Title	Structuring the KSI
Version	2
Author(s)	SK
Date	12-09-2017
Status	For discussion
Distribution	KSIC & UI
ID	CE-2017-1045

Table of Contents

Preface

This is an internal document, still under construction, prepared for the joint UI-KSI Workshop at the Annual CLARIN Conference 2017 in Budapest. Some bits and pieces are still missing, such as a list of workshops organised at the central and decentral level. The User perspective should be better developed.

1   Introduction

Distributed data research infrastructures have three main components:
- At the heart is the Technical Infrastructure, consisting of the technical facilities that provide users access to data and tools, and the people that operate them.
- A set of commonly agreed upon organizational rules and measures and conventions that should ensure seamless interaction between infrastructure users, operators and components, including e.g. standards, access, licenses, quality assurance etc. I will call this the Policies Infrastructure.
- A set of measures and facilities aimed at securing a continuous transfer of knowledge between the main players involved in the construction, operation and use of the infrastructure. We call this the Knowledge Sharing Infrastructure.

In this document, we focus on the Knowledge Sharing Infrastructure (abbreviated KSI). Throughout the document it will become clear that there is no sharp distinction between knowledge sharing and user involvement, as the ability to make use of the knowledge and expertise accumulated by the CLARIN community at large is a necessary condition to optimally benefit as a user from what CLARIN has to offer.

After a brief introduction in section 1 we give in section 2 a brief description of what the KSI is and should do for us. In section 3 we discuss the basic organisational building bricks of the KSI: the K(knowledge)-centres. In section 4 we briefly describe the various types of members of the CLARIN community and their roles. In section 5 we present the various instruments that we can use in addition to the K-centres. I section 6 we make some concluding remarks. Two annexes show examples of themes and instruments that could be used as sources of inspiration when further developing the KSI.

2   The Knowledge Sharing Infrastructure

In CLARIN we see at least six distinct roles that people or organisations can play in relation to the infrastructure. For each of the roles we will use the term in brackets to refer to it:
1. Those who use, want to use or should be encouraged to use the infrastructure (Researchers)
2. Those who educate or help others to use the infrastructure (Teachers)
3. Those who populate it with data and services (Providers)
4. Those who build and operate the technical infrastructure (Operators)
5. Those who instruct others how to build and operate the infrastructure (Instructors)
6. Those who manage the infrastructure (Managers)

We will use the term CLARIN Community to refer to every person or organisation that plays one or more of the above roles, none of which are mutually exclusive. Most members of the CLARIN community (both institutions and individuals) will generally play more than one of the above roles, depending on the situation: someone who uses the infrastructure as a researcher can at the same time teach classes in Digital Humanities, create new data and tools and deposit them at the infrastructure, participate in the construction of parts of the infrastructure, instruct colleagues in

other countries to build up the infrastructure, and be at the same time in charge of an infrastructure or one of its subcomponents. In the next section we will elaborate a bit more on this.

To carry out the tasks that follow from their roles people need knowledge and expertise to do their jobs, and they generate new knowledge and expertise as they go along. Contrary to big infrastructures in the form of physical installations, where the knowledge and expertise can be brought together in one physical location known to all, and can easily be shared within and across the various groups, a distributed infrastructure such as CLARIN exists in dozens or maybe even hundreds of locations, distributed all over Europe, and the same applies to specific CLARIN knowledge and expertise.

The KSI will be our instrument to ensure that such knowledge and expertise will not exist as a fragmented collection of unconnected bits and pieces, but will be made accessible in an organized way to the whole CLARIN community and to the Social Sciences and Humanities [SSH] research community at large.

3   The people who play a role in the CLARIN community

Researchers are typically (but not necessarily) academic employees or advanced students who have as one of their main tasks to do research within a specific discipline within the Humanities or Social Sciences, or in another discipline where language plays a role. Researchers can also be citizen scientists, journalists, policy makers or working in archives, libraries or museums, who have to conduct research in order to do their work. They can benefit from using CLARIN by using its data or tools, its depositing facilities or the expertise present in the community

Teachers (or facilitators) are people from the domain of humanities with communication abilities to connect technical work with the non-specialised user world, who engages researchers in the CLARIN community (and beyond) in the activity of Digital Humanities and the use of NLP tools properly and provides a guidance to find and access the knowledge sources that users need for their work. They also exchange information through mass media to the scientific community, via internet and other mediums for disseminating information, news and advances in DH and NLP tools. They will exploit potential synergies and complementaries of different groups.
They help a research group understand their needs and objectives, and they assists researchers to plan how to solve these needs and achieve the objectives.
They will serve researchers who are transitioning into Digital Humanities and use of NLP tools in research.
The teachers organize problems or topics of study and shows how other researchers obtained results and use NLP tools and promote discussion addressing the needs of the research team.

A **provider** can be a researcher, a group, a library or an institution that provides language data (e.g. corpora), software or services accessing and using the Internet in an open way to CLARIN community. This service has to be available for any CLARIN user who needs it.

Every infrastructure needs a group of people that **build, operate and maintain** the infrastructure. These people can be technical experts, or scientists. In a multi-national and multi-sited infrastructure such as CLARIN, where not every country or site is in a position to act at the same pace as the others there is a natural need for groups that are already well-established to share their expertise with those who are less advanced due to e.g. late start, absence of specific expertise at the national level or less favourable economic conditions.

Infrastructures and their components cannot function without proper **management**, both at the transnational, the national and site level. This comprises both the management task proper, as well as the support functions such as technical, financial, administrative, legal and HR management.

4 Main building blocks: K(nowledge)-centres

4.1 What it is

Parallel to what we have for the technical infrastructure the main building blocks of the KSI will be knowledge centres, which can be physical (i.e. concentrated in one physical place) or distributed (spread over various institutions or centres in one or more countries, with a common focus).

Their role is gather and maintain a specific body of knowledge and expertise, and make it available to and share it with the CLARIN Community at large.

Knowledge centres may target specific groups (e.g. the six listed in the first paragraph above), or cross- or sub-sections. They can be based on language ("the K-centre for the Danish language"), modality ("the K-centre for speech analysis"), domain ("the K-centre for Dutch historical texts"), technology ("the virtual K-centre for machine learning"), data types ("the K-centre for treebanking), problems of common interest ("the K-centre for OCR"), etc.

A typical characteristic of K-centres is that they have to provide their services to the CLARIN community at large, which normally implies that services should be provided at least in English and that they should have a high degree of availability. Operating a K-centre is a typical contribution in kind to the CLARIN infrastructure.

Not every CLARIN site will be in a position to do this, not only because of the potential additional workload, but also because of the language. Apart from e.g. in linguistics or computational linguistics, where English seems to be the normal vehicle for communication, local languages may dominate in many subareas of the humanities and social sciences, and it would not always be obvious how (and why) to open up a course in e.g. processing of mediaeval Polish poetry, taught in Polish, to people who don't speak the Polish language.

Sustainability and critical mass are other important factors for K-services: activities that rely on the expertise of a single expert may not be as sustainable as those that are supported by a whole team of experts. Sustainability is also linked to the user base: if there is no demand for the expertise offered by a (candidate) K-centre from their envisaged target audience there is little point in sustaining it.

K-centres can also be contributed by countries or regions not covered by CLARIN ERIC, or by individual institutions.

Institutions or groups of institutions from inside or outside CLARIN can submit an application to be recognized as a K-centre. The requirements and the whole process are described in detail in document CE-2015-0504. Applications are received by the KSI committee, assessed on the basis of the criteria as an iterative process, and when the committee is satisfied with the application in its final form it is submitted to the Board of Directors for final approval.

Recognised K-centres receive a certificate which is valid for 3 years, and the description of the centre is included in the list of centres on the CLARIN website.

Not all knowledge sharing activities need to be implemented through K-centres. In section 4 we will discuss a number of other instruments that are already being used or that can be used to share knowledge both within the CLARIN community and with other interested parties.

4.2 What we have

The following table lists the first nine K-centres. Three more applications are in the pipeline.

Centre	Institution(s)	Area of competence	Link to Certificate
Spanish CLARIN K-Centre	(1) CLARIN Competence Centre IULA-UPF and HDLab @ Department of Humanities (Universitat Pompeu Fabra, Barcelona) (2) UNED – LINHD: Laboratorio de innovación en Humanidades Digitales (Universidad Nacional de Educación a Distancia, Madrid) (3) UPV – Grupo IXA (University of the Basque Country, San Sebastián)	Spanish texts (all), Basque texts (IXA) Catalan texts (IULA-UPF)	certificate (PID)
CLARIN Knowledge Centre for Treebanking	(1) University of Bergen (UiB) (2) LINDAT @ Charles University in Prague (CUNI)	Treebanks: construction, integration, search, processing, formats	certificate (PID)
Phonogrammarchiv CLARIN K-Centre	Phonogrammarchiv – Institute for audio-visual Research and Documentation (Austrian Academy of Science)	Audio-visual fieldwork and long-time preservation of audio-visual data	certificate (PID)
CLARIN Knowledge Centre for Speech Analysis (CLARIN-SPEECH)	KTH Royal Institute of Technology, Dept. of Speech Music and Hearing	Technical advice on speech analysis	certificate (PID)
CLARIN K-Centre DANSK - DANish helpdeSK	(1) The Center for Language Technology (CST) @ Department of Nordic Research (University of Copenhagen) (2) The Department of Nordic Studies and Linguistics (DNSL) (University of Copenhagen) (3) The Danish Language Council (DSN)	Danish and language technology for Danish	certificate (PID)
The CLARIN Knowledge Centre for Language Learning Analysis (CLARIN-Learn) (TalkBank)	The TalkBank Project @ Carnegie Mellon University (USA)	Technical advice on tools and methods for the study of first and second language learning, and recovery from aphasia and other language disabilities	certificate (PID)

Centre	Institution(s)	Area of competence	Link to Certificate
CLARIN Knowledge Centre for Languages of Sweden *(via https://sweclarin.se/)*	The Language Council of Sweden (Stockholm), in cooperation with the other sections of the Institute of Language and folklore in Uppsala and Göteborg.	Information service offering advice on the use of digital language resources and tools for the Swedish language, minority languages in Sweden, the Swedish sign language, Swedish dialects, as well as other parts of the intangible cultural heritage of Sweden in text and speech, as well as language policy and planning	certificate (PID)
The CLARIN Knowledge Centre of Lund University Humanities Lab	Lund University Humanities Lab at Lund University	Advice on multimodal and sensor-based methods, including EEG, eye-tracking, articulography, virtual reality, motion capture, av-recording	certificate (PID)
CLARIN Knowledge-Centre,for linguistic diversity and language documentation (CKLD) [no website available yet]	University of Cologne, Hamburg Centre for Language Corpora (HZSK), Endangered Languages Archive (ELAR) & SOAS World Languages Institute, Leibniz Center General Linguistics (ZAS)	Expertise on methods, technology and background information on language resources and tools in connection with linguistic diversity research and language documentation: – Answering of incoming questions on these topics – Online-documentation and web-pages on relevant tools and best practice methods – Training in workshops and intensive courses, both as regularly offered general trainings and as (project) specific workshops on request	Not yet available

A number of applications is still in the pipeline.

4.3 Things to do

At this moment we don't know yet how many other sites in CLARIN ERIC as a whole are in a position to act as K-centres, and what types and volumes of knowledge sharing activities they offer. This may differ from country to country, depending on their funding conditions and their ambitions.

Thus far the process of the creation of K-centres has been a bottom-up process. As we don't have the funds to create our own centres we are crucially dependent on what the institutions in the CLARIN countries (and in some cases outside CLARIN) are able and willing to offer. But rather than just wait and see what happens we can formulate general guidelines (e.g. in the form of a wish list or recommendations) and encourage institutions that are sitting on specific expertise to apply (individually or jointly) for recognition as a K-centre if we feel that there is a demand for it.

Some examples of areas of expertise that could be candidates to be covered by K-centres:
- K-centres for each national or regional language spoken in any of the CLARIN member countries
- Coverage of all modalities (e.g. building, annotating, exploring and exploiting corpora)
- Coverage of main linguistic data types (e.g. corpora, treebanks, lexical data, manuscripts)
- Coverage of main supporting technologies (e.g. OCR, machine learning, text mining, visualisation)
- Coverage of specific DH (sub-)disciplines (e.g. oral history, newspapers, parliamentary records, social media)
- How to deal with IPR
- How to make a Data Management Plan
- How to teach Digital Humanities

and many others.

It should be noted that it only makes sense to create K-centres if there is evidence that there is sufficient demand for their services or that it relates to outreach activities that are aimed at creating such a demand.

At this moment we don't have guidelines or templates for periodic reporting by K-centres. This should be taken up, as it is necessary for the further development of the K-centre concept to see how they function (or don't function) in practice. It is also of crucial importance to hear the voices of those who have made use of the services of K-centres. Such insights are useful both for the centres and for the KSIC to see what works and what doesn't work, which may help us to provide better criteria and best practice guidance to existing and future K-centres.

5 Knowledge sharing topics and instruments

Many organisations in CLARIN will have activities that could be described as 'knowledge sharing' even if they don't have the ambition or capability to be recognised as K- or L-centre. Even without official status these activities are part of the KSI and can all be seen as in kind contributions to CLARIN. All players in the KSI may use specific instruments to make their knowledge and expertise available, both proactively (e.g. organizing courses, workshops, web courses, social media, mobility schemes) and reactively (e.g. help desks, FAQs, technical support on demand, portals).

The challenge for the CLARIN organisation is to ensure that as much of this knowledge and expertise as possible is made accessible to the CLARIN community at large.

An analysis of the CLARIN Agreements made in 2013 on the basis of the specific competences listed by the members of the various national consortia suggests quite a number of possible topics. It is

attached as Annex 1 as potential candidates for new knowledge centres or as topics to be taken up as individual knowledge sharing activities. In this document we will not try to assign priorities to topics and themes, although at some stage we may want to do this in order to be able to give our activities more focus – but this should always be driven by a combination of the needs we perceive and the practical opportunities we see.

The 2013 analysis of the CLARIN Agreements and documents from the CLARIN Preparatory Phase (2008-2011) shows a number of instruments. The list is attached as Annex 2 as a possible source of inspiration.

Many of the instruments actively deployed at this moment are implemented and coordinated at the institution or consortium level. The role of CLARIN ERIC is necessarily limited although as part of its responsibility to ensure the added European value of the CLARIN infrastructure as a single, distributed infrastructure as opposed to a loose collection of individual infrastructures, it could undertake a number of activities in connection with what is happening in the national consortia, such as:
- Make a continuously updated inventory of knowledge sharing activities across national consortia and publish them through our website
- Suggestions to national consortia to open up specific instruments to 'external' people, i.e. people from outside their own consortium
- Suggestions to national consortia to team up in cases where this could lead to more effective, more efficient or more impactful activities
- Suggestions to initiate activities that would be beneficial for the user communities and that would be best placed at the national level
- Facilitate exchange of information between national consortia

At the central CLARIN level KSI activities should be complementary to what is happening in the countries, and that means that the focus should be on strengthening their impact (along the lines sketched above), integration, innovation and outreach.

Below we will briefly discuss the main instruments we use for Knowledge Sharing in addition to the K-centres

5.1   Knowledge sharing instruments

5.1.1   What we have
The annual conference
The CLARIN Annual Conference is organised for the Humanities and Social Sciences community in order to exchange ideas and experiences on the CLARIN infrastructure. This includes its design, construction and operation, the data and services that it contains or should contain, its actual use by researchers, its relation to other infrastructures and projects, and the CLARIN Knowledge Sharing Infrastructure. Although people from outside the CLARIN community can submit papers and attend the conference it is mostly an internal event, to which normally also representatives are invited from countries where there is an interest in joining CLARIN.
The format is based on submitted papers related to CLARIN, sessions where various bodies report on their activities and plans for the future, and pre- or appended meetings of committees. The NCF acts as the Programme Committee.

Its main objective is strengthening integration and cohesion across national consortia, and outreach to new countries. In the longer run we hope that it will help to attract more people and to make

ourselves more visible to the outside world, but it is not a priori evident how a broader event should be positioned in relation to conferences such as Digital Humanities and LREC and other conferences

The annual conference is centrally organised and funded, with contributions (financial and in kind) by the national consortia.

Workshops [lists not complete yet]
Workshops as we use them are typically gatherings dedicated to addressing a specific problem area or training. In some cases, typically if organised in the margin of major conferences, they take the shape of focused mini-conferences
At this moment 4 types of workshops have been or are being organised and funded centrally:
   A. Workshop type I: addressing a topic that is in line with one or more of the strategic priorities of CLARIN. Examples are:
      o
   B. Workshop type II: meant to prepare a work plan for a development project of up to three person months, again contributing to one or more of the strategic priorities of CLARIN. Examples are:
      o
   C. A series of 4 training workshops, covering different topic areas, organized for those who are in charge of building and operating CLARIN consortia and centres. The 4 topic areas are:
      o Organisational issues
      o Tools and services
      o Repositories
      o How to reach out to new users
   D. In order to allow for a focused approach in bringing in new users four topics have been identified that can be used to demonstrate how the application of language and speech technology (LST) tools and services on digital language material can advance humanities and social sciences research in other fields than linguistics. The 4 topics are:
      o Work on digital collections of newspapers (as a dataset highly relevant to many SSH communities, without major IPR issues for historical newspapers)
      o Oral history archive materials (a topic which combines language and speech technology, and is hard to do without digital tools)
      o Parliamentary records (from national assemblies and the European Parliament, and which are normally freely accessible and available for re-use)
      o Analysis of social media

Out of these four types A and B are funded from the central budget through calls for proposals, and C and D by the CLARIN-PLUS project.

Series A and B contribute to cross border integration (at least 3 countries per proposal) and to strengthening of the infrastructure.
Series C addresses outreach to new countries and integration at the infrastructure level. D addresses outreach to new users and better integration of users on the one hand and providers on the other.

On a number of occasions CLARIN has (co-)organised workshops in the margin of major conferences. Examples are:
o

Some national consortia organise workshops that are open to members of the CLARIN community. Examples are:
o

5.1.2 Things to do

Special workshops can be organised if and when the need arises, e.g. to prepare major infrastructure or research project proposals. National consortia should be encouraged to publish announcements of their workshops on the CLARIN website if they are relevant for the CLARIN community.

Given our limited financial and human resources we should not aim at organising more than one conference per year, but we should not exclude the possibility to participate in joint conferences (or smaller events) together with other RIs, as we have done in the past (in 2010 in Vienna and in 2011 in Copenhagen, together with DARIAH).

5.2 Mobility grants

5.2.1 What we have

Mobility grants are grants that allow people to spend time in another place to do research, give instruction or be instructed in the use, construction or operation of the CLARIN infrastructure.

In order to enhance integration within CLARIN at the human resources level, and thus to reduce the risk of creating a two-speed CLARIN, a Mobility Grant scheme was set up in 2013. The aim is to allow researchers and technical staff working at centres in CLARIN ERIC member countries (or organisations) to pay short visits (typically up to 1 week) to CLARIN sites in other countries in order to do one or more of the following:

(A): "Horizontal mobility", between CLARIN sites in different countries:
   (a) For technical and scientific staff from CLARIN sites to receive training in all aspects related to setting up or operating centres and/or services, or
   (b) For experts from CLARIN centres to provide in situ training in these areas to local teams;
   (c) For technical and scientific staff from CLARIN sites to collaborate with technical or scientific staff from other sites on the solution of specific problems related to the construction or operation of CLARIN
   (d) For technical and scientific staff to collaborate with technical or scientific staff from other sites on porting solutions between sites or languages

(B): "Vertical mobility", between CLARIN sites on the one hand, and humanities or social sciences research institutes (not necessarily CLARIN sites) in other CLARIN ERIC countries on the other:
   (e) For scholars to receive training in using the infrastructure to conduct research in the humanities or social sciences, or
   (f) For experts from CLARIN sites to provide in situ training in these areas to local humanities or social sciences research teams
   (g) For scholars to collaborate with technical or scientific staff from a CLARIN site on using the CLARIN infrastructure as an instrument in a specific humanities or social sciences research project.

CLARIN sites can be both infrastructure centres, data centres, service centres and knowledge centres in CLARIN ERIC member or observer countries; sites can also be academic departments or research institutes in Europe with a focus on the humanities or social sciences, or repositories, computing facilities, or other technical facilities in Europe.

More details can be found in document https://office.clarin.eu/v/CE-2013-0151-MobilityGrants-CurrentCall.pdf and the application form https://office.clarin.eu/v/CE-2013-0117-MobilityGrants-Application-Current.docx

5.2.2   Things to do

Mobility grants are currently under-used, and we should promote them more, e.g. on our website. At this moment mobility grants can only be found by those who know that in our jargon we call them "mobility grants".

Within the constraints of our budget there are no possibilities to support longer stays, e.g. to support collaboration on a research project between a humanities scholar and a technical person for a longer period. We may want to look into the possibility to attract EC funds to support this type of training (and maybe other types as well).

In the past some national consortia had funding in their budget to host researchers from other places, but it is not certain that this still exists. We should check with the national consortia whether this is still the case and if so encourage them to publicize these opportunities more widely.

5.3   Mass dissemination

5.3.1   What we have
- The website. This our main continuously accessible instrument for knowledge and information sharing. We may want to ask ourselves whether all 5 roles that visitors can have (see section 2) are equally well covered.
- The Newsflash. This appears to be a good instrument to disseminate information within the CLARIN community, rather than knowledge.
- Social media. Not sure what we have at the moment. Twitter, Facebook, Forums, Blogs, Video channels?

5.3.2   Things to do
The User Involvement Action Plan for 2017 (https://office.clarin.eu/v/CE-2017-0949-UI-Action_plan_2017_NCF.pdf) seems to cover most of these aspects very adequately, and there is no need to repeat them here.

5.4   Document and software libraries
By libraries we mean structured catalogues of (references to) documents or software collections. Examples are:
- Best practice documents and how-to guides, e.g.
    o   Which standard should I adopt for my project
    o   Which licence should I select for my data
    o   How do I refer in my publications to data or tools accessed via CLARIN
    o   How do I make a data management plan
    o   How can I use named entity recognisers in my history project
    o   What is the best approach to OCR
    o   How do I create a parallel corpus from the web
- Descriptions of successful use cases as examples of what one can do with CLARIN
- Scientific papers related to CLARIN
- Collections of software tools

5.4.1   What we have
It is hard to tell from what one sees on the front page of the website what we have in each of the 4 example categories above. CLARIN-PLUS has produced best practice guides in the context of WP5. They are available on the website, in the form of project deliverables, and address a small subset of our audience (Managers and Operators, not Researchers).

5.4.2   Things to do

Decide which ones we want to have, collect sufficient items to give them critical mass and then publish them on the website in a place where people can find it and in a form that appeals to the target audience.

5.5   Education and training

The User Involvement Action Plan for 2017 (https://office.clarin.eu/v/CE-2017-0949-UI-Action_plan_2017_NCF.pdf) presents the planned actions for 2017. Even if the scope of KSI is a bit broader in some aspects (as it also targets 'internal' people, such as infrastructure managers and operators), but narrower in others (outreach as such is not the focus of the KSI) the UI plan covers a broad spectrum of KSI activities and instruments and there is no need to repeat it here.

We have to make sure that the interests of others than consumers of our services are also properly covered. We see opportunities for this in close collaboration with sister infrastructures and infrastructure projects, such as DARIAH and PARTHENOS.

6   Concluding remarks

A data research infrastructure such as CLARIN cannot exist as just a constellation of technical facilities but requires a closely connected knowledge infrastructure that ensures a proper flow of knowledge and expertise between all categories participants, so that new research can build on work done by others. In a single-site physical infrastructure this is relatively easy to accomplish, and people will know where to find the information, and if necessary the training they need to optimally benefit from the infrastructure.

In a widely distributed data infrastructure, where users have access to the infrastructure from all parts of Europe, across national and language barriers a special effort is required in order to ensure access to information and expertise needed to use the infrastructure. In the case of CLARIN we can make use of the fact that many CLARIN centres and sites are based in universities, where there is a strong tradition of knowledge transfer. The challenge for CLARIN is to make optimal use of the knowledge and the knowledge transfer mechanisms that already exist, and to try to make sure that a much broader community can benefit from them, so that all thresholds for using the infrastructure in order to enhance and innovate European research in the humanities and social sciences will disappear.

In all our activities it is important to keep in mind that the KSI targets 6 different audiences in terms of roles: Researchers, Teachers, Providers, Operators, Instructors and Managers, and that they have different needs and priorities depending on in which role they are in contact with CLARIN. We should also keep in mind that there will be a large overlap between the KSI activities for Users and our User Involvement activities, as is clearly illustrated by sections 4.3.2 and 4.5.

ANNEX 1: Instruments

From the CLARIN Agreements and earlier CLARIN documents we have extracted a whole battery of possible instruments for knowledge sharing.

Below we list the types encountered, in alphabetic order.

instruments

automated helpdesk	mailing lists
awareness seminars	manuals
background documents	movies
best practice documents	newsletters
books	panel sessions
conferences	personal advice
consultation	physical centres of expertise
cooperation with associations etc	portal
course modules	questionnaires
curation service	scientific papers
curriculum development	short guides
customisation of resources	showcases
directory of expertise	social media
directory of players	summer schools
documentation	training
ec projects (e.g. marie curie)	tutorials
education	virtual centres of expertise
faq	web presentation of knowledge
forums	web courses and lectures
grants	websites
help desk	wiki
hosting	workshops
interviews	
journals	
ksi network	

ANNEX 2: Topics

In the CAs we found a large number of different topics on which expertise is offered, using various instruments, and for a variety of languages. We list them here in alphabetic order.

topics

aai
annotation tools
archiving
audio data
chunking
cmdi
collecting data from web
concept registries
conversion
corpora in education
crowdsourcing
curation
digital humanities
extraction of lexical semantic networks
field work
formal semantics
gis
human-human comms
human-machine comms
individual languages
information extraction
information retrieval
ipr
isocat
knowledge management
language documentation
large corpora
legal issues
lexicography
licenses
linking resources
linking web services
localisation to own language
longitudinal studies

machine learning
metadata infrastructure
multilingual corpora
multimodal
multi-word expressions
named entity recognition
nlp analysis tools
ontologies
orthography
parsing
pid
porting between languages
registers
repositories
segmentation
semantic networks
shallow parsing
speech
standards
statistics
tei integrator
text
text mining
topic extraction
translation
treebanks
validation
video processing
visualization
web services
wordnets
workflows
wsd