



CLARIN

Cost Estimations

Deliverable D2R-9a

2010-02-18 - Version: 3



Editor: Peter Wittenburg

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.



WG2-9

Cost Estimations

EC FP7 project no. FP7-RI-2122230

Deliverable: D2R-9a - Deadline: 1.1.2010

Responsible: Peter Wittenburg

Contributing Partners: MPI, INL, OTA, RACAI, WROCUT, UPF, ELDA, ILSP, ILC, USFD, DFKI, CSC, UIL-OTS, ULeuven, ATILF, UTuebingen, HASRIL, CST, UTartu



Scope of the Document

This document describes the costs that need to be taken into account when building up a research infrastructure as intended by CLARIN. The costs will occur at CLARIN ERIC, national and institutional level. The costs at CLARIN ERIC level are described in a separate document (D8S-2.1a, *Cost estimations for construction and exploitation phase - v1*). This document refers to the centre types and the requirements for centres as described in earlier CLARIN documents.

This document will be discussed in the appropriate working groups and in the Executive Board. It will be subject of regular adaptations dependent on the progress in CLARIN.

Chapter 1 of this document gives a short introduction and describes some constraints; chapter 2 is devoted to discuss national and organizational aspects; chapter 3 specifies costs along a number of dimensions; chapter 4 makes some rough calculations for large, medium and small countries; and chapter 5 gives rough summary of the annual costs for the construction and exploitation phase. Chapter 6 is an appendix where we refer to costs calculations that we are aware of.

CLARIN References

- CLARIN Centre Types CLARIN-1-2008 May 2008
- CLARIN Centres CLARIN-3-2008 May 2008
- Cost estimations for construction and exploitation phase - v1 D8S-2.1a Feb 2010

Contents

Executive Summary	6
1. Introduction	7
2. National and Organizational Aspects.....	7
2.1 National Aspects	7
2.1 Organizational Aspects.....	8
3. Cost Components	9
3.1 Resource & Service Centres	9
3.2 Resource & Tool Integration	9
3.3 Infrastructure.....	10
3.4 Support & Advice	10
3.5 Training & Education.....	10
3.6 Coordination & Management	11
3.7 National and EU Level Offices.....	11
3.8 External Services and Licenses.....	11
4. Example Calculations.....	12
4.1 Large Country	12
4.2 Medium Size Country	12
4.3 Small Country	12
5. Cost Summary	14
6. Appendix: Existing Calculations	15
5.1 Beagrie Cost Estimates.....	15
5.2 MPI/DOBES Repository Calculation.....	16
5.3 CLARIN NL Calculation	16
5.4 Calculations in Germany	17
5.5 CLARIN Startup Calculation.....	17

Executive Summary

Estimating costs for research infrastructures is still not a matter where we can refer to exact data gathered from experience. However based on own activities in CLARIN and some indications from other sources we can make improved guesses which widely confirm first estimates. One source of complication for estimating costs can be found in the various solutions to be found dependent on the national situations and the forms of institutional organizations. Ignoring these differences in this version we can make estimates for a number of relevant cost dimensions (resource & service centres, resource & tool integration, infrastructure, support & advice, training & education, coordination & management, national & EU level offices and external services).

Based on these individual cost dimensions we can make example calculations for a large country, a medium size country and a small country where the terms large, medium and small do not refer only to the geographical size or to the number of people, but also to the ambitions of maintaining the languages spoken. Assuming a certain number of centres as the backbone for the CLARIN infrastructure, where the number is correlated with the type of country, the calculation of the average annual cost amounts to 6.6 mio € for a large, 3.3 mio € for a medium size and 1.2 mio € for a small country.

Of course these sums need to be seen on the background that many countries have already invested in centres which come close to the CLARIN requirements, i.e. the costs will be reduced by some factor.

1. Introduction

One of the major tasks of the preparatory phase is to make new cost calculations based on own tests and experiences and those of others. Yet there are not so many open and reliable calculations from other initiatives that work on distributed research infrastructures. In this first version of the deliverable we describe the costs of a number of components and base our calculations on (a) own experience, (b) numbers that were presented in the Netherlands and Germany and (c) on statements made at the Budapest conference of the Alliance for Permanent Access (<http://www.alliancepermanentaccess.eu/index.php?id=3>) by Keith Jeffery, Neil Beagrie and Peter Wittenburg. We can also refer to the statements made in the Interim Report of the Blue Ribbon Task Force on Sustainable Digital Preservation (Dec 08) which only makes statements about the preservation costs but which can be extended to a certain degree to our cost calculation challenges. It is stated that *“they did not have data available regarding the costs of digital preservation per se, largely because their organizations are in the very early stages of understanding exactly what digital preservation entails for them. Often the preservation activities are tied to other production or access-related activities and determining costs specific to preservation would require a largely manual endeavour of parsing the activities from budgets that are not structured to highlight preservation as a separate activity”*. Traditional types of centres who have been established in our field have other tasks not related to infrastructure activities as well making it hard to get reliable figures at this very moment.

In case of personnel all numbers presented in this document are estimated in FTE and in € (rounded) based on an assumed salary of 60 k€ although we know that the salary levels are very much differing, i.e. every expert per country can make his own specific calculations. The calculations simply include personnel and larger equipment costs, thus they do not include overhead, housing, etc. No revenue costs are considered although we know that in some countries this is expected.

2. National and Organizational Aspects

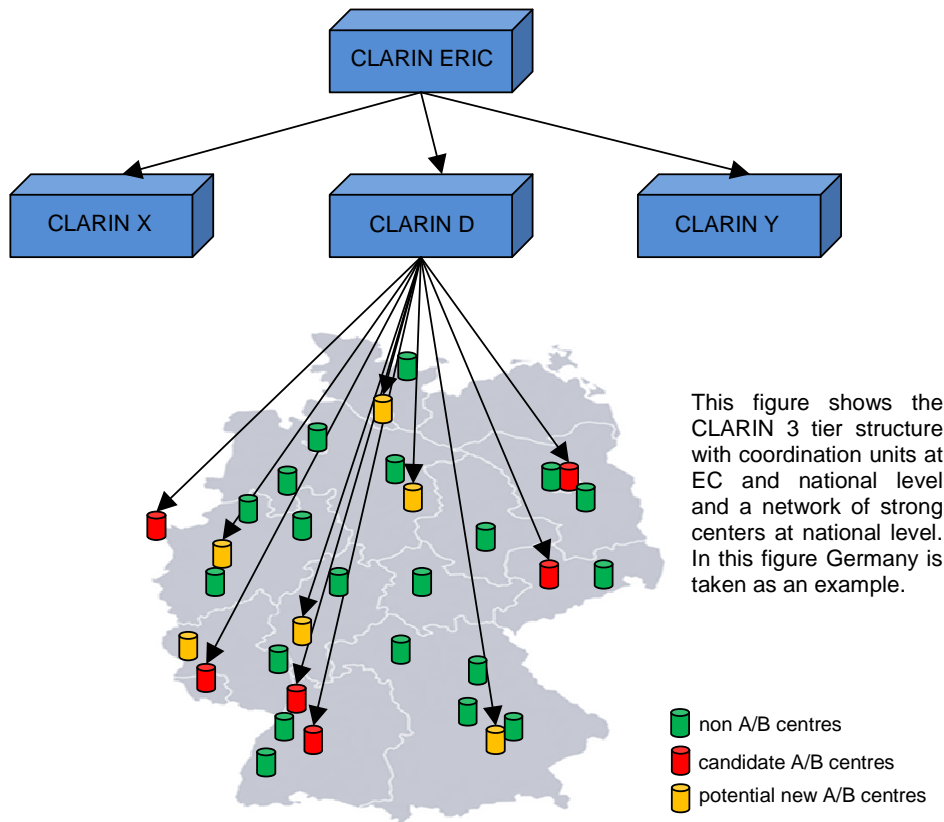
One of the reasons why it is so difficult to come to agreed cost calculations can be found in the differences in organizational forms and cultures that are applied in the different member states. Within the following years one needs to respect these differences, since (a) infrastructures can only be established successfully when the national stakeholders are convinced and (b) when the infrastructures - and here in particular the centres - can be built up on already existing institutions and knowledge. Starting to build an infrastructures and proper centres from scratch would in general not be efficient and require much time, and would cost additional money. Therefore we need to look at the national and organizational aspects. In this version we will only present a summary of the relevant aspects, but we will not try to make differences with respect to the cost calculations.

2.1 National Aspects

Similar to GEANT, CLARIN will be organized as a 3 tier structure as shown in the following figure. At the European level we will have an ERIC, where CLARIN prefers to have a lean organization model (see below). The CLARIN ERIC will be a legal body and coordinate all CLARIN activities such as the integration and interoperability tasks. At the national level we will have national representatives that coordinate all national activities. Here we can clearly see different forms of organizations already: (1) there are some countries that just have one national centre for language resources and tools; (2) there are other countries which will have a number of centres for LRT so that coordination is required; (3) we can already foresee in a few countries that there is a trend to build one national organization that will represent CLARIN and DARIAH, i.e. an organization responsible for the arts and humanities infrastructure. At the third tier we have the CLARIN centres that comply with the A/B type centres requirements which form the pillars of the infrastructures. In most countries we can see that these A/B type of centres are being implemented at locations where already a certain tradition has been established.

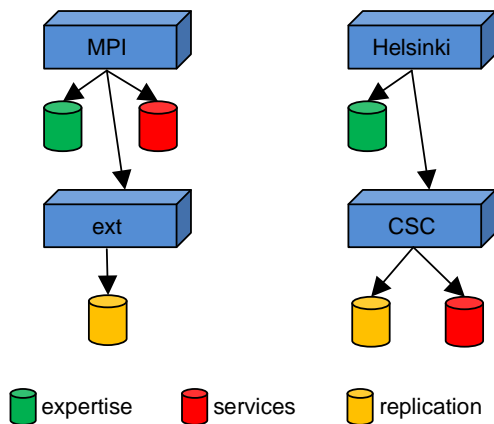
In addition we will have many institutes with linguistic departments which will not have the strength to become A/B type centres, but still want to participate by offering for example metadata of certain special resources or tools for example.

A final cost scheme needs to consider all these variants, but to accomplish this we need a national plan for each country.



2.1 Organizational Aspects

Another dimension that is crucial for cost calculations has to do with the distribution of roles when acting as a centre. In the following we will discuss 2 models, but others can be thought of.



In this figure we indicate differences in organizing a centre. The MPI model includes a unit that takes care of all aspects of a repository, i.e. it provides expertise at all levels and also offers access services. Only data replication for preservation purposes is outsourced. In the Helsinki model the linguistic expertise is kept at the linguistic department, but all services and the replication are outsourced to a service provider. None of these two models is per se optimal or non-optimal, optimality depends on many local constraints.

Also the basis of calculations and efficiency aspects depend very much on the local situation. In the MPI case we have an almost self-contained unit that is focussing on linguistic research which needs a strong service centre anyhow to satisfy the researcher's needs. In this case economy of scale and related arguments suggest offering these services and expertise also to the outside world since data needs to be shared anyhow. In the Helsinki case where university departments do not want to maintain expensive IT-staff to take care of data management and to offer services to others and where the national computing centre has a clear service task and has some linguistic expertise the decision may tend towards sharing the services task with others, thus economy of scale is achieved in a different way.

Whatever solutions are applied, we will see that these may change over time dependent on how technology will develop. Also differences with respect to organizational solutions are not included in this version.

3. Cost Components

In this chapter we want to briefly introduce the various cost components (dimensions) that need to be taken into account when estimating the costs of an infrastructure for language resources and technology. This list needs to be comprehensive to achieve a complete coverage. For each cost component we describe a few formal attributes that will make it easy to summarize at the end.

The attributes are the following:

- **variable/constant:** indicating whether the costs are varying over the different phases
- **centre/total:** indicating whether the costs need to be multiplied by centres where the number of centres will differ per country or whether it is meant as a total per country or EU level
- **existing/new:** indicating whether the costs are partly already spent or whether they are in addition to what is already existing
- **level:** indicating whether the costs appear at institutional, national and/or European level

3.1 Resource & Service Centres

costs	0.5 m€/year	these are costs just to run a centre, it does not cover the curation and integration of resources or tools
personnel	7 fte	
equipment	0.08 m€	
variable/constant	constant	additional funds may be required to build up a centre from scratch, these depend on a number of factors
centre/total	per centre	the costs may vary dependent on the commitments and the scope of services
existing/new	existing	in general one will build centres on ongoing activities of research organizations/institutions which will reduce the explicit costs
level	institutional/national	CLARIN service centres mainly will exist at institutional level, but there will also be national and EU level centres, however these centres need to be lean and just offer aggregation services; in many cases some of these centres will collapse, i.e. one of the institutional centres will become the national centre

A network of centres will guarantee the stability and persistence of an infrastructure. There will be a variety of solutions for offering services as has been described in the CLARIN centres requirements documents. A linguistic institute may decide to outsource the real services to an existing computer centre or library for example that is willing to fulfil the CLARIN requirements and restrict itself only to giving advice etc. Other institutes may want to take over all services. Whatever solution will be found the necessary expertise needs to be provided. Costs will largely depend on the concrete institutional embedding, since economy of scale factors and cost sharing will play a large role. Therefore the costs mentioned above per centre can only be seen as a baseline indication.

3.2 Resource & Tool Integration

costs	0.6 -> 0.2m€/y	costs will drop to a base level over time
personnel	10 - 2 fte	
variable/constant	variable	at the beginning the costs to curate resources and adapt tools will be higher so that a maturation effect can be expected
centre/total	country total	costs are estimates for meant per country
existing/new	existing	we expect resources and technologies to be already existing so that simply the curation and adaptation needs to be funded; new resources and tools need to be created under other funding schemes
level	national	funds need only to be available at national level

The costs for adapting and integrating existing resources and technology are huge due to the necessary curation efforts and it is a matter of national policy to set priorities and time scales for the transition period. Therefore the indicated costs will start at a high level and settle on a much lower stable level over time. There will always be a curation and integration effort due to new standards and new resources not being in the correct standards or not having sufficient quality.

3.3 Infrastructure

costs	0.8 -> 0.4 m€/y	costs will drop to a base level over time
personnel	12 - 7 fte	
variable/constant	variable	at the beginning the costs for constructing the infrastructure will be much higher than for operating and maintaining it after the start up phase
centre/total	country total	the construction and maintenance costs occur at national and EU level, the costs for running an EU level centre are indicated below
existing/new	new	infrastructures in general need to be built from scratch
level	national	funds need only to be available at national level

The construction costs are at the beginning comparatively higher than for running it. The costs will drop down to a basic level that needs to be reserved at country level. Funds for aggregated services at European level are mentioned below.

3.4 Support & Advice

costs	0.0 -> 0.2/y	costs will increase to fixed level
personnel	0 - 2 fte	
variable/constant	variable	at the beginning the costs will be minimal and start increasing with the number of users until an upper value
per centre/total	country total	the construction and maintenance costs occur at national and EU level, the costs for running an EU level centre are indicated below
existing/new	new	this has to be setup from scratch
level	national	funds need only to be available at national level

The DEISA project indicates how important support and advice is in particular in distributed scenarios where it is not easy to trace errors and where it is not easy to enforce changes. So a fixed level of costs needs to be assumed to help researchers.

3.5 Training & Education

costs	0.2/y	costs will stay relatively constant
personnel	2 fte	
variable/constant	constant	costs will not change too quickly since new generations of students need to be trained
per centre/total	country total	the construction and maintenance costs occur at national and EU level, the costs for running an EU level centre are indicated below
existing/new	new	this has to be setup from scratch
level	national	funds need only to be available at national level

Here the same statement holds as in 2.4. Currently we need to educate this researcher generation. Later we need to train our young people until the new methods are commonality.

3.6 Coordination & Management

costs	0.1/y	
personnel	1.5 fte	
variable/constant	constant	
per centre/total	country total	
existing/new	new	
level	national	

At national level some funds need to be available to manage an infrastructure, the functioning of the centres network, quality assurances, aggregation services and portals, definition of standards etc. The tasks are very well known.

3.7 EU Level Governance and Coordination (CLARIN ERIC)

With respect to the cost at the CLARIN ERIC level we refer to a document D8S-2.1a (*Cost estimations for construction and exploitation phase - v1*).

3.8 External Services and Licenses

costs	0.1/y	
personnel	0	
variable/constant	constant	
per centre/total	country total	
existing/new	new	
level	national	

We will need to make use of services offered for example by horizontal infrastructure service providers, to buy licenses for various software components and to reserve funds to access other than academic data and tools. Yet we cannot make proper estimates, but we know that a variety of services such as for persistent identifiers, for data persistency services etc funds may need to be available if they are not funded as a general service such as computer networks.

4. Example Calculations

Three example calculations are made that may serve as a guideline for many of the CLARIN partners to choose the right model for their country.

4.1 Large Country

For a large country such as Germany we assume to have 6 centres in the LRT field. This will amount to the following costs:

topic	constr	oper
centres	0.5 x 6	0.5 x 6
infrastructure	1.6	0.8
resource curation & tools adaptation	1.2	0.4
support line	0.4	0.4
education/dissemination/training	0.4	0.4
coordination	0.2	0.2
external services	0.2	0.2
total with centres	7.0 per year	5.4 per year
total without centres	4.0 per year	2.4 per year

For reasons of simplicity we calculate a factor of 2 for the costs for the running infrastructure compared to its construction. The same holds for the effort for integrating resources and tools. The calculations are made with and without costs for centres, since in many cases it may be expected that the centres already exist and receive regular funding from their research organizations or funding agencies.

As baseline costs for a large country we see costs ranging from 4.0 (7.0) Mio€ in the construction phase going down to 2.4 (5.4) Mio in the operational phase.

4.2 Medium Size Country

For a medium size country with ambitions such as the Netherlands we assume to have 3 centres in the LRT field. This will amount to the following costs:

topic	constr	oper
centres	0.5 x 3	0.5 x 3
infrastructure	0.8	0.4
resource curation & tools adaptation	0.6	0.2
support line	0.2	0.2
education/dissemination/training	0.2	0.2
coordination	0.1	0.1
external services	0.1	0.1
total with centres	3.5 per year	2.7 per year
total without centres	2.0 per year	1.2 per year

The same rules as above are applied.

As baseline costs for a medium size country with ambitions we see costs ranging from 2.0 (3.5) Mio€ in the construction phase going down to 1.2 (2.7) Mio in the operational phase.

4.3 Small Country

For a small country such as perhaps Latvia we assume to have one national centre in the LRT field. This will amount to the following costs:

topic	constr	oper
centres	0.5	0.5
infrastructure	0.2	0.1
resource curation & tools adaptation	0.1	0.05
support line	0.1	0.05
education/dissemination/training	0.2	0.1

coordination	0.1	0.05
external services	0.1	0.05
total with centres	1.3 per year	0.9 per year
total without centres	0.8 per year	0.4 per year

The same rules as above are applied.

As baseline costs for a small country with ambitions we see costs ranging from 0.8 (1.3) Mio€ in the construction phase going down to 0.4 (0.9) Mio in the operational phase.

5. Cost Summary

Here we summarize the annual costs for the construction and exploitation phase for large, medium size and small countries. Please note that the estimations are based on annual salary costs of 60 k€ per fte.

country size	annual cost based on 60 k€ per fte per year			
	construction phase		exploitation phase	
	with centres	without centres	with centres	without centres
large country	7.0	4.0	5.4	2.4
medium country	3.5	2.0	2.7	1.2
small country	1.3	0.8	0.9	0.4

6. Appendix: Existing Calculations

In this appendix we want to refer to a number of cost calculations of different sort that were presented during the last two years in particular at the Budapest conference of the APA (Alliance for Permanent Access; <http://www.alliancepermanentaccess.eu/index.php?id=3>).

Here we mainly refer to numbers about the costs of centres storing and providing access to data and about infrastructures. With respect to costs for data persistency Keith Jeffery pointed out that the "market is imperfect" since we cannot specify the value of research data. We can extend it to stating that we cannot specify how much money data centres or infrastructures giving smart access to data may cost. It is a matter of priorities of how much money needs to be taken from the overall research budget to be spent. Knowing that for example 80% of our recordings about languages and cultures are highly endangered of become unreadable we can only conclude that obviously we would need to invest a huge amount of money if we wanted to preserve all research data.

5.1 Beagrie Cost Estimates

Resource Life Cycle Costs

At the APA meeting in Budapest Neil Beagrie presented results from a UK overview which shows how the costs are distributed across the lifetime of resources. It turns out that the costs for acquisition and ingest are highest due to the required curation costs. Also the costs for providing access to users are comparatively high. The lowest are the cost for archival storage and preservation.

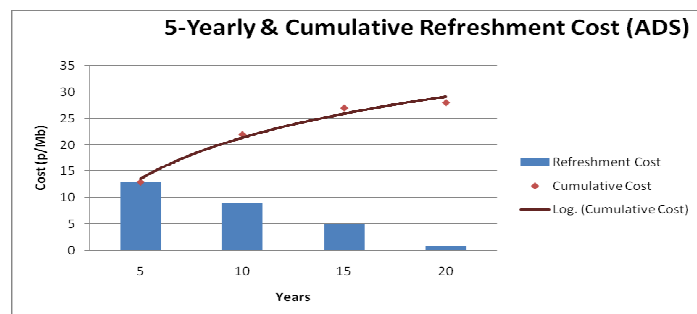
Acquisition and Ingest	Archival Storage and Preservation	Access
c. 42%	c. 23%	c. 35%

Actually we assume that the real costs for acquisition and ingest (including curation) and the costs for access will be even higher in an eResearch scenario where all kinds of existing resources need to be curated to become usable and where user communities expect a lot of different access ways.

From this summary we can conclude that it is important to reserve a continuous amount of money to carry out curation on existing data and maintenance of useful tools.

Storage Costs

It is well-known that the storage costs follow an asymptotic function. In addition to what Beagrie describes we can refer to the conclusion of the MPG IT council which stated that the storage capacity which we currently need will be less than 10% of the capacity which we will have after a complete technology innovation cycle (~10 years). The prices per MB will follow roughly.



From this we can conclude that the pure storage costs of the data which we collected until now can be neglected in our calculations.

Data Repository Costs UK

In another case study Beagrie compared costs for running data repositories in the UK. He found out that U Cambridge spent 4 FTE to run such a repository and Kings College London 2.5 FTE. Nothing is said about the type of data and the size of the repository.

In addition Beagrie describes the possible economy of scale effect: a 600% increase in accessions will only increase costs by 325% (ULCC), i.e. much saving is possible by concentrating resource storage and management¹.

We don't know exactly what creates the differences between institutions such as U Cambridge and Kings College, but from all we know we can assume that 2.5 FTE is just the bare minimum to keep a repository running. If we add aspects as indicated in 5.2 the calculation of 4 FTE (~ 0.3 Mio€) seems to come close of what is needed. Of course economy of scale factors are important.

5.2 MPI/DOBES Repository Calculation

The MPI/DOBES archive² is housing about 50 Terabyte of multimedia/multimodal resources in a Hierarchical Storage Management system, is replicating them at 4 other institutes and has the LAT (Language Archiving Technology³) software suite to manage the data and give access to it. It serves currently about 46 mostly external DOBES documentation teams, about 100 internal researchers and about 20 external researchers who have deposited their data at the MPI archive. Thus in total there are about 270 depositors if we calculate an average of 3 persons per DOBES team. Based on 10 years of experience the following cost model was presented at the APA meeting in Budapest:

type	k€/y	comment
basic IT infrastructure	80	4-8 years innovation cycle
digitization and workflow	10	new recorders, capturing dev
copies (30 TB) at 4 large computer centers	<5	currently no explicit costs are paid
system management	60	shared for different activities
archive management	80	advice, curation, consistency
repository software maintenance	60	without new functionality
utilization software maintenance	>120	wide spectrum of tools
building, energy, etc	?	ignored here
total	415K€	

This cost calculation assumes that the centre is embedded in the institute, i.e. no costs have been calculated for water, electricity, building etc and management of such a group. The specified costs are not explicit since the institute already employs some of the experts involved and has a clear IT infrastructure policy for its own purposes. In many respects economy of scale factors can be applied here, i.e. the costs to host additional resources does not result in significantly higher explicit costs. The costs for software maintenance are at the bottom line. In particular if one wants to maintain a complex software suite as the LAT with its many professional components the costs are much higher.

Adding some additional costs for items not mentioned here the baseline costs for such a centre can be specified as 500.000 € per year. If only the personal costs are calculated these numbers are similar to those from U Cambridge.

5.3 CLARIN NL Calculation

CLARIN NL presented a calculation for running costs with the following pillars:

topic	infra	09-10 (prep)	11-12 (constr)	13-14 (oper)
technical construction	x	0.17	0.27	0.1
resource curation	(x)	0.27	0.32	0.17
tools & services adaptation	(x)	0.30	0.34	0.1
use cases		0.25	0.34	0.1
expertise centres	x	0.03	0.13	0.13
education/dissemination/training	x	0.1	0.14	0.1

¹ This finding is inline with the observations at the MPI which allowed MPI to open its archive for external depositors.

² <http://www.mpi.nl/resources>

³ <http://www.mpi.nl/tools>

coordination	x	0.15	0.14	0.13
ERIC contribution	x	0.06	0.2	0.41
total		1.3 per year	1.9 per year	1.3 per year

Thus the Dutch will invest considerable funds (about 60%) to carry out resources and tools curation and integration and to carry out use case projects to involve the researchers. These are not requirements for building a running infrastructure, but will be essential to populate the infrastructure with useful things and to convince researchers. But every country needs to make its own choices and set its own priorities. Infrastructure builders are involved in the curation and integration projects as well as in the use cases to keep a close relationship. Some funds will be reserved to help centres to meet the requirements.

The running costs for an LRT infrastructure are estimated as about 1.3 Mio € per year taking into consideration that the centres will mainly be funded by their institutions. Currently 3 centres are in discussions, 2 may join as typical tool/service providers later. It is a matter of political decision whether 60% of the funds are used for data curation and technology adaptation.

5.4 Calculations in Germany

It may be known that in Germany it was suggested to the ministry that CLARIN and DARIAH will go together to make a joint proposal that also needs to cater for including a variety of humanities disciplines. A rough sketch has been worked out and needs to be further elaborated on, i.e. the calculations are not that detailed at this moment. A few general lines can be described however, but everyone should keep in mind that many fields in humanities need to be covered.

Per year the following rough amounts have been estimated for the coming 10 years:

- 1.8 Mio€ for research communities to fund use case and curation/adaptation projects of limited size and duration
- 0.5 Mio€ for the discussion and development of joint topics of varying nature in Virtual Competence Centres such as common metadata solutions, standards and wrappers, curation etc which will be elements of the infrastructure
- 0.5 Mio€ for research communities to fund fix costs such as continuous curation, community coordinators, community portals and services
- 1.7 Mio€ for a centres and virtual competence centres to run and maintain the infrastructure and all its services
- 0.3 Mio€ for education and training activities
- 0.6 Mio€ for the management of such a complex enterprise including a large variety of disciplines
- 0.6 Mio€ to support both the European umbrella sites for DARIAH and CLARIN which is about 10%

Also in the German calculation it is expected that the centres to a large extent are funded by their home institution. We can see that there is about 2.7 Mio€ proposed to fund infrastructure work and some funding to improve centres and that the EC contribution is estimated at 10%.

If we simply assume 10 different disciplines we can see that the amount of money reserved is in average about 0.27 Mio€ per discipline which is comparable to the Dutch estimation.

5.5 CLARIN Startup Calculation

When proposing CLARIN to the commission we were asked to make a first initial cost estimation. We will briefly present the estimations here again where the costs are indicated per centres per year..

topic	infra rel	prep phase	construct phase	op phase
centres	x	0.36	0.36	0.18
knowledge centres	x	0.27	0.27	0.18
service infrastructure	x	0.9	0.9	0.27
gathering&curation resources	(x)	0.18	0.45	
gathering&curation tools	(x)	0.18	0.45	
sample applications		0.18	0.45	0.09
training & education	x	0.02	0.18	0.18
management	x	0.09	0.09	0.09

standards	x	0.02	0.02	0.02
total		2.2	3.17	1.01

If we merge the two first numbers we can see that the costs are comparable with the 0.5 Mio€ estimated at other places. We believe now that the costs for running centres will not go down over time, however, we can assume that for each centre with experience the amount of resources dealt with (economy of scale) can scale up without adding more costs. We also believe that the old assumption that there are no costs anymore required for data curation and tools maintenance was not correct.

In general we believe that the cost estimates at the very beginning were widely correct and still can be used as baseline data.