

The CLARIN logo features a stylized network of blue circles connected by lines, resembling a molecular or network structure, positioned above the word "CLARIN" in a bold, blue, sans-serif font.

CLARIN

Requirements and best Practice for Transnational Coordination and Collaboration with Third Parties

DELIVERABLE 8S-3.1
August 2009



Editors: Bente Maegaard, Hanne Fersøe, Lina Henriksen

Common Language Resources and Technology Infrastructure

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.



Requirements and best Practice for Transnational Coordination and Collaboration with Third Parties

CLARIN-2009-2

EC FP7 project no. 212230

Deliverable: D8S-3.1 - Deadline: 31.12.2008 (postponed due to late start)

Responsible: Bente Maegaard, Hanne Fersøe, Lina Henriksen



Contributing Partners: All CLARIN partners have contributed with information

Contributing Members:

Scope of the Document

This report, *Requirements and best Practice for Transnational Coordination and Collaboration with Third Parties*, is a deliverable in work package 8, which aims to prepare the organizational foundations for the construction and exploitation phase of CLARIN.

The main objective of work package 8 is the preparation of a ready-to-sign agreement between the participating countries whereby they commit themselves to the joint construction and exploitation of the CLARIN Infrastructure. A number of topics have been defined that need to be addressed in the preparation of such an agreement document. This deliverable reports on the first findings regarding current practice in the countries with regard to transnational coordination, cooperation with third parties and openness, and it makes some preliminary statements regarding the future requirements of an infrastructure such as CLARIN w.r.t. the topics addressed.

CLARIN References

CLARIN D8S-1.1. *Requirements and best practice overview for governance*.

Contents

1. Introduction.....	6
1.1 General Background	6
1.2 Structure and contents of this report	6
2. Memberships of Existing Transnational Structures.....	7
2.1 Summary	7
2.2 List of examples of existing structures	7
2.3 Some instances of collaboration activities.....	10
3. Bi- or multilateral Transnational Collaboration Agreements	11
3.1 Summary.....	11
3.2 List of examples of existing bi- or multilateral agreements	11
4. Next Steps – towards an Analysis.....	12

1. Introduction

1.1 General Background

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 organizational 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.

This report on requirements and best practice for transnational coordination and collaboration with third parties is a deliverable in work package 8, which aims to prepare the organizational foundations for the construction and exploitation phase of CLARIN.

The main objective of work package 8 is the preparation of a ready-to-sign agreement between the participating countries whereby they commit themselves to the joint construction and exploitation of the CLARIN Infrastructure. A number of topics have been defined that need to be addressed in the preparation of the agreement document:

1. Governance and management
2. Financial plan
3. Transnational coordination, cooperation with third parties and openness
4. Potential legal issues arising from differences in national legislation.

This deliverable reports on item 3.

1.2 Structure and contents of this report

Data collection for this report was organized through two questionnaires sent to all the CLARIN partners. The objective was to map the funding landscape and the transnational collaboration landscape, respectively, of each country. The report presents an inventory of first findings regarding current practice solutions in transnational coordination and collaboration, and it presents a list of requirements that follow from the way we see the construction and exploitation of CLARIN.

The organizing principle for the grouping of the data is based on the following definition of the types of collaboration, where we distinguish between three types.

1. Participation in the European Commission's Framework Programmes, and all formation of national information offices and support structures that are created in order to promote and support participation in these programmes. None of these, neither the framework programmes nor the national support structures, are included in this report because they are considered to include every CLARIN partner and be basically the same for all.

2. Existing structures that a country may join as a member where membership is voluntary even if a country meets the eligibility criteria, and where the group of countries eligible for membership of the structure is largely either *all European countries* or simply *all countries*. The findings in this category are described in section 2 of this report.
3. Structures that were created as an effort between specific countries who share some common and rather definite interest in that particular co-operation. The findings in this category are described in section 3 of this report.

Section 4 describes future requirements expressed by a few countries, and next steps.

2. Memberships of Existing Transnational Structures

2.1 Summary

The questionnaire reports CLARIN partners' memberships of existing structures requiring transnational coordination. Some basic traits are that CLARIN partner countries are typically members of many such structures and that the list of different structures is very long. The partners have mainly reported memberships of European collaborations – often with associate countries from other parts of the world, but many purely international collaboration projects also form part of the picture.

The national organization responsible for a transnational coordination is usually a ministry of research, a funding agency affiliated to the ministry or a research council, but it may also be e.g. a research and education network, as in the case of GEANT2.

2.2 List of examples of existing structures

Below please find a list of European and international collaborations which have all been mentioned by several CLARIN partners; the list is not exhaustive.

CERN - Centre Européen pour la Recherche Nucléaire (or in English: European Organization for Nuclear Research). CERN organizes and sponsors international co-operation in research, promoting contacts between scientists and interchange with other laboratories and institutes. CERN is run by 20 European countries.

COST - European Cooperation in Science and Technology – is a flexible, fast, effective and efficient tool to network and coordinate nationally funded research activities. COST is based on networks called COST Actions centered around research projects in fields that are of interest to at least five COST countries. 34 European countries are members of COST.

ECMWF - European Centre for Medium-Range Weather Forecasts - is an organization supported by 31 European states and with collaboration agreements with other countries.

EDL - European Digital Library - The EDL foundation's aims include acquisition of access to Europe's cultural and scientific heritage through a portal, maintenance of the portal and support of digitization of Europe's cultural and scientific heritage. The partners cover 100 representatives of heritage and knowledge organisations and IT experts from throughout

Common Language Resources and Technology Infrastructure

Europe. The EDL foundation was established with a background in **Europeana**, a Thematic Network funded by the EC. The Europeana portal is currently being developed and will be launched in 2010 with links to over 10 million digital objects.

EGEE – Enabling Grids for E-science – is the largest multi-disciplinary grid infrastructure in the world and provides researchers in academia and business with access to a production level grid infrastructure. A European project with international associates and comprising more than 50 countries.

EMBO – European Molecular Biology Organization – promotes excellence in the molecular life sciences in Europe. EMBO is mainly funded by EMBC – European Molecular Biology Conference – and together EMBO and EMBC comprise 27 countries including most of the European Union and some of the neighbouring countries.

ERA-NETs - European Research Area. The objective of the ERA-NET Scheme is to step up the cooperation and coordination of research activities carried out at national or regional level in the member states and associated states. Among these nets are:

HERA – Humanities in the European Research Area - is a partnership between 15 Humanities Research Councils across Europe and the European Science Foundation, with the objective of firmly establishing the humanities in the European Research Area and in the 6/7th Framework Programmes.

CIRCLE – Climate Impact Research Coordination for a Larger Europe. Climate impact analysis and adaptation response must be informed by a coherent body of research and it is CIRCLE's prime objective to contribute to such efforts by networking and aligning national research programmes in the 13 CIRCLE European partner countries.

ESA - The European Space Agency - its mission is to shape the development of Europe's space capability and ensure that investment in space continues to deliver benefits to the citizens of Europe and the world. ESA has 18 European member states.

ESF – European Science Foundation - provides a common platform for its member organizations (organizations devoted to scientific research) in order to advance European research and explore new directions for research at the European level. ESF has 80 member organizations from 30 countries.

ESO - European Organisation for Astronomical Research in the Southern Hemisphere - provides state-of-the-art research facilities to European astronomers and astrophysicists and is supported by 14 member states.

ESRF – European Synchrotron Radiation Facility - is an ambitious project comprising 12 member states and representing a very real technological, scientific and human challenge within many research areas. Some common questions that are addressed are: what is our planet made of, how can we explain the properties of matter and what are the processes that sustain life?

Common Language Resources and Technology Infrastructure

EUMETSAT - European Organization for the Exploitation of Meteorological Satellites - delivers weather and climate-related satellite data, images and products. EUMETSAT is an European organization with users from all the world.

EUREKA - aims to enhance European competitiveness through its support to businesses, research centres and universities who carry out pan-European projects to develop innovative products, processes and services. EUREKA has 39 member countries.

EurOcean – European Center for Information on Marine Science and Technology – comprises 12 member states and is a focal point for information on marine science and technology in Europe and its Internet portal is aiming to provide information on topics related to marine science and technology in Europe with a priority given to two main domains: marine research infrastructures and European research, technology and development information.

GBIF – Global Biodiversity Information Facility - is an international network of data providers that builds biodiversity information infrastructure and promotes the growth of biodiversity information content on the Internet by working with partner initiatives and coordinating activities worldwide. GBIF is an international organization with 30 participating countries and continents.

GEANT2 - Consortium of European Academic and Research Networks - GÉANT2 is the successor of GÉANT (ended 2005) and is a multi-gigabit research and education network. It provides state-of-the-art data communications to the European research and education community. GÉANT2 connects 30 European national research and education networks across 34 countries.

GRID - Global Resource Information Database - aims to provide and facilitate access to environmental data and information for decision-making and policy setting, and to underpin UNEP's review of the state of the world's environment and provide early warning on emerging environmental threats. GRID is a worldwide network of 15 environmental data centres managed by UNEP's Division of Early Warning and Assessment.

IARC - The International Agency for Research on Cancer - is part of the World Health Organization and a major goal is the identification of causes of cancer, so that preventive measures may be adopted against them. IARC is an international organization with 5 members.

ITER - International Thermonuclear Experimental Reactor - is a joint international research and development project that aims to demonstrate the scientific and technical feasibility of fusion power. ITER is an international organization comprising American, Asian and European countries.

2.3 Some instances of collaboration activities

In order to get some first ideas about the funding mechanisms and governance mechanisms used in transnational collaborative activities, we provide below a first short description of a few of the activities mentioned above.

ESF – European Science Foundation. As mentioned above, ESF provides a common platform for its member organizations in order to advance European research and explore new directions for research at the European level. ESF has 80 member organizations from 30 countries.

Seat The seat of the ESF is Strasbourg, France.

Members and fees. The member organisations are research councils or academies. The general fee is calculated on the basis of the GNP of the countries. In each country the fee can be split so that several national members share it; this split is decided by the National Group. This fee covers only the central activities and is paid to the ESF secretariat.

Governance. ESF is governed by the Governing Council. Each member organisation has one vote, apart from questions concerning budget and accounts, where the voting is by country according to size of the fee paid for general activities. The chair of the Governing Council is president of the ESF.

The Governing Council employs the Director, who employs the rest of the staff.

Financing and types of activities. The activities can be divided into two types: the general central activities, and “à la carte” activities. The general activities are covered by the fee, and on top of that the countries can opt for financing programmes that they want to support.

Comments. Some countries feel that the ESF is quite administration heavy.

ITER - International Thermonuclear Experimental Reactor. The goal of the ITER fusion program is to produce a net gain of energy, and set the stage for the demonstration fusion power plant to come. ITER has been designed to produce 500 MW of output power for 50 MW of input power - or ten times the amount of energy put in.

Seat St. Paul-lez-Durance, France

Members and fees ITER has seven partners which are countries, the EU being considered as a ‘country’ (The parties are China, EU, India, Japan, Korea, Russia and USA). An international treaty was signed between the parties, setting out the objective and agreement to build ITER and the contribution from each party. The parties contribute very much in-kind i.e. they contribute components of the machine, based on an estimated value of the components. There are, of course, cash contributions from the parties to pay salaries etc.

Staff. In general, the staff are employed by the ITER Organization, which was established after signature of the treaty. Some people are seconded from their parent organisations.

Governance There is an ITER Council, roughly equivalent to the Council of Ministers in the EU. This meets 6 monthly. There is a more hands-on committee that assesses proposals to the Council from ITER management.

Comments

IPR is an issue. There is a tendency for all parties to want a share of the key technologies, so all parties want to participate in everything.

Like for the ESF, some may feel that ITER is quite administration heavy.

To manage the in-kind contributions, each party has established a Domestic Agency. Its job is to agree the specification for components with ITER and then to tender for and supply these. In principle ITER has little access to the end-supplier and no contractual arrangement. This may pose problems.

3. Bi- or multilateral Transnational Collaboration Agreements

3.1 Summary

Data shows that the CLARIN partner countries are engaged in many transnational collaboration agreements. The agreements can be bilateral or multilateral, they can be transnational between European countries alone or they can involve third parties practically all over the world.

The collaborations in general exist between countries with a common language, e.g. Flemish and Portuguese, or with a common geography, e.g. Black Sea, Central Europe, Baltic Sea, or with a common culture, e.g. the Nordic countries, or they can be motivated by a specific purpose such as knowledge transfer, e.g. from old to new European countries. In most cases, several of these factors can be overlapping in the various agreements and programmes presented here.

Like for section 2, the national organization responsible for a transnational coordination is usually a ministry of research, a funding agency affiliated to the ministry or a research council.

3.2 List of examples of existing bi- or multilateral agreements

Below we present the findings from the questionnaires through many best practice examples of current transnational collaborations. Although the list is comprehensive it should not be perceived as exhaustive.

Belgium (Flanders) and the Netherlands have formed the institution Nederlandse Taalunie with the purpose of supporting and promoting the Dutch/Flemish language. One of the instruments used for this purpose is management of research funding, e.g. through the STEVIN research programme which is managed by the NTU. NTU finances TST-Centrale – a Dutch/Flemish HLT Agency.

Belgium (Flanders) and the groups of countries 1) Bulgaria, Estonia, Hungary, Lithuania, Latvia, Poland, Romania, Slovenia, Slovakia, Czech Republic, and 2) Albania, Bosnia and Herzegovina, Kosovo, Croatia, Macedonia, Montenegro, Serbia, and 3) Moldova, Ukraine have entered into a collaborative research program called Co-operation Programme between Flanders and Central and Eastern Europe. The purpose of the program is to consolidate and further establish relationships with the countries of Central and Eastern Europe through project cooperation.

The Nordic Countries Denmark, Finland, Iceland, Norway and Sweden have formed the Nordic Council of Ministers with the purpose of promoting and supporting Nordic collaboration in the areas of education and research in the Nordic countries and autonomous areas. The Council of Ministers have established numerous instruments such as e.g. basic funding given to a number of Nordic Institutions (their web page lists 18 institutions) and to 25 co-operating bodies, programs and other partners. Funding is administered through the institutions, e.g. NordForsk, or through individual

programs, e.g. NordPlus. The Nordic Council of Ministers is also responsible for other international collaborations, which may or may not include research. The web site lists the following countries and regions under International Co-operation: Russia, Belarus, Estonia, Latvia, Lithuania, Baltic Sea Region, Poland, Germany, West Nordic Region (Faroe Islands, Iceland and Greenland), and the Arctic.

In the Czech Republic, the Czech Science Foundation has as one of its purposes to cooperate and support international scientific cooperation on research projects through agreements with research councils all over the world. A closer cooperation with the German Research Foundation (DFG) and Korea Research Foundation (KRF) is mentioned, and from Germany it is confirmed that the DFG is a partner of the Czech Science Foundation, and in addition we are told there is a special DFG program for scientific co-operation with the Middle- and Eastern European countries. The Italian CNR (Consorzio Nazionale de Recherche) has similar tasks in its mission.

Greece participates in the SEE-ERANET programme; other participating countries are Albania, Austria, Bosnia and Herzegovina, Bulgaria, Croatia, France, Germany, Greece, Hungary, Former Yugoslav Republic of Macedonia, Montenegro, Romania, Serbia and Slovenia.

The Hungarian Scientific Research Fund collaborates with the American National Science Foundation about health and welfare issues and they have a broad scientific collaboration with the Deutsche Forschungsgemeinschaft (DFG) in Germany.

Portugal has bilateral scientific agreements with countries in Africa, America, Asia and Europe (France and Italy), and in the context of CLARIN the collaboration agreements with Portuguese speaking countries in America (Brazil), in Africa (Cabo Verde, Guiné-Bissau, S. Tomé e Príncipe, Angola, and Mozambique), and in Asia (Timor Lorosae) are of particular relevance.

Romania collaborates with Greece, Adzerbaijan, Armenia, Bulgaria, France, Germany, Georgia, Italy, Malta, Moldova, Turkey, and Ukraine in the Black Sea ERA-NET collaboration.

4. Next Steps – towards an Analysis

In this report we have listed the various types of transnational collaboration, apart from the European Commission Framework Programmes. We have also mentioned e.g. ERA-NET, although they can be seen as part of the European Framework Programmes. This is because they constitute a special type of collaboration within the FP, and because they specifically focus on the collaboration between ministries or other governmental funding agencies.

Some respondents have raised problem areas which may deserve some consideration, such as local national problems arising from the general economic situation and from lack of coordination and long term strategies. Although these problems are reported as national, they do of course also have an effect on the ability to enter into and to carry out international collaborations and agreements.

In the questionnaire we distributed we also mentioned problems and requirements that could originate from topics such as e.g.

- coordination with national and EU R&D programmes and third parties,
- problems arising from different economic conditions in different countries,
- problems arising from lack of synchronisation in national budgetary procedures,

Common Language Resources and Technology Infrastructure

- problems arising from differences in R&D funding allocation procedures in different countries,
- problems arising from lack of openness.

Respondents confirmed that these were relevant, but only highlighted the problem of the economic situation and the lack of long term strategies.

Future work in this area consists of mainly two activities:

First, we will analyse in greater depth, selected existing transnational collaboration, as far as possible in collaboration with research agencies, in order to determine the important factors for successful collaboration. This is a deeper and more comprehensive analysis of what was mentioned above.

One of the activities we want to analyze more in-depth is GEANT2, because GEANT2 has many features in common with CLARIN, except that they are more physical in nature than CLARIN is. A visit to the GEANT2 management team is being planned. Additionally, we will analyse three to four other projects/activities of a similar and/or different nature in order to cover the spectrum as well as possible. We may work further with some of the above, and may choose others.

Secondly, these findings will have to be compared with and integrated with the work which is ongoing in the other part of this work package: *Analysis and proposals for governance*, and with the ongoing work with the European Commission and other research infrastructure projects on the European Research Infrastructure (ERI, ERIC).

Finally, any additional input which may be provoked by this report, will of course be taken into account.

Issues: We welcome input about which transnational activities to analyze in-depth during the next months.