



CLARIN-NL

Second Open Call

Jan Odijk

CLARIN-NL Call 2 Info-session

Amsterdam, 26 Aug 2010



Overview

- Background
- Project Types
- Project Goals
- Roles
- Resource Curation Projects
- Demonstrator Projects
- CLARIN Centres
- Requirements and Desiderata
- Practical Matters

Background

- CLARIN-NL Phase 1: Specification and Design
- Incorporate existing data and tools
- Standards and Interoperability
 - <http://www.clarin.eu/recommendations>
- Conversions and Adaptations
 - Standardized data and metadata
 - Web applications structured to become web services
 - Demonstrators
 - Requirements and Desiderata

Project Types

- Resource Curation Projects
 - Data or tools
 - Curation of digital language resources
- Demonstrator Projects
 - Web applications
 - User interface
 - Core component (possibly web-service)
 - Demonstrator
- Mix of the two

Project Goals (1)

- Resource Curation Projects
 - Adapt existing resources to make them visible, uniquely referable and accessible via the web, CLARIN-compatible and properly documented
- Demonstrator Projects
 - Create a Web application
 - Starting from an existing tool / application
 - Demonstration scenario



Project Goals (2)

- For all projects:
 - establish requirements and desiderata for the CLARIN infrastructure
 - By applying standards and best practices
 - By making use of CLARIN architecture and agreements
- So that we understand their limitations



Roles

- User
- Data Provider
- Technology Provider
- Infrastructure Specialist

User

- Is a researcher from
 - Linguistics, or
 - Humanities more broadly
- Wants to investigate research question(s)
- Some data and tools can facilitate this
- Must be the project leader

Data provider

- Has digital language data
- Can be used to address the user's research question
- Has the rights to use these data and to make them available on CLARIN-servers

Technology provider

- Has technology and expertise on this technology
 - E.g. language or speech technology
- Can be used to address the user's research question
 - Can form the basis for a web application and demonstrator, or
 - Can be used for resource curation
- Has the rights to use this technology and (for web applications) to make them available on CLARIN-servers

Infrastructure Specialist

- Has a deep understanding of
 - CLARIN service-oriented architecture and its requirements, and/or
 - Data, metadata or tool format standards and best practices supported in CLARIN
- Will assist a project and its members if needed
- Always via the CLARIN-NL HelpDesk

Roles and People

- Often roles will be played by different persons (possibly of different organisations), but
- Several roles may coincide in one person
 - E.g. user & data provider
 - Data provider & technology provider



Brokerage

- CLARIN-NL will assist in bringing persons with different roles together
 - Contact the CLARIN-NL office
 - This meeting!

Resource Curation Project (1)

- Resource must be brought in CLARIN-standards compatible format
- Metadata descriptions must be created and made available, including persistent identifiers, using CMDI
- Linguistic encoding must be related to the data category registry as implemented in ISOCAT: <http://www.isocat.org>

Resource Curation Project (2)

- Resource must be properly documented
- Resource must be made available on a server of a designated CLARIN centre
- All aspects must be tested, e.g.
 - Metadata harvesting test
 - Testing against XML schemata
 - ...

Demonstrator Project (1)

- Adapt existing tool / application into
 - A web-based application with
 - Clearly separated user interface
 - Core component for which an API is defined
 - Can be turned into a web-service if desired/required
- Provide Data for the demonstrator

Demonstrator Project (2)

- Provide Demonstration scenario
 - Optionally movie or screen capture sequence
- Provide proper documentation
- Metadata description
- Make it available on a server of a designated CLARIN center

CLARIN centres

- No recognized CLARIN centers exist yet
- The following institutes are working towards it:
 - MPI, Nijmegen ('typological data')
 - INL, Leiden ('lexical and textual data')
 - Meertens, Amsterdam ('dialect data')
 - DANS, The Hague

Requirements and Desiderata

- Data and Metadata
 - Requirements for data formats and encoding standards
 - IPR / restricted use / ethical issues requirements
 - Metadata elements and formats
- Tools / Applications
 - Processing requirements
 - Memory requirements
 - Network Bandwidth requirements
 - API requirements (e.g. Calling conventions)
- Infrastructure
 - Web-service wrappers
 - User workspace requirements
 - Repository Requirements
 - Requirements for registering and resolving PIDs
- Semantic interoperability
 - Data Category Registry
 - ISOcat

Specific Priorities

1. Literary studies
2. History and political studies
3. Communication and media studies (focused on language data and tools)
4. 1st and 2nd language acquisition
5. Historical linguistics

Evaluation Procedure

- Evaluation and Ranking by International Advisory Panel (IAP)
- Evaluation and Ranking by National Advisory Panel
- Decision by the Board
- Evaluation Criteria described in the Call

Practical Matters

- **Duration:** max 12 months
 - Exceptions allowed if justified
- **Budget per project:** max 120,000 Euro
- **Total budget (projects):** max 600,000 Euro

Practical Matters

- **Who:**
 - researchers employed by CLARIN-NL participants that have signed the CLARIN-NL consortium agreement.
- **Eligible Costs**
 - Personnel in accordance with NWO-VSNU agreement
 - Fee for travel and subsistence
 - Max 3000 Euro / FTE / year



Practical Matters

- Use Proposal Template
- Proposal language: English



Time Line

Activity	Date
CLARIN-NL Call 2 Open	Tuesday July 6, 2010
Call 2 Info Session	Thursday August 26, 2010
Deadline Proposal Submission	Wednesday September 29, 2010 13:00hrs
Assessment IAP	Friday October 29, 2010
Assessment NAP	Tuesday November 30, 2010
Decision on Funding by Board	Wednesday December 15, 2010



Official Documents

- Can be found on the CLARIN-NL website
- Since July 6, 2010

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

Questions

- For any technical questions:
 - Contact the CLARIN-NL Helpdesk!
 - helpdesk.clarin@uu.nl
- For any questions related to this proposal, or for any desire to know about user's research questions / data / tools:
 - Contact the CLARIN-NL office!
 - clarinnl@uu.nl



Thanks for your attention