1 Introduction ........................................................................................................................................... 3
2 Liaison activities within CLARIN ............................................................................................................. 6
   Liaison activities with CLARIN ERIC .................................................................................................... 6
3 Liaison with other projects and initiatives .............................................................................................. 8
   DARIAH-DE: Joint Technical Advisory Board and joined activities .................................................... 8
   EUDAT – European Data Infrastructure and other Horizon 2020 projects ............................................. 8
   Special interest groups of DHd ............................................................................................................... 9
   Research Data Alliance ........................................................................................................................ 9
   Standardization ..................................................................................................................................... 9
   Liaison activities with other initiatives ................................................................................................. 10
4 User Liaison .......................................................................................................................................... 11
5 Further perspectives ............................................................................................................................... 13
1 Introduction

The CLARIN-D report R2.5 “Fifth Year Activity Report – Liaison” documents liaison activities during the fifth year of the CLARIN-D project. It presents an overview of cooperation within the CLARIN-D project as a whole, as well as the cooperative efforts of the individual CLARIN-D centres.

The CLARIN-D project1 (Web and Centre-based Infrastructure for the Social Sciences and Humanities) is motivated by a vision of providing humanities scholars with single-sign-on access to diverse repositories of language and textual data and the tools to work with them. It is a collaborative effort involving institutions and researchers at nine academic centres:

- Bavarian Archive for Speech Signals (BAS), Ludwig Maximilians University of Munich
- Berlin Brandenburg Academy of Sciences and Humanities (BBAW)
- Institute for the German Language, Mannheim (IDS)
- Max Planck Institute for Psycholinguistics, Nijmegen (MPI)
- University of Tübingen, Department of General and Computational Linguistics
- University of Hamburg, Hamburg Centre for Speech Corpora (HZSK)
- University of Leipzig, Institute of Computer Science
- Saarland University, Department of English and Translation Studies
- University of Stuttgart, Institute for Natural Language Processing

CLARIN-D builds on the achievements of the European CLARIN (Common Language Resource and Technology Infrastructure) initiative as well as the German national project D-SPIN (Deutsche Sprachressourcen-Infrastruktur). CLARIN and D-SPIN developed a set of common standards and goals for humanities research infrastructures as well as plans for the sustainable provision of digital resources and their long-term archiving.

The projects D-SPIN (Germany) and CLARIN (EU) were preparatory to CLARIN-D. D-SPIN’s funding was provided by the German Federal Ministry of Education and Research (BMBF), while CLARIN was funded directly by the European Commission. Both were part of the European Seventh Framework Programme (FP7)2 as well as the ESFRI Roadmap3 (European Strategy Forum on Research Infrastructures).

1 http://www.clarin-d.de/
2 http://cordis.europa.eu/fp7/home_en.html
3 http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri
FP7 was a strategic plan to encourage research and innovation in the European Union and associated states. It follows the first six framework programmes, which operated from 1984 through 2006, and covered the period from 2007 through 2013, encompassing roughly 50 billion euros of funding activities. It was replaced by Horizon 2020 in 2014.

ESFRI is an on-going European planning framework to coordinate research activities across Europe in order to establish competitive, open-access, continent-wide infrastructures designed to better serve research communities. It promotes the development of cross-border research facilities and services covering broad areas of scientific and academic interest.

ESFRI research infrastructure projects are planned with three sequential phases: a preparatory phase, an implementation phase, and an exploitation phase. D-SPIN was a preparatory phase project, during which plans for digital humanities infrastructure were developed through technical research, prototypes, models for training researchers, and through clarification of the legal framework for the project. CLARIN-D is the implementation phase successor to D-SPIN. Funding during the implementation phase of the ESFRI roadmap comes entirely from participating European countries that directly fund national projects to fulfil the roadmap’s goals.

D-SPIN’s European preparatory phase partner (CLARIN) has been replaced by CLARIN ERIC, which is structured as an association of national members. This new organization is a European Research Infrastructure Consortium (ERIC)4, – a new category of legal entity established in European Community law as part of ESFRI in order to facilitate the establishment and operation of joint research infrastructures of Europe-wide interest.

CLARIN ERIC currently has seventeen members, sixteen of whom are national project organizations joining together researchers and institutions in their respective states and one is an intergovernmental body:

- Austria (http://www.oeaw.ac.at/icltt/node/14)
- Bulgaria
- Czech Republic (http://lindat.mff.cuni.cz/lindat/)
- Denmark (https://clarin.dk/clarindk/forside.jsp)
- Estonia (http://keeleressursid.ee/en/)
- Finland (https://kitwiki.csc.fi/twiki/bin/view/FinCLARIN/FinClarinEsittely)
- Germany (http://www.clarin-d.net)
- Greece (http://www.clarin.gr/)

4 http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=eric
• Italy
• Lithuania (http://clarin-lt.lt/?lang=en)
• Netherlands (http://www.clarin.nl)
• Norway (http://clarin.b.uib.no/)
• Poland (http://www.clarin-pl.eu/en/)
• Portugal
• Slovenia (http://www.clarin.si/)
• Sweden (http://sweclarin.se/)
• United Kingdom (Observer, http://www.clarin.ac.uk/)
• The Nederlandse Taalunie (http://taalunie.org/wat-doet-taalunie/activiteiten/clarin), a joint Dutch-Flemish intergovernmental organization responsible for promoting and maintaining standards for the Dutch language

Membership is open to European Union countries and EU-associated entities such as intergovernmental organizations as well as individual centres. Several other countries are in the process of joining CLARIN ERIC. CLARIN-D and the other national CLARIN participants are encouraged to cooperate and coordinate their activities with other projects, at the national and local levels as well as internationally.

On the national level, CLARIN-D has aligned some of its activities with DARIAH-DE (Digital Research Infrastructure for the Humanities), another project on the ESFRI roadmap with a structure similar to that of CLARIN, focused on digital research infrastructure for arts and humanities research. DARIAH-DE is funded by the German Federal Ministry of Education and Research (BMBF) and is the German national participant in the DARIAH ERIC.

CLARIN-D members, both from centres and discipline specific working groups, are also involved in a number of BMBF-funded eHumanities projects, often bringing their experience with CLARIN-D and its infrastructure to these initiatives. Furthermore, as academic centres and universities, all CLARIN-D members are involved in multiple national and international projects, and they are encouraged to promote and use the CLARIN-D infrastructure whenever it can serve the scientific goals of researchers.

In order to facilitate interaction with users and other initiatives in various Digital Humanities disciplines, CLARIN-D interacts with associations, activities and users, supported by the Liaison Coordinator, Thorsten Trippel. This report encompasses the liaison and cooperative activities of CLARIN-D centres and CLARIN-D as a whole, with CLARIN ERIC, with other CLARIN ERIC members, with DARIAH-DE, with BMBF eHumanities projects, and with other national and international initiatives.
2 Liaison activities within CLARIN

This section summarizes the liaison activities of CLARIN-D centres with CLARIN ERIC and foreign CLARIN centres. CLARIN-D centres have a great amount of autonomy to initiate contact with CLARIN ERIC and work together with foreign CLARIN centres within the framework of this project.

Liaison activities with CLARIN ERIC

CLARIN-D partners participate at all levels of CLARIN ERIC. Until April 2016, Erhard Hinrichs of the University of Tübingen was a member of the CLARIN ERIC Board of Directors (BoD) in his capacity as chair of the CLARIN ERIC National Coordinators’ Forum (NCF). Though his term came to an end, he will continue to serve as a member of the NCF. Dieter Van Uytvanck, though CTO of CLARIN ERIC, is also part of the CLARIN-D developers group and is a member of the board of directors by virtue of being the chair of the Standing Committee for CLARIN Technical Centres (SCCTC).

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 and CCR (CLARIN Concept Registry) Coordinator for CLARIN-D.
- Erik Ketzan of the Institut für Deutsche Sprache (IDS) is chair of the CLARIN ERIC Legal Issues Committee.
- Dörte de Kok of the University of Tübingen served until April 2016 as administrative assistant of the National Coordinators’ Forum (NCF).
- Heike Stadler (IDS) serves as Deputy National Content Manager.
- Andreas Witt (IDS) is the co-chairperson of the CLARIN ERIC Standards Committee.

CLARIN-D centres and personnel participated in the CLARIN ERIC conference held in Wroclaw, Poland, October 15th–17th 2015, with the following contributions:

- *Interaction and dialogue with large-scale textual data: Parliamentary speeches about migration and speeches by migrants as a use case*, invited keynote presentation by Andreas Blätte, who is part of CLARIN-D’s discipline specific working groups,
- *Influence of Interface Design on User Behaviour in the VLO*, presented by Thomas Eckart, Alexander Hellwig and Twan Goosen

---

5 http://www.clarin.eu/
• On the trail of diachronic collocations, presented by Bryan Jurish

• Bootstrap a neural net dependency parser for German using CLARIN resources, presented by Daniël de Kok

• DMPTY - A Wizard For Generating Data Management Plans, presented by Thorsten Trippel and Claus Zinn

• Operationalisation of Research Questions of the Humanities within the CLARIN Infrastructure – An Ernst Jünger Use Case, presented by Dirk Goldhahn, Thomas Eckart, Thomas Gloning, Kevin Dreßler and Gerhard Heyer

• Turkish NLP tools in the CLARIN infrastructure, presented by Cagri Coltekin

• Towards depositor-controlled data ingest in digital repositories for linguistic corpora: automation of metadata processing, presented by Dennis Zielke and Daniel Jettka
3 Liaison with other projects and initiatives

This section outlines recent liaison activities between CLARIN-D and other projects and initiatives such as the ESFRI project DARIAH-DE and the FP7 and Horizon 2020 project EUDAT, CLARIN Plus and Parthenos. Selected collaborations with local initiatives of the CLARIN-D centres are also mentioned.

DARIAH-DE: Joint Technical Advisory Board and joined activities
In the past months, CLARIN-D and DARIAH-DE have intensified their cooperation. The cooperation is based on three different areas: the joint technical advisory board, joint dissemination activities, and strategic cooperation.

The most recent meeting of the joint Technical Advisory Board for CLARIN-D and DARIAH-DE took place at the IDS in Mannheim on February 17th, 2015. Due to the renewal phase of both initiatives in Germany, the next meeting of the advisory board has been delayed.

At the German speaking Digital Humanities conference in Graz, Austria in 2015, CLARIN-D and DARIAH-DE discussed options of joining forces for an information booth at successive conferences. This was successfully implemented at the DHd 2016, from March 8th to 11th 2016, in Leipzig. With the joint presence, the user community was provided with information on the services of both infrastructures adapted to the respective needs and the emphasis of the infrastructures. It is planned to continue with this cooperation also at other dissemination events. The CLARIN-D helpdesk has also been contacted by a group from DARIAH-DE to explore options to reuse the system within DARIAH-DE.

In 2015–2016, CLARIN-D and DARIAH-DE both had to undergo the reapplication for national funding. In the application process, CLARIN-D and DARIAH-DE created a new coordination council on strategic questions. This council is working on the reconciliation of the division of labour with regards to the disciplines, coordinating parallel developments to achieve synergies and discuss common topics, such as the long term models for operating the infrastructures. This led first to a joint introduction to the applications, and later to a joint declaration on the cooperation and division of labour, regular exchanges in joint video conferences on strategic questions and a face-to-face strategic workshop in April 2016. Another workshop is planned before the end of 2016.

EUDAT – European Data Infrastructure and other Horizon 2020 projects
Central infrastructure components are not only of interest to CLARIN but also to projects in
other subject fields. For the European Data Infrastructure project EUDAT, CLARIN is a key community to apply their services in the humanities and social sciences. EUDAT is a Horizon 2020 project, developing standard solutions for backup, replication, data sharing, etc. CLARIN evaluates the solutions to see if they can be utilized within the infrastructure. To liaise with EUDAT, CLARIN-D took part in the Kick-off meeting in Helsinki in March 2015.

In connection with CLARIN ERIC, CLARIN-D participates in various Horizon 2020 projects, namely CLARIN Plus, Parthenos and Europeana. In CLARIN Plus, CLARIN works on an increased integration of data, services and tools developed on the European level. CLARIN-D centres are key in defining requirements and contributing as partners to the individual tasks. In Parthenos, ESFRI projects exchange concepts and ideas. CLARIN-D centres are involved in standardization tasks, work on data management procedures and exchange ideas for user requirements. In Parthenos, participating CLARIN-D centres are in Leipzig, Berlin and Tübingen. The cooperation with Europeana focussed on establishing technical compatibility with the CLARIN infrastructure for a broad range of metadata sets (e.g., many fulltext collections of newspapers in various languages). Selected Europeana metadata sets are now regularly harvested by CLARIN’s Virtual Language Observatory (VLO), effectively making these resources available to the CLARIN community. This cooperation is planned to be extended later this year.

**Special interest groups of DHd**

Within the association of digital humanities in the German speaking area, dedicated special interest groups establish a platform for the exchange of requirements and ideas of different communities. CLARIN-D liaises with the groups on digital humanities and computer science, data centres, and legal and ethical issues. The groups meet at the DHd conferences and at other locations, for example the data centre group met in Hamburg during FORGE 2015.

**Research Data Alliance**

The Research Data Alliance (RDA) is an international interdisciplinary initiative of data centres, archives and research data infrastructures. CLARIN-D follows the developments and discussions, especially with regards to data management, policies and international collaborations. CLARIN-D participated in the Sixth RDA plenary in Paris in September 2015. CLARIN-D also presented on workflows in research in November 2015 in conjunction with colleagues from DARIAH and RDA.

**Standardization**

CLARIN-D continues to use and drive forward standardization of formats and procedures for creating, describing and processing resources. This standardization is intended for long term support of research data maintenance. With the CLARIN-D centres at the IDS and in Tübingen,
CLARIN-D is involved in the relevant working groups of the German Standardization Institute DIN with the subcommittee for language resources (NA 105-00-06 AA) and Systems for Maintenance of Terminology, Knowledge and Content (NA 105-00-05 AA), both with a deputy chair from CLARIN-D (Thorsten Trippel, Tübingen). Both groups focus on the semantic interoperability of metadata and resource formats. CLARIN-D is also representing Germany in the international committees, especially ISO TC 37 SC 4. In WP 6 of this standardization committee, a CLARIN expert, Andreas Witt of IDS, is the convenor. The Berlin CLARIN-D centre represents CLARIN-D in the activities of the TEI.

**Liaison activities with other initiatives**

CLARIN-D members belong to the larger German, European and international academic communities. They continue to be routinely present and active in academic conferences, workshops, and projects on many different scales. This involves, but is not limited to the Deutsches Textarchiv (German Text Archive), special research centres (SFB) SFB 833 Emergence of meaning – the dynamics and adaptivity of linguistic structures and the Stuttgart SFB 732 *Incremental Specification in Context* in the development of repositories and tools, and the BMBF funded project KobRA (Korpus-basierte linguistische Recherche und Analyse mit Hilfe von Data-Mining, project lead: University of Mannheim) as providers of resources. Also eHumanities projects funded by the BMBF liaise with CLARIN-D as part of the discipline specific working groups (for example in F-AG 7: Applied Linguistics, Computational Linguistics) of CLARIN-D or as partners of a CLARIN-D centre (as for example CRETA, the centre for reflective text analysis, at the University of Stuttgart).
4 User Liaison

CLARIN-D reaches out to user communities who are interested in applying data and services. This includes specific workshops, conference presentations, publications and contact with user communities. In 2015, CLARIN-D presented at the annual meeting of various associations. Among them were the German Society of Romance Philologies (“Romanistentag”) and the German Society of English Philologies (“Anglistentag”). CLARIN was also represented in the group of data centres of the association for Digital Humanities in the German language region, and published an overview article with the Bavarian Academy of Science. CLARIN-D also used dissemination events to contact users, such as the European Summer University in Digital Humanities in Leipzig. At the conference of the association of digital humanities in the German speaking area (DHd) in Leipzig 2016, CLARIN-D presented the services for the DHd community with an information booth jointly with DARIAH-DE.

At the dissemination event in Leipzig 2015, CLARIN-D launched its new website (see http://www.clarin-d.net) which features the core services of CLARIN-D in the areas of accessing data, analysing data and archiving data for the humanities and social sciences. The portal also provides step by step descriptions for exemplary tasks that can be accomplished using the infrastructure. The CLARIN-D infrastructure presents itself as easily usable for the intended community. Technical and administrative parts of the initiative are still available, but not featured prominently as the services address the humanities scholar.

Three technical tools have been developed for users to contact CLARIN-D: The CLARIN-D centre selector, the data management plan developer tool DMPTY and the helpdesk. With the CLARIN-D centre selector, CLARIN-D shows the users a way of locating a centre which has the expertise they need for their project, discipline or service. Though the wizard is simple with regards to the questions asked, and by no means tested for logical consistency, it is a low threshold tool for users to find appropriate CLARIN-D experts or to be redirected to the liaison coordinator to establish contact. More complex is the DMPTY tool, also available on the CLARIN-D website. This tools allows researchers planning to submit project proposals to develop a data management plan, which includes CLARIN-D depositing services, policies and standards. The complexity of data management plans allows a more in depth conversation of the researchers with CLARIN-D already during the planning phase of a research project. The helpdesk is another tool integrated in all CLARIN-D services and allowing users to contact

---

6 German article: https://www.badw.de/de/publikationen/akademieAktuell/2016/56/00_gesamte_Ausgabe_0116.pdf
experts with specific questions. The helpdesk as such is developed and maintained by Work Package 7.

Videos and screencasts on CLARIN-D activities were also produced and published\textsuperscript{7}. Videos on the European Summer University in Leipzig became available in English and German, videos on the helpdesk and individual use cases are usually – targeting a German audience – available in German only.

\textsuperscript{7} Usually published in the CLARIN-D youtube channel https://www.youtube.com/channel/UCSXAFB9S3tkR8XCDSNFz12A
5 Further perspectives

In the coming year, activities are planned with the University of Münster and the Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science as new contacts. Additionally, established contacts with associations such as the philologies, history, political sciences will continue. The cooperation with other groups such as RDA and DARIAH-DE will continue. As in previous years, work package 8 is co-organizer of the summer school “Culture & Technology – The European Summer University in Digital Humanities”.

On the technical side, liaison with users will be increased using the helpdesk of work package 7. This makes it possible not only to support individuals when using the CLARIN infrastructure but also to support potential users of CLARIN trying to establish contact with the infrastructure. The step by step guides to usage examples of CLARIN will also continue and be enhanced. Also, an increased number of tutorials for CLARIN resources and CLARIN services by videos and screencasts, targeting the user groups, will be part of the liaison activities.