch Scholars' Circulation of Knowledge: Visualizations and Interfaces for Research," presentation held at the Mapping the Republic of Letters Conference, Fondazione Cini, San Giorgio Maggiore, Venice, March 16-18, 2011.

Van den Heuvel, C.M.J.M. and Coleman, C.N. (2011), "Visualizing Uncertainty and Complexity: Humanistic Methods for Mapping the Intellectual Geography of the Early Modern World", presentation held at the conference 'Intellectual Geography: Comparative Studies, 1550-1700', St Anne's College, University of Oxford, September 5-7, 2011.

Van den Heuvel, C.M.J.M. (2011), "Visualizing Universes of Knowledge" invited lecture held at the Winsor Lecture- Graduate School for Library and Information Sciences, University of Illinois, Urbana-Champaign, October 5, 2011.

Van den Heuvel, C.M.J.M. (2011), "Circulation of Knowledge and Learned Practices in the 17th century Dutch Republic. A Web-based Collaboratory on Correspondences. Expectations-Experiences-Recommendations", invited lecture held at the CATCH-Meeting HiTIME: e-History colloquium, International Institute of Social History, Amsterdam, November 1, 2011.

Vriend, F. de (2011), "COAVA", presentation held at the CLARIN-NL Kick-off meeting Call 2, Utrecht, February 9, 2011.

Windhouwer, M. (2011), "The ISO-DCR", presentation held at the CMDI Tutorial, Nijmegen, January 17, 2011.

Withers, P. (2011), "ARBIL, a tool for organising corpus/language data", presentation held at the CMDI Tutorial, Nijmegen, January 17, 2011.

Wittek, P. and Ravenek, W. (2011), "Supporting the Exploration of a Corpus of 17th-Century Scholarly Correspondences by Topic Modeling". In: Maegaard, B. (Ed.), Supporting Digital Humanities 2011: Answering the unaskable. Copenhagen, Denmark.

Meetings

From	To	Description	City	Ctry	Who	With Whom
19/01/2011	19/01/2011	Tweede Kamer	Den Haag	NL	JO, AvH	GJL
19/01/2011	19/01/2011	Meeting CLARIN European Demonstrator Case	Nijmegen	NL	DB	
24/01/2011	24/01/2011	Lecture OH and CLARIN	Enschede	NL	AvH	
26/01/2011	26/01/2011	Prep ERIC SC	Den Haag	NL	JO	RD
15/03/2011	15/03/2011	ED&R Bestuur	Den Haag	NL	JO	
17/03/2011	18/03/2011	CLARIN ERIC SC	Den Haag	NL	JO	
24/03/2011	24/03/2011	CLARIN-NL interview	Utrecht	NL	JO, AvH	
29/03/2011	29/03/2011	eHumanities meeting	Wassenaar	NL	HB	
06/04/2011	06/04/2011	TAB Taalportaal	Leiden	NL	JO	
15/04/2011	15/04/2011	ERIC SC 2	Den Haag	NL	JO	
27/04/2011	27/04/2011	ED&R Bestuur	Den Haag	NL	JO	
28/04/2011	28/04/2011	CLARIN ERIC meeting with NTU (a.o.)	Roosendaal	NL	HB	
04/05/2011	04/05/2011	EC meeting	Brussels	BE	DB	
13/06/2011	15/06/2011	TC37/SC4 meeting	Seoul	KR	DB	
16/06/2011	16/06/2011	KNAW Meeting	Amsterdam	NL	JO, JN	
21/06/2011	21/06/2011	CLARIN-NL / OCW-NWO	Den Haag	NL	JO, ME, HB	RD, AD
22/06/2011	22/06/2011	CLARIAH Meeting	Amsterdam	NL	JO	many
28/06/2011	30/06/2011	Meeting of CLARIN National coordinators	Budapest	HU	HB, JO, DB	many
11/08/2011	11/08/2011	CLARIAH meeting	Amsterdam	NL	JO	many
23/08/2011	24/08/2011	CMDI dev meeting	Nijmegen	NL	DB	
26/08/2011	26/08/2011	CLARIAH meeting	Amsterdam	NL	JO, HB, DB	many
12/09/2011	12/09/2011	IPROSLA meeting	Nijmegen	NL	HB	
13/09/2011	13/09/2011	Lecture for VOICE	Utrecht	NL	AvH	
15/09/2011	15/09/2011	Meeting at KNAW	Amsterdam	NL	HB	DW

16/09/2011	16/09/2011	'Ronde tafelbijeenkomst' Hot technology for information professionals	Rotterdam	NL	AvH	
21/09/2011	21/09/2011	Prep meeting Creatieve Industrie	Nijmegen	NL	JO	LvdB, AvdB, CC
28/09/2011	28/09/2011	Meeting Nationaal Archief	Den Haag	NL	JO	MB
29/09/2011	29/09/2011	Final presentation Alfalab	Amsterdam	NL	HB	
30/09/2011	30/09/2011	OH lecture	Amsterdam	NL	AvH	
01/10/2011	31/10/2011	Many meetings with several applicants about Nederlab		NL	HB	
10/10/2011	10/10/2011	Meeting about Creatieve Industrie	Amsterdam	NL	HB	TM, DW
14/10/2011	14/10/2011	Meeting UBU	Utrecht	NL	JO	DJ
14/10/2011	14/10/2011	Valerie Frissen Creatieve Industrie	Den Haag	NL	JO, AvdB, LvdB	VF
17/10/2011	17/10/2011	Meeting about digitization and CLARIN	Den Haag	NL	HB	BS, PD
31/10/2011	31/10/2011	Meeting Raad van Toezicht CATCHPlus	Amsterdam	NL	HB	
31/10/2011	31/10/2011	CLARIAH meeting	Utrecht	NL	DB	
03/11/2011	03/11/2011	Expert meeting Erasmus (CROME)	Rotterdam	NL	AvH	
09/11/2011	09/11/2011	Lecture Hogeschool Utrecht	Utrecht	NL	AvH	
10/11/2011	10/11/2011	INPOLDER meeting	Amsterdam	NL	HB	
11/11/2011	11/11/2011	Meeting about Creatieve Industrie	Amsterdam	NL	HB	AB, JOO
17/11/2011	17/11/2011	Meeting about The Language Archive	Amsterdam	NL	HB	TM
20/11/2011	20/11/2011	CLARIN- NL/ DE Meeting	Utrecht	NL	EB	Many
23/11/2011	23/11/2011	ArthurianFiction meeting	Den Haag	NL	HB	

28/11/2011	28/11/2011	Meeting about Centre for Digital Humanities	Amsterdam	NL	HB	VU, UvA, KNAW
28/11/2011	28/11/2011	Closing session STEVIN	Rotterdam	NL	AvH	
29/11/2011	20/11/2011	Taal in Bedrijf	Rotterdam	NL	AvH	
30/11/2011	30/11/2011	CLARIAH Meeting	Utrecht	NL	JO, DB	
02/12/2011	02/12/2011	Talk with Rijksdienst voor Cultureel Erfgoed	Amsterdam	NL	HB	DH
09/12/2011	09/12/2011	CLARIAH Meeting	Den Haag	NL	JO	
12/12/2011	12/12/2011	Reviews for Digital Humanities 2012	Amsterdam	NL	HB	
13/12/2011	13/12/2011	Meeting Stuurgroep CATCHPlus	Den Haag	NL	HB	
15/12/2011	15/12/2011	Meeting about The Language Archive	Amsterdam	NL	HB	DB, PW, SW, HW
19/12/2011	19/12/2011	Meeting about Early Dutch Books on Line	Amsterdam	NL	HB	JF
20/12/2011	20/12/2011	Presentation at Dutch embassy Sarajevo	Sarajevo	BA	AvH	
21/12/2011	21/12/2011	Meeting about Dienstenmodel CATCHPlus	Den Haag	NL	HB	MA, MS

(JN= John Nerbonne; AvdB=Antal van den Bosch; LvdB = Linde van den Bosch; AD= Alice Dijkstra; RD= Richard Derksen; MB=Martine Berendsen; ME=Martin Everaert; GJL=Geert-Jan Lodder; CC=Catia Cucchiarini; DJ=Dafne Jansen; DW=Demitrius Waarsenburg; TM=Theo Mulder; BS=Bas Savenije; PD=Paul Doorenbosch; AB=A. Bos; JOO=Johan Oomen; DH=Dirk Houtgraaf; SW=Sally Wyatt; HW=Henk Wals; JF=Jasper Faase; MA=Martijn Arnoldus; MS=Marco Streefkerk)

Support for other events

Date	Description	City	Country	Form of Support
17/01/2011	CMDI developers meeting	Nijmegen	NL	Subsidy of €228,50
09/05/2011	Federated Search & EDC Workshop	Nijmegen	NL	Subsidy of €3000,-
19/05/2011	Kick-off meeting	Amsterdam	NL	Subsidy of

				€116,15
	VU-DNC			
23/06/2011	Workshop E-History	Amsterdam	NL	Subsidy of €423,85

Support for travel of Dutch researchers to CLARIN-related events

Date	Description	City	Country	Who	Organization
12/06/2011	ISO TC 37 plenary	Seoul	KR	Menzo Windhouwer	MPI
19/06/2011	Digital Humanities	Stanford	USA	John Nerbonne	RUG
20/07/2011	Methods in Dialectology XIV	Ontario	Canada	John Nerbonne	RUG
16/11/2011	Supporting Digital Humanities	Copenhagen	DK	Ineke Schuurman	UU

CLARIN- NL Newsflashes

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

- Newsflash 18 (23-12-2011)

- Newsflash 17 (05-09-2011)

- Newsflash 16 (01-07-2011)

- Newsflash 15 (21-04-2011)

- Newsflash 14 (26-01-2011)

- (Newsflash 13 (21-11-2010))

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

CLARIN-EU Newsflashes

In 2011, two CLARIN newsflashes have appeared.

A contribution from CLARIN-NL has been included in the following CLARIN EU Newsflash:[6]

Jan Odijk (2011), 'CLARIN Developments in the Countries: the Netherlands', CLARIN *NewsFlash* March 2011, p. 2. [pdf]

CLARIN-EU Newsletter

In 2011 two CLARIN-EU Newsletters[7] have appeared. From CLARIN-NL there was one contribution:

---

[6] http://www.clarin.eu/view_newsflash

[7] http://www.clarin.eu/external/index.php?page=publications

Jan Odijk (2011). 'CLARIN-NL Story: How to successfully convince your government', *CLARIN Newsletter* 13 (June 2011), p. 5. [pdf]

eData & Research

CLARIN-NL has joined the eData & Research Magazine. 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, and CLARIN-NL aims to continue its participation in 2012. On behalf of CLARIN-NL, Erica Renckens participates in the editorial team and Jan Odijk in the board.

The following articles have been published in eData & Research in 2011. They can be accessed via http://www.clarin.nl/node/188.

- "CLARIN-NL deelnemer in edata & research", article by Erica Renckens in eData & Research 5(4), March 2011, p. 1.

- "Negen nieuwe CLARIN projecten van start", article by Erica Renckens in eData & Research 5(4), March 2011, p. 7.

- "CLARIN-NL - Common Language Resources and Technology Infrastructure in Nederland", focus article by Inge Angevaare in eData & Research, 6(1), June 2011, p.6.

- "Zoekmachine vindt relevante WO2-verwijzingen in Handelingen der Staten Generaal. Dat doet denken aan de oorlog", article by Erica Renckens in eData & Research 6(2), October 2011, p. 1.

CLARIN-NL in the press

- "Waarom politici graag over de oorlog praten", article by Bart Funnekotter in NRC Handelsblad about the subproject WIP, February 25, 2011.

- "Oorlog in de Kamer", article by Maarten Marx at 'PoliticalMashup', March 1, 2011.

- "Het Digitale Drama", article by Karel Berkhout in NRC Handelsblad, September 10, 2011 and various responses, in particular the response by Jan Odijk (in Dutch).

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.

The EB has also commissioned the production of 6 short movies on CLARIN-NL projects (to Dyzlo). Below one can find the CLARIN-NL projects that are selected for these movies. The current status of these movies is as follows:

Project	Status	Target Delivery date
MIMORE	Finished	01/02/2012
SignLinc	Being finalized	01/03/2012
Arthurian Fiction	Recordings finished, editing on-going	01/03/2012
WFT-GTB	Recordings to start Feb 2012	15/03/2012
WIP	Recordings finished, editing on-going	15/03/2012

Search & Develop	Recordings to start end of February 2012	01/04/2012
Gabmap (movie by the ADEPT project)	Finished	

The movies will be made available in streaming mode via a VIMEO channel, and the originals will be made available on hard disk or DVD/BluRay disk.

Website

A website has been set up for CLARIN-NL. The design and maintenance of the website is not optimal. We have hired an external company (Joygroup) to do the maintenance and extend the website's functionality, which it has done for an initial set of tasks. However, for a second set of tasks, they turned out to be completely nonresponsive. We have therefore hired a new company (Clearwind), which will start on a redesign of the website early 2012.

The CLARIN-NL website will become a project-internal website. We also aim 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 design and editorial work for this to CAOS and aim to further involve Erica Renckens for the editorial work.

At the same time we intend to build, in 2012, a portal to the CLARIN infrastructure, which will integrate all the loose bits and pieces (components) that have been and are being developed in the various projects. This portal will give access to increasingly rich and sophisticated infrastructure functionality, where the introduction of each new component will be accompanied by PR-activities.

Education

JO held a LOT Winterschool course, for one day together with Arjan van Hessen, (*Introduction to Language and Speech Technology*, Amsterdam, 17-21 January 2011), in which the CLARIN infrastructure was a recurrent theme.

Helpdesk

The CLARIN-NL Helpdesk[8] has been set up and has run all year. Since the amount of requests is rather low, there was no justification to keep an employee specific for this task. Since Mid 2011, the Helpdesk is managed by Jan Odijk and Daan Broeder.

The EB has proposed to CLARIN-D to cooperate on the Helpdesk, and to combine FAQ sections. Discussion on this will be continued in 2012.

After a meeting with the KNAW, the KNAW declared themselves prepared to include "CLARIN-compatibility" for language-related resources in their data plan. To that end, the concept "CLARIN-compatible" had to be defined explicitly in a single document.

---

[8] http://trac.clarin.nl/trac

A document explaining the concept "CLARIN-compatible" has been written[9], and the KNAW researchers have been pointed out the existence of this document. The document is obviously also relevant to other researchers (and has been referred to in Call 3) but so far exists only in Dutch. This activity should be followed up by 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.

National Coordination Point

The national coordination point is unchanged in comparison to 2010. Its members are prof.dr. Jan Odijk (1FTE) as Programme Director and Drs. Jolien Scholten as project secretary (.5 FTE).

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.

Board The Board consists of 8 members. A full list of its members can be found on page 31 and on the CLARIN-NL website (http://www.clarin.nl/node/15).

The Board has held 3 meetings in 2011.

NAP The National Advisory Panel consists of 17 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 A and on the CLARIN-NL website (http://www.clarin.nl/node/16). The NAP elected Dr. Karina van Dalen-Oskam as NAP chair since Dr. Willemijn Heeren had to give up this position due to her extended stay abroad.

The NAP has held 3 meetings in 2011. In one of these meetings the focus was on evaluating the proposals submitted in the third call.

IAP The International Advisory Panel consists of 8 members, though there is one vacancy. A full list of the members can be found in Appendix B and on the CLARIN-NL website (http://www.clarin.nl/node/52). The EB has selected a candidate for the vacancy and invited the candidate, but the candidate did not respond to e-mail or telephone calls. The IAP has held one

---

[9] http://trac.clarin.nl/trac/attachment/wiki/WikiStart/CLARIN%20compatible%20NL%20110805.pdf

meeting in 2011, which was dedicated to the evaluation of the Call 3 project proposals. The task of the IAP is hard (many project proposals must be reviewed by few IAP members), and a few IAP members were not able to participate in the project proposal evaluation. Therefore the EB has proposed to the Board to invite three more IAP members, and the Board has approved this. They can be selected from the former CLARIN-prep IAP and from researchers active in national CLARIN activities in other countries (e.g. Norway, Germany, etc.).

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 3 new partners (Erasmus Universiteit Rotterdam (EUR), Intelligent Systems Lab Amsterdam (ISLA) and Instituut voor Gebaren, Taal & Dovenstudies (IGTD) have joined the consortium in 2011, bringing the total number of partners to 28 (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 2011 The Working Plan for 2011 was approved by the Board on August 19, 2009. It can be found on the CLARIN-NL website: www.clarin.nl/system/files/CLARIN%20NL%20Workplan%202011%20120323.pdf

Overall, the governance structure as set up functions appropriately but it has become clear that the effort of the Project Secretary is now, with the increasing number of subprojects, significantly larger than the amount of FTE assigned to this function. The EB therefore intends to increase the amount of FTE for the project secretary.

3   2011 Finances

A detailed report on the CLARIN-NL finances in 2011 can be found in the 2011 Financial Report.

4   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. This proposal has been submitted and it has received excellent and good-excellent reviews by external experts. The consortium has written a rebuttal to the comments of the external experts, and it has presented and defended the proposal in an interview with the Committee on January 9th, 2012. The results of this procedure are expected to be made available by the end of the spring of 2012.

5 Assessment of the Current Situation

A lot of work has been done in the context of CLARIN in the last two years, in building bricks and pieces of the CLARIN infrastructure but all of these activities are distributed and have been done in isolation. It is now time to start working towards an integration of all the work done in a real and visible CLARIN infrastructure, and also to organize the remaining work around this CLARIN infrastructure. Integrating all the functionality and data will be a test for the interoperability and will help us identify functionality that is still lacking and should be added. An incremental build-up of this infrastructure can be accompanied by PR-activities so that the CLARIN-infrastructure becomes visible to the user community and starts to 'live' with them. The demonstrators, showcases, movies, screen captures and other material developed so far can play a big role in these PR-activities. In 2012, a plan will have to be developed and its execution started to achieve this integration.

Composition of the Board

Name	Affiliation
Prof.dr. Geert Booij	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	ACLC, Universiteit van Amsterdam
Prof.dr.ir John Nerbonne	CLCG, Universiteit van Groningen

Appendix A Composition of the NAP

Name	Affiliation
Prof.dr. Willem Adelaar	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	UiL-OTS, Universiteit Utrecht
Prof.dr. Piek Vossen	Vrije Universiteit Amsterdam
Prof.dr. Sally Wyatt	eHumanities (KNAW)
Dr. Joris van Zundert	Huygens Instituut voor Nederlandse Geschiedenis

Appendix B 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