

Appendix D: Final Workshop

Peter Wittenburg, Daan Broeder, Remco van Veenendaal, Sven Stromqvist, Vincent Wagelaar
DAM-LR Project
18.12.2007

The final workshop of DAM-LR was organized in Lund at 8/9. December with the intention to inform in particular those institutions interested in CLARIN about the basic layer of a research infrastructure. 36 colleagues from all over Europe participated in the meeting and created a productive atmosphere. While the first day was devoted more to information on a management level, the second day was devoted to the details of the various components. The times were seen as indications to allow as much as possible space for discussions and interactions.

1. Agenda

Day 1 (10.30 - 17.30)

Session1: Overviews

10.30	Sven Stromqvist	Welcome and introduction
11.00	Peter Wittenburg	Introduction in Research Infrastructures objectives of CLARIN
11:30	Peter Wittenburg	The DAM-LR project, what it tried to achieve and why
	Discussion	
	Lunch 13:00	14:00

Session2: Current practices in integration and operability

14:00	Daan Broeder	Creating a single virtual resource catalogue, using metadata mapping, harvesting metadata via IMDI or OAI-PMH protocol
14:30	Thomas Soddemann	Interorganizational Authentication and Authorization
15:00	Daan Broeder	Identifying resources and services in a unique and persistent way
16:00	Peter Wittenburg	Agreements and Contracts
16:30	Discussion	

Day 2 (9.00 - 13.00)

Session 3: DAM-LR solutions

9:00	Vincent Wagelaar	Metadata interoperability - The DAM-LR resource catalogue
9:30	Thomas Soddemann	Using Shibboleth for AAI
10:00	Eric Auer	Using the Handle System for unique resource identifiers
10:30	Daan Broeder	Repository Systems - Fitting things together. The DAM-LR reference implementation and other solutions
11.00	Discussion	

The organizers were surprised about the many participants, since they organized the workshop at a weekend and people had to come on own funds. It was the timing on the weekend that many other interested people could not join and of course several institutes informed us that they could not participate due to a lack of funds at the end of the year. We were asked to repeat this information exchange in the CLARIN framework.

2. Comments

The reason for the great interest was very obvious:

- increasingly often research institutes realize that they need to invest human resources to setup real repositories/archives to maintain their language resources in an accessible state
- increasingly more institutes realize that a network of "powerful" centers in different research domains will emerge and that this network will structure the ERA
- many core people in our field do not have any idea what it means to set up the basic layers of a research infrastructure

This brief report is not meant to discuss the details of the presentations. The slides can be found on the DAM-LR site and the topics were very close to what is reported in the annual reports and the deliverables. Instead we would like to give a few general impressions.

- The workshop was seen as very important for most, since it was the first time that people got a deep understanding what it means to setup the basic layers of a research infrastructure and to participate in a "federation".
- In the audience there were a number of decision makers and for them the first day was relevant to understand what type and how much effort is required to set up a center and to run it so that it can participate in a federation.
- In the audience there were also a number of (younger) people that will help setting up centers. For them in particular the second day was relevant to see how the components operate in detail, what the practical difficulties are etc.
- It was another moment to keep telling the relevant people how important the availability of proper machine readable metadata and a proper repository system is. Metadata is the glue to bundle related resources, to inform others about the availability of resources, the basis for forming virtual collections across archives, the basis for handling unique and persistent resource identifiers and linking them with citation information and for managing large amounts of resources in a repository. A proper repository system is the basis for all kind of management and service tasks and for integrating the holding into a federation based on integration components. For many in the field it is the first time that they were confronted with details about such infrastructure components and their requirements.
- It was the first time that the ideas about a flexible component based metadata infrastructure were discussed. It became obvious that on the one hand such a flexible component based metadata set was seen as necessary, but that this idea needs more elaborate discussions before a design can be started in CLARIN.
- The workshop was a good starting point for CLARIN, since a number of aspects were discussed that need to be taken up by CLARIN again:
 - need of a component-based metadata infrastructure (see above)
 - need of a repository system
 - need for unique and persistent resource identifiers although it requires resolving instances
 - components and architecture for the basic layers of integration
 - problems of building a language resource and technology federation and the need for a close collaboration of people working on technological issues and people working on trust agreement, IPR and license issues
 - the scope of what "language resources and technology" exactly means, what it covers and what not¹⁴

¹⁴ It was argued that from the beginning we have to link up with large providers such as Google and libraries since they have lots of language resources and since they will largely determine the world for humanities scholars due to their enormous coverage.