



# CLARIN-NL Requirements and Desiderata

Jan Odijk

CLARIN-NL Call 3 Info-session

Utrecht, 25 Aug 2011



# Overview

- Requirements and Desiderata (RDs)
- Overview of RDs
- Examples and Actions Taken
- Conclusions

## RDs: come from from Project Goal (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

# RDs: Examples

- 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)



# RDs: Examples

- Infrastructure
  - Web-service wrappers
  - User workspace requirements
  - Repository Requirements
  - Requirements for registering and resolving PIDs
- Semantic interoperability
  - Data Category Registry
  - ISOcat

# RDs: Actuals (5 projects)

• Row Labels	Count of Topic
• ISOCAT	8
• Web Services /Tools	7
• Metadata	5
• Project Management	3
• IPR & Ethical Issues	2
• Standards	2
• PIDs	1
• Desired Functionality	1
• Search	1
• <b>Grand Total</b>	<b>30</b>

# RDs: Examples

- ISOCAT
  - Semantics of DCs ill-defined
  - Too many closely related DCs
    - Which of these is going to be standardized?
  - I need the same mappings as another project, but probably more: which ones?
- Action:
  - Several workshops to discuss problems and possible solutions
  - Several ISOCAT tutorials held
  - Appointment of ISOCAT coordinator Ineke Schuurman

# RDs: Examples

- Tools / Web services
  - Authentication not implemented-everyone can see everyone's files
  - Problems with performance in case of large outputs
  - CLAM documentation can be improved
- Action:
  - Support for web services and CLAM by Maarten van Gompel (now Uvt soon RUN)
  - To start as soon as possible



# RDs: Examples

- Metadata
  - Metadata profile for tools is needed
  - Improvements for Metadata Profile / Component Editor desired
- Action:
  - Regular Metadata tutorials
  - Desired improvements discussed
  - Budget reserved for creating a metadata profile for tools / applications and web services
    - Plan under development
    - Looking for specialist

# RDs: Examples

- Project Management
  - Coordinate latter tasks with initial tasks at the beginning of the project!
  - Organize the project in such a way that one person leaving is not disastrous
- Action:
  - We mention these here so that you can take them into account



# RDs: Examples

- Standards
  - My data do not fit any standard supported by CLARIN
  - (in such a case: First ask a recommendation from the HelpDesk before continuing!!!)
- Action:
  - We are working on making the supported standard list a dynamic registry
  - We are working on a policy for approving additions to the CLARIN Supported Standards

# RDs: Examples

- IPR & Ethical Restrictions
  - My data are partially publicly accessible but
  - Full accessibility only if a special license is signed and approved
- Action:
  - We recommended to work with two sets of metadata
  - We take this into account in designing procedures / protocols for handling IPR & Ethical Issues in the CLARIN-infrastructure

# RDs: Conclusions

- We all learn a lot from actually carrying out resource curation and demonstrator projects
- Make sure your Req and Des document contains all problems / wishes/ requirements you encountered
- The EB already has taken action and will continue to take action to address matters raised in these documents
- Using these documents will help us achieve compatibility of the CLARIN infrastructure with YOUR data/tools



Thanks for your attention!