



CLARIN - NL

What is going on?

Jan Odijk

Amsterdam 26 Aug 2010

Overview



- The CLARIN-NL Project
- CLARIN Infrastructure
- Targeted Users
- Running Subprojects
- New Subprojects

CLARIN-NL Project



- Project in the Netherlands
- Aims to play a leading role in the creation of the European CLARIN technical infrastructure
- Budget: 9.01M Euro
- 2009-2014
- Coordinated by Utrecht University
- 23 participants
- <http://www.clarin.nl/>

CLARIN-NL: Infrastructure



- CLARIN-NL aims to design, construct, validate, and exploit
 - a research infrastructure that is needed to provide a sustainable and persistent eScience working environment
 - for researchers in the Humanities
 - who want to make use of language data and tools.

CLARIN-NL: Infrastructure



- The CLARIN Infrastructure
 - Will make data and tools on different locations easily accessible via web interfaces and services (CLARIN-portal(s) with intelligent searching, browsing, viewing and querying services)
 - Will make it possible for non-technical researchers to extract / combine/ enrich data (supported by dissemination and training)
 - Will make available **interoperable** data and tools based on existing standards and best practices

CLARIN-NL: For whom?



- For researchers that work with language data
 - Humanities
 - Linguistics (broadly construed)
 - Literary and Theatrical Studies
 - Media en Culture
 - History
 - Political Sciences
 - ...

CLARIN-NL: For whom?



- Current Partners (23)
 - Targeted Users
 - Linguistics (10)
 - Culture (2)
 - Lexicography (2)
 - Social History (4)
 - Literature (2)
 - Technology Providers
 - Language technology (6)
 - Speech technology (2)
 - Data Centres and Service providers
 - Data Centres (5)
 - Libraries (2)

CLARIN-NL: Subprojects



- Information on [subprojects](#):
- Infrastructure Implementation
 - General
 - Partners: Candidate CLARIN Centres
 - MPI, MI, INL, DANS
 - Directly assigned subprojects
 - Provide guidelines / training for others

CLARIN-NL: Subprojects



- Infrastructure Implementation (cont.)
 - [Infrastructure Implementation Project](#) (IIP, 3 yrs)
 - infrastructure services, an open archiving service, registries, federation of centres, set up a schema registry, profile matching, [ISOCAT](#) maintenance, add relation registry RELCAT.
 - coordinate and give guidance for work on web services, wrapper and service bus specification and implementation, select work flow tools and experiment with them.

CLARIN-NL: Subprojects



- Infrastructure Implementation (cont.)
 - Metadata project (MD, 0.5 yr)
 - Testing CMDI against existing national data
 - Create initial set of required metadata components
 - Results: (see <http://www.clarin.eu/cmdl>)
 - Component Registry
 - CMDI XML Toolkit
 - Documentation for users and developers
 - Metadata Tutorial held
 - Search&Develop (S&D, 3 yrs)
 - centralized metadata search
 - distributed content search
 - Text based and structured search

CLARIN-NL: Subprojects



- User Survey & Base Line (.5yr)
 - Directly assigned subproject
 - User survey
 - Interactive interviews
 - Current use/non-use of digital data and tools
 - Identify causes for non-use
 - Identify obstacles for (wider) use

CLARIN-NL: Subprojects



- Call 1 Data Curation & Demonstrator Projects
 - Involve a targeted user and address the user's research questions
 - Open call for small subprojects (.5 yr / 60k Euro)
 - 12 projects running, some already finished
 - Will make available
 - a range of curated resources
 - a range of showcases of CLARIN functionality
 - evidence-based requirements and desiderata
 - for the CLARIN infrastructure and
 - for supported standards and best practices

CLARIN-NL Subprojects



- [TTNWW](#) (3 yrs)
 - Cooperation between Netherlands and Flanders
 - Implement user friendly workflow services for
 - Speech
 - indexing and search of (a limited set of) audio and video data
 - For social historians (Aletta, KDC, KADOC, M2P)
 - Text
 - enriching text corpora with annotations
 - For literature researchers (Huygens) and archeologists (Salagassos)

CLARIN-NL Subprojects



- New in 2010
- Second Open Call for data curation & demonstrator projects (see later)
- Directly assigned projects
 - Selection to be based (inter alia) on results of user survey
 - Budget available: 400k Euro
 - Initial selection ASAP

CLARIN-NL: Subprojects



- Education, Training & Awareness
 - CLARIN-NL Helpdesk – see later
 - CLARIN-NL organizes events
 - CLARIN-NL support events (logistically and financially) <http://www.clarin.nl/node/123>
 - Support individual researchers for visiting events <http://www.clarin.nl/node/50>
 - Newsflashes. Subscribe at <http://www.clarin.nl>



Thanks for your attention!

<http://www.clarin.nl/>