



State of integration and next steps

Steven Krauwer, Bente Maegaard,
Dieter van Uytvanck

Point of departure



- Towards Integration of the CLARIN Technical Infrastructure, Content and Services
- Version 1.1
- August 2012

Purpose of the integration effort



- Create common solutions
- Avoid duplication of work
- Obtain common and homogeneous interfaces
- Avoid fragmentation
- Support interoperability at the European level
- At the same time recognize special needs at the national level

Objective of the first plan



- Identify layers of integration of the technical infrastructure and of the integration of language services (LR, tools, knowledge sharing)
- At each level identify those aspects that are most crucial, and identify those responsible

The layers



1. Purely infrastructural integration:
Authentication and Authorisation
2. Basic services (metadata provisioning,
metadata harvesting, and metadata
browsing and search)
3. Integration of object data and tools

Milestones and progress for each layer (tech)



M1-1

List of national identity federations

- Done

Members who do not have national identity federation should get one

- Missing: PL, BG

Milestones and progress for each layer (tech)



M1-2

Review the need for technical infrastructure workshops

- Done, and several workshops have been org.

M1-3

Integration of national identity federations into CLARIN SPF

- Done and ongoing

Milestones and progress for each layer (metadata)



M2-1

Compile a list of different types of resources and tools offered, and metadata

- Done

M2-2

Begin conversion of metadata into CMDI

Familiarise participants in Sofia with tools for metadata provisioning, harvesting and browsing

- Done

Milestones and progress for each layer (metadata)

M2-3

Establish repository for resources and tools within the national type B (A) centre, implement OAI-PMH provider endpoint

- Done for some centres but not all. Missing: BG, DLU, EE, PL (AT on its way)



Milestones and progress for each layer (interoperability)



M3-1

Nominate resource types and tools as candidates for standardisation effort (showcases)

- Done nationally, not published, but showcases are being implemented

M3-2

Presentation of the first federated content search specification (Sofia)

- Done, but also ongoing

Milestones for centre assessment



Added in the Oct. 2013 version

M4-1 (March-June 2013)

- Done: 9 centres obtained B status

M4-2 (October-Nov 2013)

- Ongoing (CZ, DK, 3 NL centres)

M4-3 (February-March 2014)

- The remaining centres

Next steps for integration



Elements for integration can be developed centrally or in collaboration between national teams

The next slides give some examples of central infrastructure components that have already been developed or are in the process of being developed and that may be good candidates for central efforts

Central infrastructure components



- **Registries**
 - Centre registry
 - Virtual Collection registry
- **Metadata processing**
 - Virtual Language Observatory
 - OAI harvester
- **Federated Content Search**
 - Aggregator
- **Federated Login**
 - Discovery Service
 - CLARIN Identity Provider

Central infrastructure components



- **Validators**
 - OAI-PMH (metadata)
 - SRU-CQL (federated content search)
- **Development services**
 - CLARIN Subversion + Trac
 - Nexus Maven repository
- **WebLicht and other workflow tools**
- **LRT inventory**

CLARIN-ERIC SCCTC Plan 2013



	1	2	3	4	5	6	7	8	9	10	11	12
Fill Center Registry												
AAI: ERICize SPF												
Assessment workshop												
PIDs: analyze offers												
PIDs: MoU or advice												
Monitoring												
VLO overhaul												
FCS Aggregator												
Assessment wave 1												
Webservices overview												
Repositories												
AAI: extend SPF												
FCS Endpoints												
CMDI 1.2												
Assessment wave 2												