
April 2013
Deliverable: R1.2:
Second Year Activity Report –
Progress Report of the Project and the Various Work Packages

Responsible: Erhard Hinrichs

© All rights reserved by the University of Tübingen on behalf of CLARIN-D

Editors: Scott Martens, Kathrin Beck, Erhard Hinrichs
Contributors: Volker Boehlke, Christoph Draxler, Axel Herold, Hannah Kermes, Erik Ketzen, Scott Martens, Timm Lehmberg, Dieter van Uytvanck, Heike Zinsmeister
Introduction ........................................................................................................................................ 4
Work Package 1: Management ........................................................................................................ 6
Work Package 2: ERIC / DARIAH Liaison ....................................................................................... 8
Work Package 3: Technical Infrastructure ....................................................................................... 10
Work Package 4: Discipline-specific Working Groups ................................................................. 11
Work Package 5: Services and Resources ..................................................................................... 12
Work Package 6: Legal and Ethical Issues ...................................................................................... 13
Work Package 7: Support and Help Desk ..................................................................................... 15
Work Package 8: Training and Education ..................................................................................... 16
Work Package 9: Dissemination and Acceptance ......................................................................... 18
Introduction

CLARIN-D – Web- und zentrenbasierte Infrastruktur für die Geistes- und Sozialwissenschaften is the follow-up project of D-SPIN (Deutsche Sprachressourcen-Infrastruktur).

In the three years of the preparatory phase (2008-2011), D-SPIN was the German partner of the European CLARIN project. The German Federal Ministry of Education and Research (BMBF) funded D-SPIN, while the European Commission funded CLARIN directly. The preparatory phase developed plans for digital humanities infrastructure through technical research, prototypes, models for training researchers, and through clarification of the legal framework for the project.

The implementation project phase began in 2011 and will continue through 2016. On the European level, the CLARIN European Research Infrastructure Consortium (ERIC)\(^1\) has replaced the multinational CLARIN project, and CLARIN-D is the national successor project of D-SPIN.

CLARIN\(^2\) is part of the European Seventh Framework Programme (FP7)\(^3\) as well as the ESFRI Roadmap\(^4\) (European Strategy Forum on Research Infrastructures), which provide for detailed planning on the national and international levels in order to construct competitive, open-access European research infrastructures to better serve research communities.

ESFRI Research Infrastructures are built in three consecutive phases – the preparatory phase, the implementation phase, and the exploitation phase.

In this phase, funding is provided entirely by the participating European countries directly to the national partners. CLARIN-D is funded directly by the German Federal Ministry of Education and Research (BMBF), but also receives contributions from German state governments (Länder) and individual universities, as well as support from the Berlin Brandenburg Academy of Science, the Leibniz Foundation, and the Max-Planck-Society.

The CLARIN report R1.2 “Second Year Activity Report – Progress Report of the Project and the Various Work Packages” provides a short overview of the activities carried out in the CLARIN-D project during the second year of the implementation phase.

\(^1\) [www.clarin.eu](http://www.clarin.eu)

\(^2\) [http://clarin-d.de](http://clarin-d.de)


According to the schedule of deliverables, work packages 2 – 9 have produced more detailed reports about specific aspects of their work. These deliverables include:

- R2.2a/b Liaison with the CLARIN ERIC and Liaison with DARIAH-DE
- R3.2 Establishing the technical infrastructure
- R4.2 Annual report on the activities of the discipline-specific Working Groups
- R5.2 AP5 Second Year Activity Report
- R6.2 Second Year Activity Report
- R7.2 Support and Helpdesk
- R8.2 Documentation of AP8 preparation activities
- R9.2 Dissemination

All deliverables are available online on the CLARIN-D web page: http://clarin-d.de/en/documents.html
Work Package 1: Management

Work package 1 is devoted to the overall management of the project.

CLARIN Agreement
CLARIN ERIC and CLARIN-D have created and signed a CLARIN Agreement, coordinated by Tübingen on behalf of the whole project. Among technical details concerning the project as a whole, each CLARIN-D member listed the language resources and tools that they are contributing to CLARIN ERIC.

Administrative Support
Tübingen has coordinated all project reporting and provided support to CLARIN-D partners and external parties concerning various aspects of the project, particularly administration and legal issues. The financial and legal office for CLARIN-D curation projects was jointly administrated by Sebastian Drude (MPI Nijmegen) as administrator of the curation funds, Kathrin Beck (Tübingen) as project coordinator, and Volker Boehlke (Leipzig) as administrator of the Discipline-Specific Working Groups (F-AGs).

Internal Steering Committee
In addition to the consortium of nine project partners, a project-internal steering committee was elected in the first project year to deal with operational questions like curation projects (cf. WP4). They have monthly virtual meetings.

Members are elected for two years. The steering committee consists of three project partners, one member of the F-AGs, ex officio the project manager, as well as the administrative coordinator and the two technical coordinators. Due to the retirement of Peter Wittenburg, leader of the CLARIN-D centre MPI Nijmegen, his successor Sebastian Drude took over his seat in the CLARIN-D Steering Committee in summer 2012. The current members of the Steering Committee are Ludwig Eichinger, Thomas Gloning (F-AG), Gerhard Heyer, Sebastian Drude, Erhard Hinrichs, Kathrin Beck, Dieter van Uytvanck, and Thomas Zastrow.

Scientific Advisory Board
A Scientific Advisory Board for the project has been established in the first project year. According to the project work plan, the mission and tasks of the board are to control and to safeguard the progress of the project, to remain informed about the status of the project by receiving regular reports, to advise the project accordingly, and to facilitate contact and cooperation with other projects. The second meeting of the Scientific Advisory Board will take place following the M24-workshop in Nijmegen at the end of the second project year.

The current members of the Scientific Advisory Board are Helen Aristar-Dry (Eastern Michigan University, Linguist List), Christiane Fellbaum (Princeton University), Björn Granström (KTH, Royal Institute of Technology Stockholm), Helge Kahler (BMBF, Bonn), John Nerbonne (University of Groningen), Jan Christoph Meister (University of Hamburg), and Achim Streit (KIT, Karlsruhe Institute of Technology)

Joint Technical Advisory Board (TAB) of CLARIN-D and DARIAH-DE
CLARIN-D and DARIAH-DE established a joint Technical Advisory Board consisting of the following national and international experts on computer and language infrastructure:
Jonas Beskov (University of Stockholm), Carol Goble (University of Manchester), Jan Hajic (Head of the Prague CLARIN Centre), Ed Hovy (University of Southern California), Michael Lautenschlager (German Research Centre for Geosciences, Potsdam), Gerhard Schneider (University of Freiburg), Toma Tasovac (Digital Humanities Centre, Belgrade), Melissa Terras (University College London), Claire Warwick (University College London)

Details on its first meeting can be found in Chapter 2.

**Project Meetings**

In addition to the monthly meetings of the internal steering committee, the consortium of all centre leaders meets every three months. The following project-internal meetings were organised jointly by Tübingen as coordinator and Stuttgart as representative of the CLARIN-D members:

- A virtual consortium meeting, 28 June 2012  
- A consortium meeting in Stuttgart, 25 September 2012  
- A consortium meeting in Hamburg, 26 November 2012  
- A consortium meeting in Tübingen, 05-06 March 2013

**M24 Organisation Committee**

At the CLARIN-D consortium meeting in November 2012, a number of task forces and an M24 organisation committee were established. The M24 organization is coordinating presentations of the technical infrastructure to take place at the M24 workshop in Nijmegen in June 2013.

Task force leaders coordinate related work in all centres:

- BAS (Christoph Draxler): Federated Content Search (FCS)  
- BBAW (Axel Herold): Resources  
- BBAW (Axel Herold): CMDI Metadata  
- IDS (Oliver Schonefeld): Authorization and Authentication Infrastructure (AAI)  
- MPI (Dieter van Uytvanck): Virtual Language Observatory (VLO)  
- MPI (Dieter van Uytvanck): PIDs  
- Hamburg (Timm Lehmberg): User questions and helpdesk  
- Leipzig (Volker Boehlke): Use Cases  
- Stuttgart (Heike Zinsmeister): Stand-alone tools und web applications  
- Tübingen (Thomas Zastrow): Web services

The M24 organisation team (Kathrin Beck, Volker Boehlke, Sebastian Drude, Dieter van Uytvanck, Thomas Zastrow) coordinates the schedules of the task forces and has been holding weekly virtual meetings since November 2012.
**Work Package 2: ERIC / DARIAH Liaison**

Work package 2 contains activities designed to connect CLARIN-D with related national and international projects, including CLARIN ERIC, the German national DARIAH-DE project, and eHumanities initiatives funded through German Federal Ministry of Research and Education (BMBF).

**Liaison with the CLARIN ERIC**

CLARIN ERIC is a European Research Infrastructure Consortium (ERIC)\(^5\) – a new category of legal entity established in European community law in order to facilitate the establishment and operation of joint research infrastructures of Europe-wide interest. CLARIN ERIC is the successor to the European CLARIN project in which CLARIN-D’s preparatory phase predecessor D-SPIN participated.

CLARIN-D personnel are involved at all levels of CLARIN ERIC. Erhard Hinrichs of the University of Tübingen is a member of the CLARIN ERIC board of directors in his capacity as chair of the CLARIN ERIC National Coordinators Forum. Dieter van Uytvanck of the Max Planck Institute for Psycholinguistics also sits on the board of directors by virtue of being the chair of the Standing Committee for CLARIN Technical Centres. Dr. Helge Kahler of the German Federal Ministry of Research and Education is President of the CLARIN ERIC.

Other CLARIN-D personnel serve in various roles within CLARIN ERIC, including:

- Axel Herold of the Berlin-Brandenburg Academy of Sciences and Humanities (BBAW) is a member of the CLARIN ERIC Standards Committee as well as National Content Manager for CLARIN-D.

- Erik Ketzan of the Institute for the German Language (IDS) is chair of the CLARIN-ERIC Legal Issues Committee.

- Heike Stadler (IDS) serves as Deputy National Content Manager.

- Thorsten Trippel of the University of Tübingen is a member of the CLARIN Office, providing support services for CLARIN ERIC.

- Andreas Witt (IDS) is chair of the CLARIN-ERIC Standards Committee.

Following the foundation of CLARIN ERIC, CLARIN-D signed the CLARIN Agreement (coordinated by Erhard Hinrichs and Kathrin Beck, University of Tübingen) in June 2012, specifying the national contributions of CLARIN-D members to CLARIN ERIC.

CLARIN-D centres and personnel participated in the CLARIN ERIC conference held in Sofia, 26-28 October 2012, and in the CLARIN ERIC Federated Content Search (FCS) workshop in Mannheim on 14 June 2012, discussing issues surrounding the CLARIN-D distributed resource search system.

---

Liaison with DARIAH-DE

CLARIN-D and DARIAH-DE have established a joint advisory board on technical issues, which had its first face-to-face meeting in Berlin on 10 April 2013. The two projects discussed a number of areas of common interest concerning technical infrastructure, particularly persistent identifiers for resources, identity verification services, and cooperative repository services.

The two projects have also three joint infrastructure-related working groups to discuss issues pertaining to:

- Infrastructure, standards, services, interoperability
- Curricula and training activities
- EU-Liaison activities

These groups are active in mediating discussions of technical standards for the two projects, in activities involving joint computing centres, and cooperative training and outreach.

CLARIN-D and DARIAH-DE organized a joint workshop at the conference *Digital Humanities 2012* in Hamburg, held 16 July 2012 and titled “Service Oriented Architectures (SOAs) for the Humanities: Solutions and Impacts. Joint CLARIN-D / DARIAH workshop at Digital Humanities Conference 2012.” Heike Neuroth from DARIAH-DE, and Peter Wittenburg, Erhard Hinrichs, and Thomas Zastrow from CLARIN-D organized this workshop.

Liaisons with BMBF eHumanities initiatives

In 2011, the German Federal Ministry of Research and Education (BMBF) launched a call for eHumanities research and development projects in the eHumanities (*Bekanntmachung des Bundesministeriums für Bildung und Forschung von Richtlinien zur Förderung von Forschungs- und Entwicklungsvorhaben aus dem Bereich der eHumanities*), which funds 24 projects. These projects involve a diverse array of humanities interests, running the gamut from automatic processing of legal argumentation, to employing data mining in lexicography, to analysing poetry and literature published during the Holocaust.

CLARIN-D centres participate in many of these projects as experts in digital text and language technologies, the Berlin-Brandenburg Academy of Sciences, the Institute for the German Language, the Max Planck Institute Nijmegen, Saarland University and the universities of Gießen, Leipzig, Stuttgart and Tübingen are collectively involved in nine such projects, providing advice, infrastructure and tools.

---


8 [http://www.bmbf.de/foerderungen/16466.php](http://www.bmbf.de/foerderungen/16466.php)
**Work Package 3: Technical Infrastructure**

The main achievements of the WP3 (technical infrastructure) work package in the period of May 2012 to April 2013 have been:

- Establishing **CLARIN-compatible repositories** in all of the 9 CLARIN-D centres. This significant step includes support for metadata (CMDI) harvesting, the allocation of persistent identifiers, and support for federated identity management.

- Participation in the Data Seal of Approval and CLARIN centre **assessment procedures**. Final outcomes are expected by May 2013.

- Filling, organizing and improving the **Centre Registry**, a human- and machine-readable store for information about the centres. The existing data category registry, relation registry and schema registry have been extended.

- The Component Metadata framework has been extended, as has the software that supports CMDI modelling (component registry, Arbil, OAI harvester, etc.). Special efforts went into remodelling the backend of the **Virtual Language Observatory (VLO)**.

- Practical tests of the **Authentication and Authorization Infrastructure (AAI)** have been performed to find out if enough attributes are available from the different Identity Providers (IdP). In general the CLARIN IdP and Discovery Service were improved, and some experiments with web services and AAI took place.

- OwnCloud has been implemented as a solution to host **personal workspaces**. The hosting offers from the computing centres were used to deploy more **CPU-intensive web services** and to establish **monitoring**.

- The **WebLicht** web service orchestration engine has been extended and many new data converters have been added. New web services have also been integrated. Connections to other CLARIN modules (Federated Content Search, VLO) are planned.

- All CLARIN-D centres have set up an endpoint for the **Federated Content Search (FCS)**. Next to that a global search content engine (“aggregator”) has been developed. The CLARIN-D specific SRU/CQL protocol extensions have been documented and sample code has been distributed. An online compatibility check has been launched.

- Regular monthly technical **developer meetings** (both real and virtual), where issues, ideas and plans are discussed among the developers at each of the centres, have helped to achieve interoperability and a good atmosphere.
Work Package 4: Discipline-specific Working Groups

In the second year of CLARIN-D the activities in work package 4 (WP4) were focused on three main tasks: curation projects, dissemination, and documentation.

Curation Projects
During the second year, the CLARIN-D steering committee approved nine curation projects. The WP4 management team provided support during the application process and reviewed all application documents for all proposed curation projects (primarily considering formal criteria) before passing them to the CLARIN-D steering committee. In order to improve this workflow, template documents and examples were created and/or updated, covering the whole process of applying for a curation project and resolving the relevant legal issues.

Communicating and discussing the legal issues surrounding curation project contracts was a recurring problem. Although answers to many important questions were collected and provided in advance, many participants needed further advice or documentation and this proved to be very time consuming for all parties involved.

Dissemination
The WP4 management team organized the CLARIN-D M12 Workshop in 2012 on 27 – 28 June in Leipzig. The workshop was mainly dedicated to potential CLARIN-D users with backgrounds in the social sciences and humanities and attracted over 100 guests from Germany and several other European states. The workshop gave a brief introduction to the CLARIN-D infrastructure and included presentations by CLARIN-D discipline-specific working groups, representatives of WP5 “Services and Resources,” and there was a poster and demo session for some of the tools already available in CLARIN-D.

Additionally the discipline-specific working groups and the WP4 management team contributed several articles to the CLARIN-D newsletter and presented CLARIN-D at various internal and external meetings, workshops and conferences.

Documentation
Documentation efforts included “The Conception of Potential Users' Questions” and documentation for digital resources (described in section “3.2 Towards a Documentation System for Digital Resources” of the full report). Since these tasks are community-specific and refer to other CLARIN-D work packages (e.g. WP7 “Support and Help Desk”), the discipline-specific working groups (especially WG1) worked in close collaboration with other CLARIN members to address documentation issues.
Work Package 5: Services and Resources

In the second year of the project, WP5 activities have remained focused on curating own tools and linguistic resources of participating centres, a task that is largely complete for those tools and resources in place at the beginning of the project. By the end of the second year, each centre had prepared and provided CMDI metadata for its resources and tools through the OAI-PMH interface of their respective repositories. In the third project year, curation services will only be provided for resources and tools from other projects that are still being actively developed. With the assistance of the Virtual Language Observatory (VLO), all CLARIN-D centre metadata will be aggregated and accessible via a faceted search engine.

In close cooperation with WP3, CLARIN-D centres’ text corpora have become accessible to distributed queries over various repositories via Federated Content Search (FCS) starting in the second project year. As part of the integration plan, additional resources and tools will be made available to participants in WP5 involved in the development of data formats for the WebLicht infrastructure project.

On 15 December 2012, the CLARIN-D User Guide\(^9\) was published. It provides a clear and concise overview of all aspects of CLARIN-D infrastructure. Enriched with notes and links to more information, it is an ideal introduction to the CLARIN-D world, as well as a practical guide to adapting and integrating existing language resources into the CLARIN-D infrastructure. Personnel from all CLARIN-D centres participated in the elaboration of this handbook, which will be regularly updated and expanded in the future. This book is presently only available in English.

WP5 organized a workshop at the conference Historical corpora for the humanities and social sciences. Use issues and perspectives (18-19 February 2013, Berlin), which was held in cooperation with the Deutsches Textarchiv project (German Text Archive). An introduction to CLARIN-D and its infrastructure was presented together with hands-on demonstrations of technical aspects in integrating tools and resources into the Weblicht system. Over 60 researchers from various humanities disciplines attended the event.

WP5 will devote the third year of the project primarily to evaluating the curation and integration of resources and tools from outside of the CLARIN-D centres. Planning and coordination for this phase of the project was already underway at the end of the second year. To this end, there is an expanded dialogue with the technical working groups as well as the coordinators of WP4. In close consultation with members of the F-AGs, WP5 will develop evaluation methodologies in order to assess the value of various external resources for particular research communities. These domain-specific metrics and evaluation schemes will be presented for discussion at a workshop planned for September 2013. The resulting evaluation methodologies will be documented in the CLARIN-D user manual and form the basis for selecting new resources and tools for curation in the third year.

\(^9\) [http://de.clarin.eu/de/sprachressourcen/benutzerhandbuch.html](http://de.clarin.eu/de/sprachressourcen/benutzerhandbuch.html)
Work Package 6: Legal and Ethical Issues

Work Package 6 focuses on legal and ethical issues regarding the usability and re-usability of language resources in CLARIN-D. The overall goal of WP6 is to provide useful legal information and resources for CLARIN-D partners. The tasks and goals of WP6 are guided by the outline and milestones laid out in the CLARIN-D project description document (Vorhabenbeschreibung).

The Institut für Deutsche Sprache (IDS) has created a comprehensive Legal Issues Platform, currently available on the CLARIN-D Wiki.\(^\text{10}\),\(^\text{11}\) The Legal Information Platform provides information, analysis, and best practices guidelines for legal issues affecting language science, as well as sample forms and reference contracts. In addition to existing sections on copyright, data privacy, and other fundamental legal issues, resources and best practices guidelines have been added over the last year concerning:

- Building corpora from open resources (Wikipedia, etc.)
- Building corpora from online text in general.
- Building corpora from online forums.
- Best practices for dealing with publishers and authors.
- Planned changes to German, EU, and UK law.
- Sample disclaimer forms for spoken interviews for oral corpora.
- Orphan works and the Orphan Works Directive.
- Sharing software through open source licenses.
- Sharing non-software through open source licenses.

We have also added a legal issues bibliography specifically for language science.

The content created, developed, and fine-tuned for the Legal Information Platform is eventually released as formal publications. This past year, WP6 team members have co-authored a book with chapters on legal issues and published an article on best practice guidelines on data privacy issues for oral interviews through the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft). More publications are in the works and in various stages of completion.

The WP6 team gives frequent presentations and lectures at CLARIN-D events, digital humanities conferences, and as invited speakers at university classrooms, providing education to the CLARIN-D and other research communities on legal issues affecting language resources and digital humanities in general.

The WP6 team is active in the new CLARIN Legal Issues Committee, and has begun representing CLARIN at European Commission hearings in Brussels focused on clarifying the legal situation regarding academic use of text and data mining as well as...


\(^\text{11}\) [http://de.clarin.eu/de/home/arbeitspakete/ap-6-rechtliche-und-ethische-fragen](http://de.clarin.eu/de/home/arbeitspakete/ap-6-rechtliche-und-ethische-fragen)
streamlining, simplifying, and reducing the cost of licensing text and data for researchers. Representation and lobbying of this kind will continue in the future.

Members of the WP6 team assisted with internal CLARIN and CLARIN-D legal issues, including the finalization of the CLARIN Service Provider Federation Agreements. They have begun a set of legal submission guidelines for resources within CLARIN, and will continue work on integrating legal information into metadata as well as legal policies for long-term archiving of data.
**Work Package 7: Support and Help Desk**

The focus of the work undertaken by WP7 between M12 and M24 was on the implementation of the first productive version of the CLARIN-D help desk with the help of the OTRS\(^{12}\) ticketing system and a CLARIN-D-wide test run. This led to a re-evaluation of the ‘Taxonomy of Support Areas’ and ticketing workflows that were initially implemented.

**Implementing and refining the ‘Taxonomy of Support Areas’**

A ‘Taxonomy of Support Areas’ that mirrors and orders CLARIN-D activities has been implemented as an integral part of the ticketing system through a set of hierarchically ordered queues. At present, the full taxonomic structure of the queue system has not been comprehensively defined. We expect actual usage by researchers and the CLARIN-D community after milestone M24 to reveal which categories are necessary and which are superfluous.

We have also taken advantage of the ability to assign tickets to freely definable ticket types and to customize them by so-called dynamic fields with the OTRS system, so that we do not have to over-specify the taxonomic structure of the help system.

**Rethinking Ticketing Workflows**

Following discussion during the ‘Support and Help Desk’ workshop in November 2012 (see below), we decided to blur the boundary between first and second level support by having a researcher at each CLARIN-D centre act as a supervisor for first level support as well as acting as a second level support agent replying to complex user inquiries. This means that student assistants will not be exclusively responsible for first level support.

**Implementation and Test Run**

The testing system was hosted on a Kernel-based Virtual Machine (KVM)\(^{13}\) using QEMU\(^{14}\) for emulation on hardware running Scientific Linux.\(^{15}\) We are in the process of moving the system to the Forschungszentrum Jülich, where the system will be hosted after the start of the production phase at M24.

100 typical user requests that had been collected before from members of the ‘Discipline-specific Working Groups’ (F-AGs – ‘Fachspezifische Arbeitsgruppen’)\(^{16}\) were entered into the system as a test of the system with relevant inquiries. Replies to those inquiries were documented with the help of the OTRS FAQ-module so they could serve as knowledge base for handling future requests.

---

12 http://www.otrs.com/

13 http://www.linux-kvm.org/

14 http://www.qemu.org

15 http://www.scientificlinux.org/

16 http://de.clarin.eu/en/discipline-specific-working-groups
Work Package 8: Training and Education

CLARIN-D, as an infrastructure for language resources and technology, has an explicitly user-oriented mission. The fundamental challenge for this project is to awaken in new potential users the possibilities a digital text and language oriented research infrastructure can offer them. Prejudices and obstacles to the acceptance and use of digital language technologies have to be addressed and users have to become acquainted with their use.

During the first year of the project, we identified three different contexts where we can provide training and education:

- Integrating language resources and language tools into university degree programs.
- Training of teaching staff.
- Assessing the qualifications of researchers.

To address these three contexts, we undertook the following measures:

- The Teaching and Learning Material Collection (TeLeMaCo): a collaborative portal for all kinds of training and teaching materials relevant to linguistics and digital humanities.
- Hosting individual researchers
- Constructing training modules and tutorials.
- Cooperating with WP7 (Support & Helpdesk) and DARIAH-DE.

Teaching and Learning Material Collection (TeLeMaCo)

Large amounts of specialized, domain-specific data, resources and tools are available on the web and are in principle accessible to researchers. However, since academic resources are distributed over many websites and often are not interlinked, nor indexed in one place, and do not comply with any common standard, gathering relevant resources remains a cumbersome task for academic users.

Our portal – TeLeMaCo (Teaching and Learning Material Collection) – provides access to the following types of material related to language resources and tools:

- Technical documentation (including tutorials, quick starts, etc.)
- Learning material for self-study.
- Short teaching modules (2-4 hours) that can be integrated into existing courses.
- Full courses covering a broader spectrum of language resources and tools or focusing on specific topics of application of language resources and tools.

TeLeMaCo is designed as with “social network”-style applications in mind. Registered users may both make their own educational materials available on the web and use
TeLeMaCo to find relevant materials, as well as provide comments on the materials available via TeLeMaCo.

**Hosting**

WP8 provides for humanities researchers with specific research interests to receive financial aid to visit a CLARIN-D centre for up to two months. Requirements for the funding are:

- A concrete research plan.
- A report about the stay.
- A presentation of the work at the hosting centre.
- The integration of resulting resources or tools in CLARIN-D, where applicable.

**Training modules and tutorials**

**CLARIN-D Doktorandentage (PhD days)**

The CLARIN-D *Doktorandentage* are intended as a small series of training events targeting young researchers, mostly PhD students. Each event has a specific thematic focus (e.g. corpora, speech, historical data).

The first CLARIN-D PhD days took place on 25–26 March 2013 at the Institute for Natural Language Processing, University of Stuttgart. The event focused on corpora and consisted of training in basic corpus linguistic methods, corpus search with regular expressions using CQP, syntactic annotation and corpus search using WebLicht and other CLARIN-D resources, and statistical analysis using R.

Follow-up events with other thematic foci are planned for the next project year, e.g. in the area of speech data and historical texts.

**Tutorials and Courses**

Individual CLARIN-D centres hold tutorials, courses, demos, and presentations on a regular basis. Those centres that are involved in university teaching routinely employ CLARIN-D resources and tools in their courses.

**Cooperation**

**Cooperation with WP7**

In order to better assist and help potential users, close cooperation between WP7 (*Support & Helpdesk*) and WP8 (*Training & Education*) is essential. The helpdesk platform, TeLeMaCo, training sessions, and hosting activities need to be coordinated to provide users with as much information and assistance as needed.

WP7 and WP8 have conducted a joint workshop in order to discuss the potential for coordinated activities.

**Cooperation with DARIAH WP2**

The CLARIN-D team in Saarbrücken have participated in two meetings of WP2 of the DARIAH-DE project. DARIAH-DE WP2 is concerned with research and education and is thus the counterpart work package to WP8 in CLARIN-D.
**Work Package 9: Dissemination and Acceptance**

WP9 focuses on dissemination of CLARIN-D activities and comprises running the CLARIN-D website, the publication of a newsletter every six months, and the production and distribution of flyers and brochures.

The Bavarian Archive for Speech Signals (BAS) in Munich and the Institute for Natural Language Processing (IMS) in Stuttgart collaborate in this work package. IMS is responsible for the website, and BAS for the newsletter, flyers and other communication channels. Heike Zinsmeister is responsible for WP9 work at Stuttgart, and Christoph Draxler is responsible at BAS.

**Principal Dissemination Strategies**

The principal dissemination strategies are:

- A website to publish detailed information about CLARIN-D, with both public and restricted areas, set up and maintained by IMS Stuttgart.
- A newsletter published every 6 months by BAS.
- Paper flyers designed and made available by BAS.

All CLARIN-D partners are able to use these information channels and all CLARIN-D partners contribute information about their activities to the partners responsible for WP9.

**Activities**

In 2012, the CLARIN-D website was completely redesigned. It now features a more attractive layout, event calendars, news and other features. Furthermore, the restricted access wiki system has been extended and is now a valuable forum for discussion and technical information exchange.

Two newsletters were published over the last year, containing descriptions of the CLARIN-D centres, reports on scientific use cases, and summaries of workshops, conferences and tutorials.

Two CLARIN-D flyers are now available for distribution: a large flyer with a list of centres and services, and a flyer with general information about CLARIN-D and how humanities researchers can use the CLARIN-D infrastructure. This second flyer is available in English and German.

Finally, CLARIN-D has actively participated in national and international conferences and workshops including LREC and Interspeech, but has also taken part in narrowly focused events such as the students’ conference StuTS, the IDS technology fair, and a conference on dialectology.

**Strategic Planning**

The focus of the WP9 will now slowly shift from technological achievements to practical applications for researchers. This entails extending the website and enhancing the visibility of CLARIN-D at major conferences and workshops. We believe young researchers are the most important potential users of the CLARIN-D infrastructure and will subsequently concentrate dissemination efforts on reaching this audience.