



**CLARIN**

# **COST ESTIMATIONS FOR CONSTRUCTION AND EXPLOITATION PHASE**

DELIVERABLE 8S-2.1a  
February 2010



Editors: Steven Krauer

The ultimate objective of CLARIN is to create a European federation of existing digital repositories that include language-based data, to provide uniform access to the data, wherever it is, and to provide existing language and speech technology tools as web services to retrieve, manipulate, enhance, explore and exploit the data. The primary target audience is researchers in the humanities and social sciences and the aim is to cover all languages relevant for the user community. The objective of the current CLARIN Preparatory Phase Project (2008-2010) is to lay the technical, linguistic and organisational foundations, to provide and validate specifications for all aspects of the infrastructure (including standards, usage, IPR) and to secure sustainable support from the funding bodies in the (now 23) participating countries for the subsequent construction and exploitation phases beyond 2010.



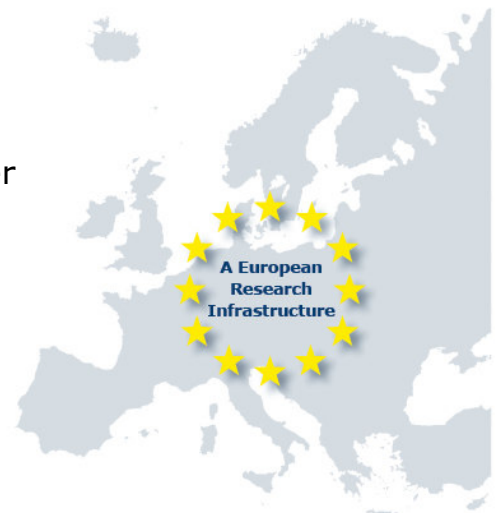
# **COST ESTIMATIONS FOR CONSTRUCTION AND EXPLOITATION PHASE**

CLARIN-2010-2

EC FP7 project no. 212230

Deliverable: D8S-2.1a

Responsible: Steven Krauwer



Contributing Partners: Peter Wittenburg, Bente Maegaard

© all rights reserved by Utrecht University (UU) on behalf of CLARIN

Contents	page
The cost of CLARIN	5
1. The governance and coordination layer	5
2. The infrastructure layer	7
3. The content layer	8

## ***The cost of CLARIN***

In CLARIN we distinguish the following three functional layers (see D8S-1.2a, *The shape of CLARIN*):

- (a) The overall **governance and coordination layer** (ERIC layer). This will take the form of an ERIC, with a General Assembly as the main decision taking body, a Board of Directors responsible for the execution of the strategies, policies and work programme decided by the General Assembly, supported by a secretariat and a coordination team.
- (b) The **infrastructure layer**. This will take the form of a number of CLARIN centers that together constitute the CLARIN infrastructure. Each of the centers will be part of a national consortium and work under the responsibility of the national coordinator.
- (c) The **content layer**. This will take the form of a number of institutions that together constitute the national consortia. They work under the responsibility of the national coordinator and create content according to the national policies and priorities determined by the national consortia.

For each of the three layers we present estimates of the costs and the way they should be covered.

### **1. The governance and coordination layer**

This is the layer of the CLARIN ERIC. The budget of CLARIN ERIC will cover the costs related to the general and overall coordination and governance of the CLARIN infrastructure. This will include both the costs of the General Assembly, the Board of Directors, administrative support and general tasks related to education and dissemination.

These costs will be covered by financial contributions from the members of the ERIC.

The calculation of the contributions is based on a 25% contribution by the hosting country, and a contribution based on the GDP from the participating countries for the remaining 75%.

In the rest of this document we will go through the various cost items, present the budget and show globally what the financial consequences would be for countries with high, middle and low GDP figures.

We have chosen for a 'lean' ERIC model, where the cost of the operations at this level is kept as low as possible. This will also minimize the cross-border flow of money.

The budget of the ERIC does not include the costs of the operation of the infrastructure itself (e.g. data and service centres, centres of expertise), as they will be borne through the contributions in kind from the members to the infrastructure.

Neither does it include the costs of the creation or conversion of digital material, tools or other content that is to populate the infrastructure. These activities will be the responsibility of the participating countries according to their own priorities and procedures.

#### **1.1. Costs**

##### *The budget items*

The main budget items are the following:

## Common Language Resources and Technology Infrastructure

### *Cost related to governance, board of directors and secretariat:*

- a. The cost of the General Assembly (GA): This covers the cost of travel for two participants (the representative and an expert) per member to the annual meeting of the GA. This cost is dependent on the number of members.
- b. The cost of the Board of Directors: This will cover the salaries of the Director (full-time, based at the headquarters, employed on the basis of a 5-year renewable contract) and 3-4 Co-directors (each contracted on a 0.4 fte basis from their home institutes, with 3-year renewable contracts). This will also cover their travel. This cost is not directly related to the number of members although we expect that the number of co-directors may go up as the ERIC grows.
- c. The cost of an annual meeting of the International Advisory Board
- d. The cost of the Secretariat of the ERIC, including a full time assistant, travel, office rent, equipment, consumables.

### *The cost of coordination activities at the general level:*

- e. Support for an annual conference, workshops, seminars, other events and travel grants
- f. Coordination of education, training, dissemination and awareness at the general level. This includes support for summer schools, workshops, training courses, the CLARIN website, a newsletter and other activities aimed at reaching out to the user community
- g. Access to legal expertise, both in relation to contractual issues at the ERIC level and in connection with IPR and related issues.
- h. The cost of international outreach. This includes the cost of meetings, workshops and other events aimed at initiating or strengthening collaboration with third countries.
- i.

### *The budget*

The table below reflects the expected annual expenditure at the CLARIN ERIC level.

Activity	Labour	Travel	Events	Other	Total
<b>governance and secretariat</b>					
General Assembly	0	20000	0	5000	25000
Director	120000	24000	0	10000	154000
Co-directors	192000	72000	0	10000	226000
International Advisory Board	0	0	15000	0	15000
Office assistant	80000	6000	0	5000	97000
	<b>392000</b>	<b>122000</b>	<b>15000</b>	<b>30000</b>	<b>559000</b>
<b>coordination</b>					
dissemination, training	16000	24000	30000	10000	80000
technical workshops and seminars	40000	12000	60000	10000	122000
website & newsletter	32000	12000	0	10000	54000
annual conference	0	6000	0	60000	66000
legal expertise	16000	12000	0	0	28000
international contacts	16000	6000	30000	0	52000
	<b>120000</b>	<b>72000</b>	<b>120000</b>	<b>90000</b>	<b>402000</b>
<b>Total annual central budget CLARIN ERIC</b>					<b>961000</b>

## 2.1 Income

The expenditure is covered by annual contributions from the members. The distribution is as follows:

- The host country (currently offered by the Netherlands) covers 25% of the total general cost of governance and coordination.
- The other members will cover the remaining 75%.
- For each country the share of the contribution is calculated as the GDP of the country divided by the sum of the GDPs of all member countries.
- For members other than countries the share will be determined during the accession negotiations between the ERIC and the entity.

If the ERIC agrees to fund other activities than governance and coordination (e.g. supporting specific infrastructure services, research projects, digitization programmes or similar) the cost of such operations will be distributed over all members, according to the same distribution key. As the individual contributions from the countries will not only depend on how many countries participate and on their GDP it is not possible to make precise predictions, but given an annual budget around 1 M€ and participation by 10 countries a big country would pay ca 370 K€, a medium sized country ca 50 K€ and a small country ca 15 M€ per year.

## 2. The infrastructure layer

### 2.1. Costs

This is the layer of the centres constituting the technical infrastructure and the coordination of these centres at the national level. Document D2R-9a (*Cost Estimates - v1*) gives an analysis of the cost of the construction and operation of the technical infrastructure and its components.

Country size and nr of centres	annual cost based on 60 k€ per fte per year			
	construction phase		exploitation phase	
	with centres	without centres	with centres	without centres
large country (6 centres)	7.0	4.0	5.4	2.4
Medium size country (3 centres)	3.5	2.0	2.7	1.2
small country (1 centre)	1.3	0.8	0.9	0.4

The table shows that the cost depends on a number of factors that are all dependent on the specific situation in a country:

- the number of centres a country wants to host
- the extent to which existing centers are already equipped to play a role in CLARIN
- the salary level (the table is based on 60 k€ per fte per year)

It also shows that we expect the costs to decrease after the initial construction.

For 10 countries (e.g. 2 large, 4 medium and 4 small) the total number of centres would amount to 28, but please note that it is up to the countries to decide how many and what type of centres they want to host. Contrary to the contributions at the governance and coordination level there is no link between a country's GDP and its contribution to the infrastructure level of CLARIN.

### 2.2. Income

During the construction and exploitation phase the costs of the centers will in each country be covered by the funding agencies and/or by the centres themselves, as part of the agreed upon CLARIN Package. No central contribution via the ERIC or from the EC is foreseen.

Three comments should be added here:

- (1) When the number of participating countries grows the CLARIN ERIC General Assembly may decide to shift the cost of certain generic infrastructure services that are shared by all to the CLARIN ERIC level, to be covered from the national contributions to the ERIC.
- (2) According to messages from the EC there are plans to make provisions in the 8<sup>th</sup> Framework Programme that would allow Research Infrastructures to receive contributions from the EC of up to 20% of the operational costs. If this is implemented through Calls for Proposals the CLARIN ERIC would be eligible to submit such a proposal and could channel the money to the centres that constitute the infrastructure.
- (3) In addition the ERIC could submit proposals in response to other EC Calls and use the funds to support new or ongoing CLARIN activities

### **3. The content layer**

This is by far the most important layer, because this is where the infrastructure will be populated with data (by integration of legacy data or creation of new data) and tools and services (legacy or new).

Yet CLARIN ERIC will not impose any other obligations on the members than that they commit themselves to the creation of new data and tools, both connected to their own language(s) and to the languages that are relevant for their research communities.

This is because the creation of language resources and tools cannot be seen as a stand-alone activity, and should always serve national or European RTD programmes, which follow their own priorities.

CLARIN ERIC will ask every member to specify at the beginning of each year what plans they have for the creation of content, and it will also offer its members opportunities to embark on joint content creation projects. Models such as the ESF a la carte will be a good instrument to initiate and fund joint projects.

CLARIN ERIC will also be in a position to submit proposals to EU programmes for projects aimed at the creation of content or tools.