Deliverable: R4.4: Annual report on the activities of the Discipline-specific Working Groups

Responsible: Prof. Dr. Gerhard Heyer

Editors: Prof. Dr. Gerhard Heyer, M.A. Gregor Wiedemann

Contributors: Prof. Dr. Thomas Gloning, Prof. Dr. Christian Mair, Prof. Dr. Nikolaus Himmelmann, Prof. Dr. Charlotte Schubert, Prof. Dr. Harald Baayen, Prof. Dr. Petra Wagner, Prof. Dr. Anette Frank, Prof. Dr. Cathleen Kantner, Prof. Dr. Gary Schaal, Prof. Dr. Simone Lässig, Prof. Dr. Martin Sabrow
Table of Contents

1. Introduction.......................................................................................................................... 4

2. Activities of the CLARIN-D WP4 management team.......................................................... 5
   2.1 General activities.................................................................................................................... 5
   2.2 New discipline-specific Working Groups of CLARIN-D....................................................... 5
   2.3 Curation projects.................................................................................................................. 6
   2.4 Meetings............................................................................................................................... 7
   2.6 Preparation of the 3rd CLARIN-D Dissemination Workshop (30.06./01.07.2015, Leipzig).... 8

3. Reports of the Discipline-specific Working Groups.............................................................. 9
   3.1 Working Group 1: German Philology..................................................................................... 9
   3.2 Working Group 2: English, Romance and Slavic Studies....................................................... 12
   3.3 Working Group 3: Linguistic Fieldwork, Anthropology, Language Typology......................... 15
   3.4 Working Group 4: Ancient History, Classical Philology, Archaeology.................................. 19
   3.5 Working Group 5: Psycholinguistics and Cognitive Psychology........................................... 23
   3.6 Working Group 6: Speech and other modalities.................................................................... 25
   3.7 Working Group 7: Applied Linguistics and Computational Linguistics................................. 29
   3.8 Working Group 8: Content Analysis in the Social Sciences.................................................. 36
   3.9 Working Group 9: Modern history......................................................................................... 39
   3.10 Working Group 10: Contemporary history......................................................................... 45
1. Introduction

The CLARIN-D work package 4 “Discipline-specific Working Groups” (WP4) is led by the CLARIN-D team located at the NLP Group (Abteilung Automatische Sprachverarbeitung) of the Department of Computer Science at the Leipzig University. The WP4 acts as a link between the CLARIN-D resource centers and the research communities which represent the users of the CLARIN-D infrastructure. Ten Working Groups (WGs) act as consultants for the needs of the humanities, social sciences and their sub disciplines. The ten WGs together consist of more than 170 academic professionals. Their main role is to advise CLARIN-D during the development and implementation of the infrastructure so that these efforts can best meet the needs of all research communities involved. They further coordinate dissemination and best practice using CLARIN-D services in their member communities. The ten working groups are:

- WG1: German Philology
- WG2: English, Romance and Slavic Studies (Other Philologies)
- WG3: Linguistic Fieldwork, Anthropology, Language Typology
- WG4: Ancient History, Classical Philology, Archeology
- WG5: Human Language Processing: Psycholinguistics, Cognitive Psychology
- WG6: Language and Multimodal Communication
- WG7: Applied Linguistics, Computational Linguistics
- WG8: Content Analysis for the Social Sciences
- WG9: Modern History
- WG10: Contemporary History

WP4 comprises of the management of joint activities of the working groups. This includes the organization of WG meetings, organization of specialized and interdisciplinary workshops and the creation of joint reports. Further, communications between CLARIN-D centers and the WG as well as groups among themselves are coordinated. This work is done in close cooperation with the heads and members of the WG. This documents reports on the results of the work done in CLARIN-D WP4 in the fourth year (01.06.2014 – 31.05.2015) of the CLARIN-D project.
2. Activities of the CLARIN-D WP4 management team

2.1 General activities

Virtual Meetings: Leaders of all ten WGs take part in a virtual meeting on monthly basis to report on progress in community activities, curation projects and coordination of the activities of the WGs with other CLARIN-D institutions. Virtual meetings were prepared and planned in close collaboration with Prof. Dr. Thomas Gloning (WG1, May – June 2014), Prof. Dr. Christian Mair (WG2, July 2014 onwards), the elected representatives of the discipline-specific working groups in the CLARIN-D Lenkungskreis. During these meetings important organizational dates and project details were communicated, current and future plans (activities, meetings, reports, ...) and problems were discussed. Protocols of these meetings are written and published in the CLARIN-D wiki.

Website/Newsletter: All WGs and their curation projects present themselves on the CLARIN-D website. WP4 coordinates these presentations and the maintenance of these parts of the website. Descriptions of recently started curation projects as well as for newly introduced WGs will be inserted until the end of this report period. Contributions to the CLARIN-D newsletter are written and organized.

Communication: WP4 distributes all relevant information between CLARIN-D institutions and the WGs, as well as information of the WGs among each other. For this, communication infrastructure such as mailing lists and information collections in the CLARIN-D wiki are maintained. Activities like the European Summer School C&T in summer 2015 in Leipzig organized by WP8 are advertised in the communities of the WGs.

Consortium meetings: WP4 reports on activities of the WGs in the quarterly CLARIN-D consortium meetings. In March 2015 a consortium meeting together with all leaders of the WGs was held in Braunschweig on invitation of WG9. The WGs presented their current curation projects in a poster session.

Staff turnover: Volker Boehlke, responsible for activities of WP4 in CLARIN-D from the first year on, left the project in December 2014. From January 2015 onwards Gregor Wiedemann took over activities of the WG coordination.
2.2 New discipline-specific Working Groups of CLARIN-D

During the reporting period three new WGs have been constituted:

- WG8 “Content Analysis for the Social Sciences” is chaired by Prof. Dr. Cathleen Kantner and Prof. Dr. Gary Schaal and focuses mainly on promoting use and best-practices of CLARIN-D infrastructure in political science and sociology.

- WG9 “Modern History” is chaired by Prof. Dr. Simone Lässig and focused on promoting use and best-practices of CLARIN-D infrastructure in modern history (around 1750-1850).

- WG10 “Contemporary History” is chaired by Prof. Dr. Martin Sabrow and focused on promoting use and best-practices of CLARIN-D infrastructure in contemporary history (around 1945 onwards).

All three new WGs rather focus on content analytic aspects of computer-linguistic applications to make use of large collections of digital text data for their discipline. From their participation we expect synergies and advantages for use and best practices through the CLARIN-D infrastructure across these disciplines.

The heads of the new discipline-specific working groups took part in virtual meetings, prepared constitutional meetings of their working groups and coordinated the development of curation projects. Detailed information on activities of WGs are given in section 3 of this report.

2.3 Curation projects

During the fourth year of CLARIN-D, all previously designed and approved curation projects were brought to an end. The WP4 management team provided information and templates for the creation of final reports and certifications by external experts that are part of the finalization process for curation projects in CLARIN-D. The final reports for all CLARIN-D curation projects are available on the CLARIN-D wiki.

For the current phase of CLARIN-D existing sketches for curation projects from the previous phase were transformed into applications for eight new projects. Applications were reviewed, each by two CLARIN-D centers, and approved by the CLARIN-D Lenkungskreis. Each of the three newly introduced WGs applied successfully for a curation project as well (see WG specific reports). Activities of development and application for curation projects were coordinated by the WP4 management team.
The eight new curation projects which started in early 2015 are:

- **WG1:** ChatCorpus2CLARIN: Integration des Dortmunder Chat-Korpus in die CLARIN-D Korpusinfrastrukturen am Institut für deutsche Sprache und an der Berlin-Brandenburgischen Akademie der Wissenschaften (Applicants: PD Dr. Michael Beißwenger, Prof. Dr. Angelika Storrer)

- **WG2:** Überführung des Old Bailey Corpus (1720-1913) in ein CLARIN-kompatibles Format (Applicant: Prof. Dr. Magnus Huber)

- **WG4:** Ausbau und Erweiterung eines Open-Source Tools zur Nachkorrektur historischer OCR-erfasster Texte (Applicant: Prof. Dr. Klaus U. Schulz)

- **WG5:** An Open Science platform for Corpus Linguistics – Broadening the scope of the Mind Research Repository (Applicant: Prof. Dr. Anke Lüdeling)

- **WG7:** Semantische Annotation für Digital Humanities (Applicants: Prof. Dr Anette Frank, Prof. Dr. Chris Biemann, Dr. Richard Eckart de Castilho, Prof. Dr. Iryna Gurevych)

- **WG8:** Plenarprotokolle als öffentliche Sprachressource der Demokratie: Klassifikation von Plenardebatten im PolMine-Plenarprotokollkorpus (Applicants: Prof. Dr. Andreas Blätte, Prof. Dr. Gary S. Schaal)

- **WG9:** Quellen des Neuen: Realkundliches- und naturwissenschaftliches Wissen für Dilettanten und Experten zwischen Aufklärung und Moderne (Applicant: Prof. Dr. Gerhard Lauer)

- **WG10:** Kuration des „DDR-Presseportals“ und Evaluierung der CLARIN-D-Services als Grundlage für die zeithistorische Forschung (Applicant: Prof. Dr. Rüdiger Hohls)

Due to administrative problems and shifts in WG composition WG3 and WG6 could employ their coordinators only very late in the current project period. This led to problems on preparation of applications for a curation project. At the moment WG6 is preparing an application to set up a new curation project during the current phase of CLARIN-D.

### 2.4 Meetings

The WP4 team takes part in CLARIN-D institutional meetings such as the Lenkungskreis and the Consortium. In special cases it further takes part in meetings of single WGs to report on CLARIN-D activities or support networking between WGs.
CLARIN-D CONSORTIUM MEETING (23.03.2015, BRAUNSCHWEIG)

The CLARIN-D consortium meets quarterly to discuss current progress and problems on realization of the project goals. The meeting in Braunschweig in March 2015 was especially conceptualized to exchange with the WGs. Hosted by WG9 (Prof. Dr Simone Lässig, Georg-Eckert-Institut) leaders of the WGs met together with representatives of the CLARIN-D centers to introduce their activities and curation projects. For introduction of the curation projects, the WGs prepared posters which were presented during the first day of the meeting. During the meeting progress on applying the CLARIN-D infrastructure for best practices in the single disciplines were discussed along with opportunities for synergies and collaboration across disciplines through the use of the infrastructure.

WG7 MEETING (09.03.2015, HEIDELBERG)

In early March 2015 WG7 held its WG meeting in Heidelberg to discuss status and progress of its activities, especially for extensions of the new version of the annotation tool WebAnno. During this workshop WP4 presented needs for annotation capabilities of the WGs more oriented towards content analysis than linguistic annotation. Opportunities to integrate functionality for semantic annotation to meet the needs of social sciences / historians were discussed.

2.6 Preparation of the 3rd CLARIN-D Dissemination Workshop (30.06./01.07.2015, Leipzig)

In summer 2015 WP4 will organize the third dissemination workshop of the CLARIN-D discipline specific working groups. For this workshop preparation started in January 2015 to determine site, dates and topics, and to set-up a preliminary workshop program. Requirements and ideas of WG leaders were coordinated. The workshop will take place on 30.06.-01.07.2015 at the Mediencampus Leipzig. The CLARIN-D infrastructure and discipline specific use cases of the WGs will be presented along the three meta topics 1) deposit, 2) access, 3) analyze which provide guidance for the main contributions of CLARIN-D services to the humanities and social sciences. On the second workshop day Prof. Dr. Christiane Fellbaum will hold a keynote on benefits of infrastructures for the digital humanities and opportunities for interdisciplinary cooperations. The workshop website can be found at http://clarin2015.informatik.uni-leipzig.de
3. Reports of the Discipline-specific Working Groups

3.1 Working Group 1: German Philology

CHAIR

Prof. Dr. Thomas Gloning
Justus-Liebig-Universität Gießen
Institut für Germanistik
Otto-Behaghel-Straße 10B
35394 Gießen

MEMBERS

• Jurgita Baranauskaite (research assistant)
• Stefanie Seim (research assistant)

WG1 has been joined by five new members:

• PD Dr. Michael Beißwenger, TU Dortmund University, michael.beisswenger@tu-dortmund.de
• Prof. Dr. Arnulf Deppermann, Institute for German Language (Mannheim) and University of Mannheim, deppermann@ids-mannheim.de
• Prof. Dr. Dietmar Rösler, University of Gießen, dietmar.roesler@germanistik.uni-giessen.de
• Prof. Dr. Ingrid Schröder, University of Hamburg, ingrid.schroeder@uni-hamburg.de
• Prof. Dr. Angelika Storrer, University of Mannheim, astorrer@mail.uni-mannheim.de

ACTIVITIES

A) Recent activities

• Virtual Meeting of WG1 (June 27th, 2014): Organisation; Participation
• Annual Meeting of WG1 (October 7th, 2014 in Gießen): Organisation; Participation
• Euralex Preconference Workshop (July 14th, 2014 in Bolzano): Participation
• Workshop of Kochbuchforschung (August 30th, 2014 in Freising): Participation
• **Meeting of DARIAH Advisory Board** (September 12th, 2014 in Würzburg): Participation


• **Second DTA/CLARIN-D Conference**: *Textkorpora in Infrastrukturen für die Geistes- und Sozialwissenschaften* (November 17th/18th, 2014 in Berlin): Participation of research assistants

• **Workshop** of the Institute for German Language: *Die Zeitung als das Medium der neueren Sprachgeschichte? Korpora, Analyse und Wirkung* (November 20th/21st, 2014 in Mannheim): Presentation on *Alte Zeitungen und Forschungsinfrastrukturen (CLARIN-D): Korpusaufbau und historisch-lexikographische Nutzungsperspektiven*

• **Conference Sprachgeschichte und Medizingeschichte: Texte – Termini – Interpretationen** (November 23rd–25th, 2014 in Heidelberg): Presentation on *Wie kann ein modernes Dokumentationssystem zum deutschen Sprachgebrauch der Medizin von den Anfängen bis zur Gegenwart aussehen? Vorschläge zum Textcorpus, zu Darstellungsformen, zu Kollaborationsformaten und zu wissenschaftsgeschichtlichen Bezügen*

• **51. Jahrestagung des Instituts für Deutsche Sprache** (March 10th–12th, 2015 in Mannheim): Presentation on *Wie verändern neue mediale Formate die kommunikativen Praktiken in der Wissenschaft? Prinzipien des Wandels und Fallbeispiele aus Geschichte und Gegenwart*

• **Meeting of WG heads** (March 23rd, 2015 in Braunschweig): Participation in the poster session

**B) Further activities**

• Participation in the monthly virtual conferences of WG heads

• Organisation of (bi)weekly consultation sessions with research assistants

• Further work on the documentation of resources; development of a documentation system in cooperation with WP4 (Leipzig)

• Contribution to the prototype of the Tübingen “resource letter”
• Further work on the compilation of user questions and usage scenarios of digitally supported research in German studies

• Establishing contacts to professional associations coping with the German language

• Consulting CLARIN-D service centres with regard to specific questions

C) Planned activities

• Annual Meeting of WG1 (roughly scheduled for autumn 2015)

CURATION PROJECTS

The third curation project of WG1, *ChatCorpus2CLARIN*, was approved by the steering committee on September 24th, 2014. Being brought forward by PD Dr. Michael Beißwenger and Prof. Dr. Angelika Storrer, Curation Project III started running on March 1st, 2015 and will be finished by the end of December 2015.

BENEFITS OF CLARIN-D FOR THE SCIENTIFIC COMMUNITY ADDRESSED BY THE WORKING GROUP

The scientific community will take advantage from the development of WG1’s documentation system of digital resources, tools and services, since the latter facilitates the process of finding relevant data via metadata in a central place and a sustainable way. Researchers will also profit from the above-mentioned catalogue of user questions towards CLARIN-D, which is intended to become a guideline for getting started with the resources CLARIN-D offers resp. for selecting which resource within the context of CLARIN-D might be used in the conduct of one’s own studies.

PUBLICATIONS


3.2 Working Group 2: English, Romance and Slavic Studies

CHAIR
Prof. Dr. Dr. Christian Mair

English Department
Freiburg University

MEMBERS

• Prof. Dr. Markus Bieswanger, University of Bayreuth, English Linguistics, bieswanger@uni-bayreuth.de
• Prof. Dr. Jürgen Handke, University of Marburg, English Linguistics, handke@staff.uni-marburg.de
• Prof. Dr. Magnus Huber, University of Gießen, English Linguistics and History of English, magnus.huber@anglistik.uni-giessen.de
• Prof. Dr. Dr. Christian Mair, University of Freiburg, English Linguistics, christian.mair@anglistik.uni-freiburg.de
• Prof. Dr. Roland Meyer, Humboldt-University Berlin, Slavic Studies, Roland.Meyer@hu-berlin.de
• Prof. Dr. Hagen Peukert, University of Bremen, Linguistics and Literary Studies, peukert@uni-bremen.de
• Prof. Dr. Stefan Pfänder, University of Freiburg, Romance Studies, stefan.pfaender@romanistik.uni-freiburg.de
• Dr. Cornelius Puschmann, Berlin School of Library and Information Science (IBI), cornelius.puschmann@hiig.de
• Dr. Christoph Schöch, University of Würzburg, Department for Literary Computing, c.schoech@gmail.com
• Prof. Dr. Elke Teich, Saarland University, English Linguistics and Translation Studies, E.teich@mx.uni-saarland.de
• Prof. Dr. Monika Wingender, University of Gießen, Slavic Studies, Monika.Wingender@slavistik.uni-giessen.de

ACTIVITIES

Recent dissemination activities:

GAL Conference (16.–19.09.2014 in Marburg): There was an information booth on CLARIN-D at the annual conference of the “Gesellschaft für Angewandte Linguistik”. In addition, Christian Mair showcased the CLARIN-D infrastructure in his invited plenary on “Teaching linguistics: data-driven, but who’s in the driver’s seat?”

Joint working group and consortium meeting (23.–24.03.2015 in Braunschweig): Magnus Huber and his programmer Magnus Nissel presented a poster of the working group's current curation
project “Transformation of the Old Bailey Corpus 1720-1913 and its search interface into a CLARIN-compatible format”. Christian Mair gave an overview of the working group’s current activities. Udo Baumann, assistant to the working group’s chair and successor to Claudia Winkle in this function from Feb. 2015 also attended the meeting.

**Planned dissemination activities:**

**CLARIN-D dissemination workshop (30.06.-01.07.2015 in Leipzig):** Magnus Huber will present a poster of the working group’s current curation project “Transformation of the Old Bailey Corpus 1720-1913 and its search interface into a CLARIN-compatible format”.

**Working group meeting (01.07.2015 in Leipzig):** A working group internal meeting is planned to take place in Leipzig; directly after the CLARIN-D dissemination workshop.

**Romanistentag (26.-29.07.2015 in Mannheim):** A workshop on CLARIN-D will be held together with the CLARIN-D Centre in Tübingen. The workshop is mainly addressed to doctoral students and early-career researchers but also open to other interested participants.

**Anglistentag (23.09.2015 in Paderborn):** Christian Mair and Thorsten Trippel (CLARIN-D Centre Tübingen) will organize a workshop on CLARIN-D’s tools and resources. Like the workshop at the Romanistentag, this workshop is mainly addressed to doctoral students and early-career researchers.

**Curation Projects**

The working group has initiated four curation projects. Two have been completed, one is in progress, and one had to be abandoned owing to administrative obstacles at the proposed host institution. The two completed projects are: CP1 – “Implementation of a web based platform for the structured documentation of languages in the mobile age” and CP2 – “Indexing of digital text archives through metadata and lemmas”.

**Curation project 3:** “Transformation of the Old Bailey Corpus 1720-1913 and its search interface into a CLARIN-compatible format”

Magnus Huber from Gießen University is responsible for the realization of the working group’s third project. After considerable difficulties with the contracts, they have finally been signed and work on the project started in January 2015. The project aims at the integration of the *Old Bailey Corpus* (*OBC*) and its search interface into the CLARIN-D infrastructure. To do so the data need to be transformed into CLARIN-compatible formats (CMDI and TEI). Although the *OBC* is already a full-fledged corpus with an own search interface, the integration of the corpus into the CLARIN-D
infrastructure promises added value. Its visibility will be increased and its sustainability secured.

**Benefits of CLARIN-D for the Scientific Community Addressed by the Working Group**

The scientific community addressed by the working group benefits from the curation projects outlined above since they fill gaps in the repertoire of resources and tools which have been on offer so far. Being integrated into the CLARIN-D website, the tools and resources can easily be accessed, their visibility is increased and their sustainability secured. Furthermore, the clear descriptions provided on the website make them attractive especially to researchers who have only little experience in working with or have been reluctant to using digital resources and tools so far.
3.3 Working Group 3: Linguistic Fieldwork, Anthropology, Language Typology

CHAIR
Prof. Dr. Nikolaus Himmelmann
Institut für Linguistik
Universität zu Köln

MEMBERS

- Peter Bouda MA, Interdisciplinary Centre for Social and Language Documentation, Minde Portugal, pbouda@cidles.eu
- PD Dr. Michael Cysouw, Ludwig-Maximilians-Universität München, Research Unit Quantitative Language Comparison, cysouw@lmu.de
- Dr. Sebastian Drude, Max-Planck-Institut für Psycholinguistik, Nimwegen, The Language Archive, Sebastian.Drude@mpi.nl
- Prof. Dr. Volker Gast, Friedrich-Schiller-Universität Jena, Department of English and American Studies, vgast@uni-jena.de
- Prof. Dr. Ralf Gehrke Goethe-Universität Frankfurt am Main, Institut für Empirische Sprachwissenschaft, gehrke@rz.uni-frankfurt.de
- Prof. Dr. Geoffrey Haig, Universität Bamberg, Institut für Orientalistik, geoffrey.haig@uni-bamberg.de
- Dr. Dagmar Jung, Universität zu Köln, Institut für Linguistik, djung@uni-koeln.de
- Dr. Sebastian Nordhoff, Max-Planck-Institut für evolutionäre Anthropologie, Leipzig, Department of Linguistics, sebastian_nordhoff@eva.mpg.de
- Kilu von Prince MA, ZAS Berlin, prince@zas.gwz-berlin.de
- Prof. Dr. Elena Skribnik, Ludwig-Maximilians-Universität München Institut für Finnougristik / Uralistik Skribnik@lmu.de
- Dr. Sabine Stoll, Universität Zürich, Seminar für Allgemeine Sprachwissenschaft, sabine.stoll@uzh.ch
- Gereon Ullmann MA, Universität Erfurt, Seminar für Sprachwissenschaft, gereon.ullmann@uni-erfurt.de
- Dr. Claudia Wegener, Universität Bielefeld, Fakultät für Linguistik und Literaturwissenschaft, claudia.wegener@uni-bielefeld.de
• Prof. Dr. Thomas Widlok, Radboud Universiteit, Nijmegen, Anthropology and Development Studies. Thomas.Widlok@mpi.nl
• Taras Zakharko MA, Universität Zürich, Seminar für Allgemeine Sprachwissenschaft, taras.zakharko@uzh.ch
• Employee of the working group: Felix Rau MA, Universität zu Köln, Institut für Linguistik, f.rau@uni-koeln.de

During the reporting period no changes of membership occurred.

ACTIVITIES

A) Recent activities

The working group chair and members of the working group participated in several workshops, conferences, and summer schools to disseminate CLARIN-D tools, services, and topics. This was done in coordination with the CLARIN-D center The Language Archive of the MPI Nijmegen. Members of the WG took part in the summer schools Coding for Language Communities 2014 in Minde Portugal from August 11 - 15th and Community-driven Language Documentation 2014 also in Minde Portugal from August 18 - 23rd. During these summer schools the results of the curation projects CLARIN-D F-AG 3 KP 1 Poio API and CLARIN-D F-AG 3 KP 2 Field Linguistic Tool Repository were taught as well as the use of CLARIN tools such as ELAN.

The chair of the working group as well as members were present and presented on the Third INNET Conference 5-6th September, 2014 in Budapest. Innovative Networking in Infrastructure for Endangered Languages (INNET) is an EU-funded project aims to intensify the worldwide archiving grid and expert networks, to disseminate state-of-the-art language technology from the CLARIN realm, and to foster the use of archives by schools and the general public. The working group has actively cooperated with the INNET project. The working group also participated in the INNET Regional Archives Workshop 8-9th September 2014 in Nijmegen and especially demonstrated the CMDI Maker which was developed as part of the CLARIN-D F-AG 3 KP 2 Field Linguistic Tool Repository.

The working group chair, working group members, as well as representatives of the CLARIN-D center The Language Archive Nijmegen organized and participated in the DFG Workshop and spring school Primer Taller bilateral Alemania-Mexico 19-21st March as well as the Primera Escuela de Documentation y Tipologia Lingüística 23-28th March 2015, both in Morelia (Mexico).

B) Further activities
The working group was represented in the monthly virtual meetings of working group chairs and representatives of CLARIN-D working package 4.

Furthermore, the working group counseled language documentation and typology research project in the deployment of CLARIN tools and services.

The working group is committed to a continued development of the metadata tool CMDI Maker. The CMDI Maker is a tool for the generation of metadata in the IMDI profile of CMDI. Following its development in the CLARIN-D F-AG 3 KP 2 Field Linguistic Tool Repository, the functionality of the CMDI Maker was funded by the Endangered Language Documentation Project (ELDP) for the extension of the CMDI Maker to include the ELDP CMDI profile.

**Curation Projects**

The working group had already submitted a proposal for a third curation project by December 2013. Changes to the rules for curation projects required fundamental revisions of the proposal. Due to the delayed signing of the contract between the University of Cologne and the Max-Planck Institute in Nijmegen and the resulting ongoing administrative expenditures the project was canceled January 2015.

**Benefits of CLARIN-D for the Scientific Community Addressed by the Working Group**

For the scientific community addressed by the F-AG 3, CLARIN-D enables the adoption of language technologies developed in the area of corpus linguistic and NLP and employs it for the linguistically diverse data that is subject of analysis in language typology and linguistic fieldwork. Furthermore, CLARIN-D fosters the employment of RESTful web services in the language archive infrastructures such as the DoBeS archive.

The fact that the DoBeS archive has a complex system of individual user access rights, which is implemented in Shibboleth, has so far prevented a proper integration of this resource into a Web based infrastructure. The solution developed by the TLA to support the CLARIN-D F-AG 3 KP 1 to enable access to this resources for web services via an OAuth bridge while preserving the individual access rights is crucial for a further integration of language documentation data into the CLARIN infrastructure. Especially, this later development has a lot of potential for future work on data from language archives hosted at the MPI in Nijmegen.

The metadata tool CMDI Maker developed as the second curation project CLARIN-D F-AG 3 KP 2 has become a standard tool for the generation of CMDI metadata for data from language documentation projects. The CMDI Maker has been taught in several summer schools and training
workshops for documentary linguists. Since its release in March 2014, the tool has been widely deployed in language documentation projects, internationally.

**Publications**

No Publications
3.4 Working Group 4: Ancient History, Classical Philology, Archaeology

CHAIR
Prof. Dr. Charlotte Schubert
University of Leipzig
Department of history / Chair of ancient history

MEMBERS
During the last year no new members admitted to the working group. Current members are:

- Prof. Dr. Peter Funke, Universität Münster, funkep@uni-muenster.de
- Prof. Dr. Dorothee Gall, Universität Bonn, Dorothee.Gall@uni-bonn.de
- Prof. Dr. Reinhard Fürtsch, Universität Köln, foertsch@mac.com
- Prof. Dr. Ortwin Dally, Deutsches Archäologisches Institut Berlin, generalsekretaer@dainst.de
- Prof. Dr. Markus Deufert, Universität Leipzig, deufert@uni-leipzig.de
- Prof. Dr. Gregor Weber, Universität Augsburg, gregor.weber@phil.uni-augsburg.de
- Prof. Dr. Foteini Kolovou, Universität Leipzig, kolovou@rz.uni-leipzig.de
- Prof. Dr. Eva Cancik-Kirschbaum, Freie Universität Berlin, altorsek@zedat.fu-berlin.de
- Prof. Dr. Kurt Sier, Universität Leipzig, sier@rz.uni-leipzig.de
- Prof. Dr. Tanja Scheer, Universität Göttingen, tanja.scheer@uni-goettingen.de
- Prof. Dr. Hartmut Leppin, Universität Frankfurt am Main, h.leppin@em.uni-frankfurt.de
- Prof. Dr. Sabine Vogt, Universität Bamberg, sabine.vogt@uni-bamberg.de
- Dr. Roxana Kath, Universität Leipzig, roxana.kath@me.com
- Prof. Dr. Christoph Schäfer, Universität Trier, christoph.schaefer@uni-trier.de
- Prof. Dr. Kai Ruffing, Universität Kassel, kai.ruffing@uni-kassel.de
- Gregor Horst kemper, Bayerische Staatsbibliothek, Horstkemper@bsb-muenchen.de
- Dr. des. Andreas Gerstacker, Universität Leipzig, andreas.gerstacker@uni-leipzig.de
- Prof. Dr. Klaus U. Schulz, Universität München, schulz@cis.uni-muenchen.de
- Dipl.-Ing. Maik Preuß, maik.preuss@gmx.de

The working group has one member of staff in part time (TVL E 13; Dr. Michaela Rücker), who
has different duties and responsibilities:

- Communication with the members of the Working Group and the employees of the centre in Leipzig
- Organisation of workshops and meetings between the members of the Working Group
- Identification of potential projects, sources and tools
- Documentation of discussions and results
- Coordination of the working process of the Curation Projects within the Working group
- Dissemination of results and improvement of strategies to engage interested persons from the scientific community.

ACTIVITIES

A) Recent activities

Members of the WG4 participated at / presented in the following events:

- 26./27.06.2014: DFG Roundtable Berlin
- 30.07.2014: Deutsche Digitale Bibliothek, Berlin
- 02.10.2014: Deutsche Digitale Bibliothek, Frankfurt
- 20.10.2014: Unterausschuß Digitale Geschichtswissenschaften, Braunschweig
- 26./27.11.14: Fachinformationsdienst Altertumswissenschaften Munich
- 03.-04.03.2015: Attendance at the DH-Summit 2015 in Berlin (TextGrid and DARIAH-DE)
- 25.-27.03.2015: Attendance at the Herrenhausen Conference: "Big Data in a Transdisciplinary Perspective"

B) Further activities

- 22.06.2014: lecture by Prof. Charlotte Schubert at Max Weber Kolleg Erfurt: “Modern information technology in the ancient studies”
• 25.07.2014: lecture by Prof. Charlotte Schubert at ESU Leipzig: “Quotation and fragment in the age of digitization. Experiences and perspectives in research”


• DHd Graz 2015: a panel with different lectures to projects from the digital humanities and discussion, presentation of two posters

• 16.04.2015: conference of the “Fachreferenten” in ancient studies in Heidelberg (lecture of Prof. Charlotte Schubert)

• 17.04.-20.04.2015: Große Mommsen-Tagung: presentation of two posters

• Contributions to the CLARIN-D website

• Presentation of the current work on scientific conferences

• Further discussions about the handling of digital resources and tools and the problems of the use

• Discussion with the other working groups about the new curation projects and potential cooperation

C) Planned activities

Currently, WG4 is preparing its contribution to the CLARIN-D Disseminationsworkshop from 30.06.2015-01.07.2015 in Leipzig: presentation of the new curation project and the new open access online journal “Digital Classics Online”. We also plan a meeting of the members of our working group to show the work of CLARIN-D and the other working groups and to speak about new projects in the ancient studies.

The communication between the members of the Working Group and the employees of the center in Leipzig provides both disciplines with an understanding of working processes.

The presentation of developed tools and the dissemination of research results help us to strengthen the position in our scientific community and to engage interested persons for further cooperation and for participation in the working group.

Further discussions with the other working groups about potential cooperation.
CURATION PROJECTS

CP 3: “Extension of an open source tool for postcorrection of historical OCR documents”

The applicant of this project is Prof. Dr. Klaus U. Schulz, center for information and language processing, Ludwig-Maximilians-Universität Munich. The project starts at the 1st of April 2015 and will end at the 31st of March in 2016.

The project pursues two directly linked goals:

1. The aim is to develop a system to facilitate and advance interactive postcorrection of historical OCR documents. The system should offer both usability and reliability to be used as a practical tool by different researching facilities.

2. The system should be made available as an open source tool to all academic institutions in the fields of Humanities, libraries and institutions concerned with digital preservation. The long-term goal should be to establish a community around the system to ensure its further development and maintenance.

The tool is available at: http://ocr.cis.uni-muenchen.de

BENEFITS OF CLARIN-D FOR THE SCIENTIFIC COMMUNITY ADDRESSED BY THE WORKING GROUP

One of the advantages that accrued in CLARIN-D is the regular contact between the heads of the other working groups in virtual meetings and on workshops. It is possible to discuss the current work in the WG and especially in the Curation Projects as well as problems with the administration.

Another benefit lies in the constitution of the working group: the members come from different scientific disciplines and different research institutions. So it is possible to discuss problems in the field of digital humanities from various angles and with different perspectives on teaching, research, data backup and so on.

The feedback to our lectures on different conferences showed the relevance of the topic “Digital Humanities” for our scientific community. It is necessary to sustain this discussion.

PUBLICATIONS

None
3.5 Working Group 5: Psycholinguistics and Cognitive Psychology

CHAIR
Prof. Dr. R. Harald Baayen
Department of Linguistics
University of Tübingen

MEMBERS

- Prof. Dr. Ingo Plag, Uni Düsseldorf, Ingo.Plag@uni-duesseldorf.de
- Prof. Dr. Harald Baayen, Uni Tübingen, baayen@ualberta.ca
- Prof. Dr. Barbara Kaup, Uni Tübingen, barbara.kaup@uni-tuebingen.de
- Prof. Dr. Lars Konieczny, Uni Freiburg, lars@cognition.uni-freiburg.de
- Prof. Dr. Pienie Zwitserlood, Uni Münster, zwitser@psy.uni-muenster.de
- Prof. Reinhold Kliegl, Ph.D., Uni Potsdam, kliegl@uni-potsdam.de
- Prof. Dr. Shravan Vasishth, Uni Potsdam, shravan.vasishth@gmail.com
- PD Dr. Erich Weichselgartner, ZPID, wga@zpid.de
- Prof. Dr. Anke Lüdeling, HU Berlin, Anke.Luedeling@rz.hu-berlin.de
- Prof. Dr. Sabine Weinert, Uni Bamberg, sabine.weinert@uni-bamberg.de

ACTIVITIES

The working group has proposed adaptations to corpus linguistics for the Mind Research Repository. Anke Lüdeling (HU Berlin), Harald Baayen (Tübingen) and Ingmar Schuster (Leipzig) collaborated to write a grant proposal in this regard, which was accepted by CLARIN. A lot of administrative hurdles had to be overcome to actually set up the project once it was accepted, mainly because of a very problematic situation where overhead is not granted and an unclear VAT situation. After this was settled, the working group concentrated on improving the Mind Research Repository software (curation project) and filling it with content, i.e. Paper Packages.

We plan on working more on the MRR and heighten its visibility, among other things by creating dissemination material.
CURATION PROJECTS

The Mind Research Repository (MRR) offers access to scientific preprints, accompanying data sets and code for statistical evaluation of the data. The MRR has evolved from the Potsdam Mind Research Repository.

Currently, we adapt the MRR to the needs of Corpus Linguistics. Among other things, this means data and preprints can be referenced by PID to enable external storage. This is useful for very large data sets and as a measure to ensure legal compliance with regard to publication policy of some journals. Also, we took first steps towards additional backups of the data stored at the MRR. Improvements in usability currently are the most important part of the project. Most of the necessary improvements have become apparent by conducting a usability study.

BENEFITS OF CLARIN-D FOR THE SCIENTIFIC COMMUNITY ADDRESSED BY THE WORKING GROUP

We currently use PIDs in the curation project as well as a server of the center in Leipzig. Furthermore, we implemented a Shibboleth Authentication, which however is not used due to open questions regarding usability.

We provide a growing number of packages with scientific data at openscience.uni-leipzig.de in a CLARIN compatible way.

PUBLICATIONS

New Paper Packages on the Mind Research Repository

- Wieling et al. *English accents and their determinants* (September 2014)
- Dietterle et al. *Zur Syntax von Plauderchats* (November 2014)
- Shaoul et al. *N-gram probability effects in a cloze task* (February 2015)
3.6 Working Group 6: Speech and other modalities

CHAIR

Prof. Dr. Petra Wagner

Universität Bielefeld
Fakultät für Linguistik und Literaturwissenschaften
Postfach 100131
33501 Bielefeld
petra.wagner@uni-bielefeld.de

MEMBERS

With the shift in the F-AG chair, there was also a major shift in F-AG topic focus (see below) and WG members. The newly constituted F-AG 6 now comprises 11 members from 7 institutions, working mostly in the area of phonetics, but also coming from applied computational linguistics, dialogue modeling, first language acquisition, speech technology, general linguistics and English linguistics:

• Dr. Kathrin Schweitzer, Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart
• Dr. Felix Burkhard, Mitarbeiter bei der Deutschen Telekom Berlin
• Dr. Benjamin Weiss, Telekom Usability Labs, TU Berlin
• Dr. Angela Grimminger, Work Group Emergent Semantics, CITEC, Universität Bielefeld
• Prof. Dr. Ulrike Gut, Institut für Anglistik, Universität Münster
• Prof. Dr. David Schlangen, Applied Computational Linguistics, Dialogue Systems Group, Universität Bielefeld
• Prof. Dr. Stavros Skopeteas, Allgemeine und vergleichende Sprachwissenschaft, Universität Bielefeld
• PD Dr. Jürgen Trouvain, Institute of Phonetics, Universität des Saarlandes
• Prof. Dr. Bernd Möbius, Institute of Phonetics, Universität des Saarlandes
• Dr. Zofia Malisz, Phonetics and Phonology Work Group, Universität Bielefeld, Institute of Phonetics, Universität des Saarlandes
• Dr. Susanne Fuchs, Zentrum für Allgemeine Sprachwissenschaft (ZAS), Berlin
• Dr. Robert Fuchs (Institut für Anglistik, Universität Münster)
ACTIVITIES

A) Recent activities

F-AG 6 Workshop in Bielefeld (01.12.2014)

Along with the change in the chairperson of F-AG 6 there was also a thematic shift from a strong focus on multimodal corpora to phonetic speech corpora, which may be, but do not have to be supplemented with multimodal data. To perform this shift, a Kick-Off meeting was held at Bielefeld University on 01.12.2014 with 17 participants, all of them experts in gathering and annotating speech corpus data. In course of the meeting, Christoph Draxler gave an overview of the CLARIN-D project purposes and contents, and informed about the infrastructure, existing tools and options for curation projects.

Besides, there was an intensive discussion about (1) aspects where CLARIN-D could be of greatest potential benefit to the community, (2) which aspects are currently causing the most pressing problems in phonetic data collections, (3) where de-facto standards already exist and (4) what kind of curation project or similar CLARIN-D initiative would provide the optimal solution towards addressing the set of problems.

With respect to (1), there was a wide consensus that the F-AG would profit eminently from an altogether better standardization, in order to simplify data exchange and establish a set of best practice guidelines enabling researchers to meet these standards and avoid typical beginner's mistakes. Alongside with a lack of best practice guidelines and missing standards, the group also diagnosed a missing general overview of existing tools suitable for building speech corpora, as many labs tend to build their own solutions, often re-inventing the wheel. For several existing CLARIN-D tools, it was found that they already define a de-facto standard in the community. The sharing and distribution of many other existing tools (PACX, TextGridTools, to mention a few) needs to be optimized.

The currently most serious problem for many researchers dealing with speech or multimodal corpora remain to be the unclarities about legal issues, especially with respect to data sharing and publication. Unfortunately, many old corpora are unsuitable for being made available publicly, as no consensus forms were gathered or stored for these. In those cases, where consent forms are available, anonymization may create a problem, as recorded speakers remain the right of having their data deleted in the future. Support by the legal help desk, alongside with the currently developed DFG guidelines for legal issues in data collections are strongly encouraged and welcomed by the F-AG members.

Despite the fact that the usage of certain tools and procedures have been established as de-facto
standards in the community (most of them already provided by the CLARIN-D initiative such as ELAN and WebMAUS), other tools of high potential remain relatively unknown. Here, a workshop would provide a suitable format to make available tools known more widely, while CLARIN-D would be an excellent platform for their further distribution. However, such a platform alone would not provide the necessary benefit without an improved dissemination of its contents, possibly realised through workshops and conference presentations. Another demand within the community was a citable reference for Best Practice Guidelines when preparing speech corpora. It was noticed that the DFG guidelines could provide such a reference, but may be in need of further specification, e.g. with respect to handling multimodal data or special participant groups (e.g. children, patients).

B) Planned activities

As discussed during the F-AG 6 meeting, it is planned to carry out a workshop on exchanging and evaluating existing tools for dissemination via CLARIN-D. This workshop should ideally co-occur with a prominent national workshop within the field of phonetics, in order to attract a larger audience (beyond the F-AG 6 members). Currently, it is planned to have this workshop together with the P&P conference in Marburg (October 2015). Additionally, it was discussed to have a workshop on Best Practice Guidelines for creating speech and multimodal corpora. This workshop should take place towards the end of the current funding phase, and establish standards among the community that could ideally also lead to a quotable publication extending and/or supplementing the DFG guidelines. That way, the CLARIN-D initiative would automatically become very visible throughout the community, e.g. as it can be explicitly mentioned in any publication's acknowledgments or reference section.

PLANNED CURATION PROJECT

Due to problems in setting up the contract between the CLARIN-D center and Bielefeld University, so far no curation project was proposed by F-AG 6. However, a first idea for a proposal that could still be realized within the current funding phase was put forward. It relates to the curation of data collected at the University of Münster (Prof. Ulrike Gut), dealing with Scottish English speech data. The corpus would be made available via CLARIN-D and annotated using a set of existing CLARIN-D tools.

BENEFITS OF CLARIN-D FOR THE SCIENTIFIC COMMUNITY ADDRESSED BY THE WORKING GROUP

The following benefits provided by CLARIN-D can be currently identified:

A set of standards or quotable best practice guidelines would provide an extremely helpful issue
when planning data collections or publishing on data collections, potentially leading to a boost in corpus quality, by minimizing the creation of data that cannot be openly published or exchanged as they do not meet those standards, unfortunately this is the case for many existing corpora as they comprise recordings carried out without proper consent forms or have been annotated in ways unsuitable for data exchange. Such a set of quotable guidelines would automatically lead to a wider dissemination of CLARIN-D throughout the community, potentially gaining extremely high impact via citations and re-use.

Standards as defined by CLARIN-D can be valuable for third party funding proposals, as they promise that data generated within a project are made available to the scientific community in a currently close-to optimal (or at least suitable) way.

Exchange, dissemination and standardization of tools would create an additional benefit for the scientific community, especially if these tools come along with hands-on workshops (as the ones carried out by the CLARIN-D center at LME Munich), leading to their widespread use and simplifying data collections.

Another potential benefit lies in synergies by exchanging experience between different F-AGs sharing similar problems or challenges.

**PUBLICATIONS**

Due to the newly implemented F-AG 6, there are no CLARIN-D related publications in 2014.
3.7 Working Group 7: Applied Linguistics and Computational Linguistics

CHAIR
Prof. Dr. Anette Frank

Department of Computational Linguistics
Heidelberg University
frank@cl.uni-heidelberg.de

MEMBERS

• All members of F-AG7 curation projects are part of the working group.
• A new member from Curation Project 3 is Silvana Hartmann from Technische Universität Darmstadt.
• Prof. Dr. Beatrix Busse from Heidelberg University is a new group member since March 2015.
• Prof. Dr. Philipp Cimiano from the University of Bielefeld joined the F-AG7. His group will also adapt the extracted linked lexical resources as Linked Open Data.
• Status or affiliation changes of existing members in the working group: Angelika Storrer moved to University of Mannheim and is now holding the chair of German Linguistics.
• On December 1st, 2014, Eva Mujdricza-Maydt was hired to work as an assistant in the F-AG7 working group project at Heidelberg University in CLARIN-D phase 2.

ACTIVITIES

The main activity within the reporting period was to prepare and conduct first steps for the Curation Project 3 “Semantic Annotation for Digital Humanities” (CP3).

A) Recent activities

Finalizing the proposal for Curation Project 3

In coordination with the F-AG7 working group members we finalized the proposal for Curation Project 3, taking into account budget cuts. The F-AG7 project in Heidelberg is contributing resources to the curation project in area B. We could also acquire support from the CLARIN-D Center Leipzig for area A. CP3 will work with a number of cooperation partners and will be aligned with the CLARIN-D Center Tübingen to ensure integration with the CLARIN-D infrastructure. A
further cooperation partner is the University of Bielefeld, who will support the project by making the constructed resources compliant with LOD representation formats. The Curation Project started in March 2015.

**Working Group meeting**

On March 9, 2015, the F-AG7 working group organized a meeting in Heidelberg to discuss further activities and collaboration within and beyond the working group. During the meeting, the new curation project was presented; partners of CP3 presented practical use of WebAnno and related research in Digital Humanities projects.

**Kick off meeting of Curation Project 3 (CP3)**

Subsequently to the working group meeting on March 9, 2015, a kick off meeting for CP3 was organized. First steps as well as theoretical issues and technical aspects were discussed in detail.

**Consortium meeting**

Anette Frank attended the CLARIN-D consortium meeting in Braunschweig as a F-AG7 representative, on March 23, 2015. Anette Frank and Richard Eckart de Castilho presented CP3 with a poster.

**B) Further activities**

The application for curation project 3 was submitted in November 2014 after a revision of financial and personal resources. During the planning phase we could attract cooperation partners for CP3. The partners are going to work on various tasks with the annotation tool WebAnno, thus the F-AG7 can profit from their experience and feedback. On the other hand, the cooperation partners can benefit from the cooperation through direct response to reported needs.

**Working group meetings**

- On October 29, 2014 a virtual meeting was organized to get feedback and support from the whole working group on the planned project CP3.

- On March 9, 2015 an on-site meeting of the F-AG working group was organized in Heidelberg (see above).

**Shared Task Organization**

Anette Frank is member of a committee that oversees shared task proposals to GSCL, DGfS-CL, and since recently CLARIN-D through the F-AG7 Curation Project 3. The committee has set up
criteria for shared tasks that form a prerequisite for funding from the above mentioned community organizations. At Konvens 2014 two shared tasks were conducted in coordination with GSCL. Named Entity Recognition for German non-standard data and Sentiment Tagging. A further shared task on PoS Tagging for IBK data is in preparation, similarly supported by GSCL.

**Working Group leader meetings (monthly)**

Anette Frank is regularly participating in the virtual meetings of the working group leaders. Occasionally she is represented by Eva Mujdricza-Maydt.

**Participation in national and international conferences**

- Anette Frank was program co-chair for the *SEM Conference 2014 in Dublin.
- Anette Frank also attended COLING 2014 in Dublin.
- Dustin Heckmann gave an oral presentation on “Citation Segmentation from Sparse & Noisy Data: An Unsupervised Joint Inference Approach with Markov Logic Networks“ at the DHd Conference “Von Daten zu Erkenntnissen: Digitale Geisteswissenschaften als Mittler zwischen Informationen und Interpretationen”, March 2015 in Graz.
- Nils Reiter and Anette Frank took part in the poster session of the same conference with a poster on “Discovering Structural Similarities in Narratives”.

**C) Planned activities**

In the scope of CP3, we plan to support shared tasks in the field of annotation. Non-standard language varieties of German are particularly interesting and were already subject of CP2. Potential shared tasks on related language varieties are annotation tasks like PoS-Tagging or dependency parsing. The F-AG7 members also raised the analysis of compounds as a topics of particular interest.

We will take part in the CLARIN-D Dissemination Workshop on June 30 – July 1, 2015 in Leipzig where we are going to present CP3 with a poster. We are planning to organize a meeting of the F-AG7 working group following the workshop, if applicable.

A further meeting of the working group is planned in conjunction with the DGfS-CL meeting on February 23-26, 2016 in Konstanz. In the meeting we will present the outcomes of Curation Project 3.
The ongoing work on CP3 is coordinated by Anette Frank; monthly/weekly contact and meetings ensure appropriate cooperation.

**CURATION PROJECTS**

**Curation project 3: “Semantic Annotation for Digital Humanities”: March 1, 2015 – February 29, 2016**

In the first phase of CLARIN-D, two curation projects were conducted: „Implementation of a web-based annotation platform (WebAnno)“ and „Development of guidelines and Best practices for annotation of non-standard varieties of German“. The aim of the new curation project „Semantic Annotation for Digital Humanities“ is to consolidate the successful work of the previous curation projects and to extend them in novel directions. The focus of the new curation project is on semantic annotation for Digital Humanities. It is divided into three work packages:

**A. Consolidation and further development of WebAnno for practical use in DH projects**

In order to provide better support for semantic annotation layers as well as user-defined annotations, *new functionalities* will be made available in WebAnno:

- **Template-based annotations** – filling predefined elements (slots) in predicate-argument structure annotation, or in event annotation;

- **Constraints** – context-based restrictions on target element annotations.

The new functionalities will be implemented in interaction with cooperation partners as active users.

For appropriate dissemination in the community, WebAnno will be integrated into the CLARIN infrastructure and offered as a CLARIN service.

**B. Curation of resources for semantic annotation and further annotation of the NoSta-D corpus**

The aim of work package B is to develop a prototype for linked lexical semantic resources for German (including a LOD representation) and a robust annotation scheme for concepts and predicate-argument structures for annotation of concepts and events in DH projects. Here, the curation project focuses on the following tasks:

1. Linking existing (GermaNet, SALSA) and newly developed (UBY) lexical semantic resources for German following the model of the **Unified Verb Index**.
2. Exploring guidelines and annotation formats for WSD (similar to OntoNotes) and SRL (FrameNet, VerbNet-style). Selected non-standard corpora will be annotated according to these schemas.

C. Supporting Shared-Tasks for German for selected annotation types

Jointly with the national organizations (GSCL, DGfS-CL) we will support shared-task initiatives for various annotation types. The first editions of shared-tasks for Named Entity Recognition (NER) and Sentiment Tagging were successfully conducted during KONVENS 2014. A further task on PoS-Tagging for internet-based communication language data is being supported by GSCL. Possible shared tasks to be supported by the curation project include dependency parsing for non-standard language varieties (building on curation project 2), or the analysis of compounds for German.

Project leaders:

- Prof. Dr. Anette Frank, Institut für Computerlinguistik, Universität Heidelberg (Coordinator)
- Prof. Dr. Chris Biemann, Fachbereich Informatik, Technische Universität Darmstadt
- Dr. Richard Eckart de Castilho, Fachbereich Informatik, Technische Universität Darmstadt
- Prof. Dr. Iryna Gurevych, Fachbereich Informatik, Technische Universität Darmstadt

Project staff:

- Silvana Hartmann
- Eva Mujdricza-Maydt
- Seid Muhie Yimam

Cooperation partners:

- Prof. Dr. Phillip Cimiano, Universität Bielefeld
- Prof. Dr. Stefanie Dipper, Universität Berlin
- Prof. Dr. Gerhard Heyer, Universität Leipzig
- Prof. Dr. Anke Lüdeling, Universität Bochum
- Prof. Bolette Sandford Petersen, Universität Kopenhagen
- Prof. Dr. Angelika Storrer, Universität Mannheim

- CLARIN-D-Zentrum Tübingen (Prof. Dr. Erhard Hinrichs)
- CLARIN-D-Zentrum Hamburg: CLARIN-D Helpdesk


**BENEFITS OF CLARIN-D FOR THE SCIENTIFIC COMMUNITY ADDRESSED BY THE WORKING GROUP**

The main effort of our discipline-specific working group is the development and distribution of a modern and sustainable research framework for natural language processing for German. Both researchers of NLP as well as users within a wider community benefit from the experience we collect with the curation projects on the annotation of non-standard language varieties, and on linking and populating lexical resources. WebAnno offers more and more flexibility and user-friendliness, which properties make it a robust, widely recommendable and sustainable tool. To ensure these aims for WebAnno, we include the feedback from our cooperation partners and other users.

**PUBLICATIONS**


Posters and presentations


3.8 Working Group 8: Content Analysis in the Social Sciences

CHAIR
Prof. Dr. Cathleen Kantner
Abteilung für Internationale Beziehungen und Europäische Integration
Universität Stuttgart
Breitscheidstr. 2
70174 Stuttgart

Prof. Dr. Gary S. Schaal
Fakultät für Wirtschafts- und Sozialwissenschaften
Helmut Schmidt Universität Hamburg
Holstenhofweg 85
22043 Hamburg

MEMBERS
• Prof. Eva Barlösius and PD Dr. Axel Philipps (Institut für Soziologie, Leibniz Universität Hannover)
• Prof. Dr. Andreas Blätte (Institut für Politikwissenschaft, Universität Duisburg-Essen)
• PD Dr. Sebastian Haunss (Zentrum für Sozialpolitik, Universität Bremen)
• Prof. Dr. Jeannette Hofmann (WZB Berlin)
• Bruno Hopp (GESIS, Abteilung Datenarchiv für Sozialwissenschaften Team Akquisition, Sicherung, Datenbereitstellung)
• Dr. Christian Rauh (WZB Berlin)
• Prof. Dr. Bernd Schlipphak (Institut für Politikwissenschaft, Universität Münster)

ACTIVITIES
A) Recent activities
• Gathering and Constitutional Meeting of the WG-8 (November 21st, 2014, Stuttgart)
• CLARIN-D European Summer School "Digital Humanities & Language Resources" (July, 22nd – August 1st, 2014, Leipzig): Participation

• Conference “Political Context Matters: Content Analysis in the Social Sciences” (October, 10th – October 11th, 2014, Mannheim, MZES, Universität Mannheim): Presentation of Prof. Dr. Cathleen Kantner and Maximilian Overbeck, Title: “The practical challenges of exploring “soft” concepts through “hard” methods: The corpus linguistic analysis of multiple collective identities in contemporary transnational media debates“

• Second Annual Conference: Digital Humanities in the German-speaking area (DHd): Workshop together with our Cooperation partners of the consortium project e-Identity: Content: “Computerlinguistische Methoden der Inhaltsanalyse in den Sozialwissenschaften: Forschungspraktische Herausforderungen, Tools und Technologien“ (February 23rd-24th, 2015, Graz)

• Launch of the Curation Project and Implementation of the Content Builder (First Version): http://clarin01.ims.uni-stuttgart.de/ccb/AnalyseCoding

• Clarin-D Meeting – WG and Consortium Meeting, 23.03.2015, GEI, Braunschweig: Presentation of the Curation project

• Participation in the monthly virtual meetings of the WG heads

B) Further activities

• Participation in the monthly virtual meetings of the WG heads

• Clarin WG Dissemination Workshop (June 30th – July 1st, 2015, Leipzig)

• Implementation and Development of the Curation Project

C) Planned activities

• Cooperations with other Working Groups, i.e. Cooperation with the Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW, WG–9 for Contemporary History)

• Dissemination of Methods and Resources
CURATION PROJECTS

Curation Project I:


Applicants:

• Prof. Dr. Andreas Blätte, Juniorprofessur für Politikwissenschaft der Stiftung Zukunft NRW, Universität Duisburg Essen

• Prof. Dr. Gary S. Schaal, Lehrstuhl für Politikwissenschaft, insb. Politische Theorie, Helmut Schmidt Universität, Hamburg

Content:

The focus of this project lies on the classification of plenary debates of the German Parliament. On the basis of an existing corpus of plenary protocols (PolMine) a sample of plenary debates will be manually classified with the taxonomy of political sciences. For that, the coding schema of the Comparative Agenda Project (www.comparativeagendas.info) will be used. With the help of these then generated training records all plenary debates will be classified. The PolMine-Corpus already is part of the German Reference Corpus (DeReKo) and available through the Institute of German Language (IDS). In the course of the project it will be hosted for a circle of users of social and political sciences.

BENEFITS OF CLARIN-D FOR THE SCIENTIFIC COMMUNITY ADDRESSED BY THE WORKING GROUP

The Curation Project of the WG-8 is building a resource for the social sciences and makes it available through CLARIN for the dh-community. Beforehand it will be tested.

PUBLICATIONS

There are no publications, yet.
3.9 Working Group 9: Modern history

CHAIR
Prof. Simone Lässig

Georg Eckert Institute for International Textbook Research
Celler Straße 3
38114 Braunschweig

MEMBERS

- Prof. Martin Baumeister, DHI Rom, Director
- Prof. Marcelo Caruso, HU Berlin
- Esther Chen, Max Planck Institute for the History of Science, Head of Library
- Dr. Stefan Cramme, German Institute for International Educational Research, Head of Research Library
- Prof. Ernesto W. De Luca, Georg Eckert Institute for International Textbook Research, Head of Digital Research Infrastructures
- Prof. Ludwig M. Eichinger, Institut für deutsche Sprache, Director
- Maik Fiedler, Georg Eckert Institute, Research Assistant
- Ursula Flitner, Max Planck Institute for Human Development, Head of Library
- Prof. Gudrun Gersmann, Cologne University
- Prof. Andreas Gestrich, German Historical Institute London, Director
- Prof. Rachel Heuberger, University of Frankfurt
- Prof. Rüdiger Hohls, Humboldt University Berlin
- Gregor Horst kemper, Bavarian State Library, Research Assistant
- Dr. des. Jörg Hörnschemeyer, German Historical Institute Rome
- Michael Kaiser, Max Weber Foundation, Research Assistant
- Maret Keller, Georg Eckert Institute, Coordinator Working Group
- Dr. Mareike König, German Historical Institute Paris, Head of Library
- Prof. Gerhard Lauer, Göttingen University, Director, Göttingen Centre for Digital Humanities
- Prof. Marina Lemaire, Trier University
- Prof. Simone Lässig, Georg Eckert Institute for International Textbook Research, Director
- Dr. Anna Menny, Hamburg University
Thomas Meyer, HU Berlin, Research Fellow
Prof. Gisela Minn, Trier University
Dr. Stefan Müller, Max Weber Foundation, Research Assistant
Dr. Michael Piotrowski, Institute for European History, Head of Digital Humanities
Prof. Sabine Reh, German Institute for International Educational Research, Director
Prof. Miriam Rüürup, Hamburg University
PD Dr Michael Schaich, German Historical Institute London, Deputy Director
Dr. Daniel Schlögl, Institute for Contemporary History
Prof. Helwig Schmidt-Glintzer, Herzog August Library Wolfenbüttel, Director
Dr. Joachim Scholz, German Institute for International Educational Research, Head of Research
Dr. Kerstin Schwedes, Georg Eckert Institute for International Textbook Research
Robert Strötgen, German Institute for International and Security Affairs, Head of Dept.
Dr. Thomas Stäcker, Herzog August Library Wolfenbüttel, Vice-Director
Dr. Heiko Weber, Göttingen Academy of Sciences and Humanities
Dr. Andreas Weiß, Georg Eckert Institute for International Textbook Research
Dr. Jörg Wettlaufer, Göttingen Academy of Sciences and Humanities
Dr. Tobias Wulf, Max Weber Foundation, Social Media Project

ACTIVITIES

A) Recent activities

An informal meeting of the Working Group took place at the DH Conference held from 23 to 27 February 2015 in Graz. Members from the GEI updated those present with information on the curation project. CLARIN-D tools and services were discussed with Torsten Trippel and Christian Thomas. Working Group members (and their associates) from Leipzig, Wolfenbüttel, Braunschweig and Göttingen presented their research in digital humanities (e.g. eAqua, Welt der Kinder) and discussed topics including the history of science and digital methods, lexicography and linked open data.

Several members of the Working Group, including its coordinator Maret Keller, attended the DH Summit and the TextGrid Grand Tour in Berlin (3-5 March), learning about and discussing a range of projects and concepts related to the digital humanities and digital research infrastructures.

A presentation on CLARIN-D was given at a discussion on Research Data Management at the Georg Eckert Institute’s colloquium on theoretical methods and approaches (11 March).
On 23 and 24 March, Prof. Simone Lässig hosted the quarterly CLARIN-D consortium and developers’ meeting at the Georg Eckert Institute in Braunschweig. As the heads of all Working Groups were invited to attend this occasion, the meeting was used to present their activities and to discuss possible cooperation.

The Working Group’s coordinator attended a workshop on text-mining historical corpora (Bochum, 10-12 April) and upon her return discussed the insights gained with the Working Group and curation project members.

The curation project was represented with a poster at an international conference on Blumenbach held in Göttingen on 23 and 24 April.

The curation project and further CLARIN-D services will be showcased at a Blumenbach Online project meeting to be held in May 2015.

B) Further activities

• Participation in the monthly virtual conferences of Working Group heads

• Production of a regular newsletter for Working Group members

C) Planned activities

• Dissemination workshop to be held on 1 July 2015 in Leipzig; a Working Group meeting will take place on this occasion

• January/February 2016: Joint conference of Working Groups 9 and 10

CURATION PROJECTS

The discipline-specific working group 9 on Modern History (German: F-AG 9) is overseeing the curation project entitled “Sources of the New: Factual and scientific knowledge for amateurs and experts from the Enlightenment to Modernism”. In January 2015 the Working Group received approval for its curation project, which launched in February 2015. Prof. Gerhard Lauer (Georg August University Göttingen) is responsible for project content; Maret Keller (Georg August University Göttingen/Georg Eckert Institute, Braunschweig) and Christian Wachter (Georg August University Göttingen) are responsible for implementation. Technical advice and assistance for this project will be provided by the CLARIN-D centre at the Institut für Deutsche Sprache [Institute of the German Language] in Mannheim (IDS) as well as the CLARIN-D centre in the Berlin-Brandenburgische Akademie der Wissenschaften [Berlin-Brandenburg Academy of Sciences and Humanities].
Project content

The project will prepare and investigate a digital corpus comprising textbooks and other texts published by the university scholar Johann Friedrich Blumenbach (1752–1840) and his circle, which will enable historians to analyse the connections and relationships between school teaching and the production and transfer of knowledge in the university context.

The project is progressing in accordance with the planned schedule submitted with the project application and will be completed by 31 January 2016. It was the subject of presentations and feedback sessions at the CLARIN-D consortial meeting in Braunschweig on 23 and 24 March and at the international Blumenbach Conference in Göttingen on 23 and 24 April.

Two websites provide information on the project for a wider audience:


**Benefits of CLARIN-D for the Scientific Community Addressed by the Working Group**

The academic community will benefit from the outcomes of the Working Group’s curation project in a number of ways. First, the curated data will become available through the CLARIN-D centres in Mannheim and Berlin and will be searchable via VLO and FCS. Second, the curation project will provide the community with a use case for the services offered by CLARIN-D and the knowledge required to make use of them. The working group’s personal and virtual communication serves to strengthen existing ties and promote future cooperation.

**Publications**


**Summary of the Work During Reporting Period**

- Contract signed with the CLARIN-D Centre at the IDS, Mannheim, on the management and coordination of the Working Group.
Recruitment process completed for the post of a part-time (0.5 FTE) academic assistant for Working Group coordination tasks (job advertised, interviews held, successful candidate selected and inducted).

Planning and supervision for the curation project entitled “Sources of the New: Factual and scientific knowledge for amateurs and experts from the Enlightenment to Modernism”

- Adjustments made to the project proposal, projected budget and timeframe in consideration of the reviews received from the CLARIN-D centres at the IDS (Mannheim) and the University of Saarbrücken

- Meeting with project collaborators from Göttingen and Braunschweig in Göttingen: 18 July (introductions with designated collaborator Maret Keller) and 5 November 2014 (discussion of issues with R&D contracts)

- Recruitment process completed for a graduate assistant at the University of Göttingen.

- Curation project meeting at the IDS with Peter Fankhauser and Florian Kuhn (9 March 2015)

- Kick-off meeting for curation project (10 March 2015)

- Presentation of curation project and further CLARIN-D services at a Blumenbach Online project meeting (May 2015).

- A presentation on the Working Group was held at a meeting of the Working Group on Digital History at the Historikertag in Göttingen on 25 September 2014.

- Kick-off meeting for the Working Group at the Academy of Sciences Göttingen on 26 September 2014, with a presentation by Dr Alexander Geyken of the CLARIN Centre Berlin.

- Establishment of a newsletter for internal communication on Working Group matters.

- Announcements made on the foundation and aims of the Working Group via reports on the websites of CLARIN-D, the Georg Eckert Institute (GEI), Verband der Historiker und Historikerinnen Deutschlands (VHD) and others; a poster presentation was held at the GEI.

- Working Group representatives took part in the monthly video conferences of Working Group heads.
Attendance at/participation in national and international events:

- Joint European Summer University on Cultures & Technology – CLARIN-D (Leipzig, 22.07-1.8.2014).
- EDIROM Summer School (Paderborn, 8.-11.09.2014).
- THATCamp, Göttingen Centre for Digital Humanities (Göttingen, 22./23.09.2014).
- Working Group meeting at the DhD Conference (Graz, 23.-27.02.2015)
- DH-d Summit and Textgrid Grand Tour (Berlin, 3.-5.03.2015)
- Workshop on text-mining historical corpora (Bochum, 10-12.04.2015)
- International conference on Blumenbach (Göttingen, 23.-24.04.2015): Poster presentation on curation project
- Workshop on semantic web applications (Göttingen 10.03.2015)
- Presentation of CLARIN-D at a discussion of research data management at the GEI’s colloquium on theoretical methods and approaches (Braunschweig, 11.03.2015).
- Hosting of CLARIN-D Consortial Meeting (Braunschweig, 23-24.03.2015)
3.10 Working Group 10: Contemporary history

CHAIR
Prof. Dr. Martin Sabrow
Centre of Contemporary History (ZZF), Potsdam
Am Neuen Markt 1
14467 Potsdam

MEMBERS
The working group has one member of staff in part time (TVL E 13), Thomas Werneke, who has different duties and tasks. Mainly, he is coordinating the working group 10 “Contemporary History”. His tasks has been to organize the constitution of the working group in August 27th, 2014, the preparation and application of a curation project (portal “GDR-press”) and the organization of an experience exchange of working group members and members of other working groups and CLARIN-Centres.

In the future he will coordinate and accompany the realization of the curation project. He will identify and evaluate further digital corpora and computer linguistic tools of the CLARIN infrastructure, which could prove themselves as useful for the work of historians. Another task will be the documentation of the results and their dissemination into the academic community via presentations, lectures and publications. Lastly, he will organize workshops and meetings of the working group.

The Curation Project of working group 10 under the direction of Prof. Dr. Rüdiger Hohls (Humboldt-University) will be executed by Daniel Burckhardt. For details please see below “Curation Project”.

Composition of Working Group 10:

- Prof. Dr. Margrit Pernau, Max-Planck-Institut für Bildungsforschung
- Dr. Sinai Rusinek, Van Leer Jerusalem Institute
- Prof. Dr. Rüdiger Hohls, Humboldt-Universität zu Berlin
- Dr. Anna Veronika Wendland, Herder Institut, Marburg
- Prof. Dr. Jörn Leonhard, Universität Freiburg
- Prof. Dr. Lucian Hölscher, Universität Bochum
• Prof. Dr. Martin Wengeler, Universität Trier
• Dr. Thomas Grotum, Universität Trier
• Dr. Christian Kreuz, Universität Trier
• PD Dr. Ernst Müller, Zentrum für Literatur- und Kulturforschung, Berlin
• Dr. Falko Schmieder, Zentrum für Literatur- und Kulturforschung, Berlin
• Prof. Dr. Heidrun Kämper, Institut für Deutsche Sprache, Mannheim
• Prof. Dr. Dirk van Laak, Universität Gießen
• Prof. Dr. Philipp Sarasin, Universität Zürich
• Prof. Dr. Jan Ifversen, Universität Aarhus
• Prof. Dr. Hagen Schulz-Forberg, Universität Aarhus
• Prof. Dr. Andreas Wirsching, Institut für Zeitgeschichte, München
• Dr. Martina Steber, Institut für Zeitgeschichte, München
• Dr. Daniel Schögl, Institut für Zeitgeschichte, München
• Dr. Jürgen Warmbrunn, Herder Institut, Marburg
• Prof. Dr. Christian Geulen, Universität Koblenz
• Dr. Dirk Bonker, Duke University
• Prof. Dr. Andreas Schulz, Kommission für Geschichte des Parlamentarismus und der politischen Parteien
• Dr. Sven Jüngerkes, Kommission für Geschichte des Parlamentarismus und der politischen Parteien
• Prof. Dr. Willibald Steinmetz, Universität Bielefeld
• Almut Ilsen, Staatsbibliothek, Berlin
• Prof. Dr. Martin Sabrow, Zentrum für Zeithistorische Forschungen, Potsdam
• Dr. Achim Saupe, Zentrum für Zeithistorische Forschungen, Potsdam
• Thomas Werneke, Zentrum für Zeithistorische Forschungen, Potsdam

ACTIVITIES

A) Recent activities

• The constituent meeting of working group 10 “contemporary history” (F-AG 10) took place under direction of Prof. Dr. Martin Sabrow on August 27th, 2014 in Bielefeld. During the meeting the group developed its working agenda.

• On October 10th the working group held a small meeting with colleagues of working group 8 “political sciences”. Dr. Matthias Lemke from the Helmut-Schmidt University in Hamburg was introducing the annotation- and topic-analysis-tool “Leipzig Corpus Miner” of the
project “ePol”.

- Preparation and application of a Curation Project, “GDR-Press”. The supervision of the program lies with Prof. Dr. Rüdiger Hohls of the Humboldt-University. The executing colleague will be Daniel Burckhardt with the support of the CLARIN-Centre BBAW and the Staatsbibliothek Berlin. Homepage CP: http://pressegeschichte.docupedia.de/clarin

- In December 10th, 2014 the F-AG 10 held a meeting at the Humboldt-University together with cooperation partners from the BBAW and the Staatsbibliothek Berlin. In that meeting F-AG 10 was introducing and discussing the curation project. Furthermore Dr. Christian Kreuz from the University of Trier gave a presentation on his corpus linguistic work with the annotation tool “ingwer”. Finally the F-AG 10 members discussed a possible major workshop together with F-AG 9 in early 2016.

- The coordinator of the F-AG 10 participated at the monthly virtual meetings of the F-AG directors.

- There has been a constant exchange of experience and ideas on a regular basis of meetings between the CLARIN-D centre of BBAW and members of the F-AG 10.

**B) Further activities**

- On July, 07th 2014, F-AG 10 coordinator Thomas Werneke participated at the fourth workshop of the “Deutsches Textarchiv” (DTA). The workshop “Aufbau historischer Sprachressourcen” took place in Berlin at the Berlin Brandenburg Academy of Science (BBAW).

- On August 29th, 2014 F-AG 10 coordinator Thomas Werneke gave a small presentation and take part at a panel discussion during the “17th International Conference on the History of Concepts” in Bielefeld. The panel discussions theme was “Historical Semantics meets Digital Humanities”.


- In the spirit of the dissemination of CLARIN-D infrastructure, F-AG 10 coordinator Thomas Werneke in cooperation with Kay-Michael Würzner from the CLARIN-centre BBAW gave several interdisciplinary presentations. On December 4th 2014 they gave a lecture “Digital corpora in contemporary history” at the Staatsbibliothek, Berlin. And on January 27th, 2015 they gave a lecture at the “Commission on the History of Parliamentarianism in Germany”
(Kgparl) in Berlin. On April 13th and 14th there will be another small presentation of the curation project at a “science slam” of the Staatsbibliothek Berlin.

C) Planned activities

- In May 2015 the curation project starts and will be executed by Daniel Burckhardt of the Humboldt-University. The F-AG 10 coordinator will support the work and execution.

- It is planned to create a tutorial for possible application scenarios of CLARIN-D tools (mainly the DTA and DWDS) to support the work of historians in Contemporary History. The tutorial will be established together with colleagues from the BBAW.

- On June 30th and July 1st 2015 the F-AG 10 will present its curation project at the Dissemination workshop of CLARIN-D in Leipzig.

- In the summer semester of 2015 Prof. Dr. Rüdiger Hohls will give a course at the Humboldt-University which focuses on an introduction into Digital Humanities. In this context it is planned to present the CLARIN-D infrastructure and its potential usability for students in the subject of contemporary history. F-AG 10 coordinator Thomas Werneke will be a guest speaker in one of the session.

- Beyond that Thomas Werneke will be a guest speaker at the colloquium of PhD students of the Centre of Contemporary History in Potsdam during the winter term 2015/16. There he will present the CLARIN-D infrastructure.

- On basis of the results of the curation project it is planned to publish an article on press language in the GDR and the benefits of distant reading methodology, corpus linguistic analysis and digital language tools. This is part of the task to evaluate resources of the CLARIN-D infrastructure.

- Together with F-AG 9 the F-AG 10 will prepare a major workshop on CLARIN-D and the historical science. Such a workshop is scheduled for early 2016.

CURATION PROJECTS

The portal „GDR Press“ is the result of a DFG funded project. The project itself is a cooperation between the Centre for Contemporary History (ZZF) in Potsdam and the Staatsbibliothek Berlin together with the Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS) as the supporting service institution. The aim of the project was a full text digitalization and online presentation of three major daily newspapers of the German Democratic Republic, including the “Neues Deutschland”, the “Berliner Zeitung” and the “Neue Zeit”. The internet portal of the
Staatsbibliothek Berlin now provides around four million articles of over 40 years of GDR-press reaching from the postwar era to the early years after the fall of the Berlin wall. The resources can be accessed freely after a small registration.

The planned curation project of the F-AG 10 „Contemporary History“ will implement the GDR-press corpus into the infrastructure of the CLARIN-D resource “Digital Dictionary of the German Language” (DWDS) with the help of the CLARIN-D centre BBAW. For this purpose the existing files in the format of METS/ALTO are to be converted into the DTABf format, a TEI-P5 profile for historic corpora of printed documents. After the automatic conversion from the DTABf format into the TCF format the texts can be directly implemented into the analytic systems (e.g. the Weblicht-Toolchain) of CLARIN-D. The project will be supervised by Prof. Dr. Rüdiger Hohls of Humboldt-University and will be executed by his colleague, Daniel Burckhardt, with the support of the coordinator of the F-AG 10.

The aim of the exemplary integration is an examination, to which extent the research in Historical Semantics, like in projects of the Historical Semantics of politics in the 20th century at the Centre for Contemporary History (ZZF) can receive new impulses in methodology, modes of operation and source analysis, when working with digital sources, tools and corpora. To reach that goal the curation project will evaluate the needs and requirements as well as the conception and possible enhancement of CLARIN-D analysis tools and infrastructure. It will hopefully also mark a first step to a methodological reflection.

**Focus of interest/overarching questions**

The curation project pursues also a historical aim. The tools and resources of CLARIN-D will be used to focus on the last years of the GDR and possible signs of the collapse of the SED dictatorship in the language of the GDR press. Therefore we have collected four main considerations, concerning the last years of the GDR.

1. **Changes of value in the language:**

Can we validate the hypothesis, that the Socialism in the GDR gets into the defensive during the 1970ies, by analysing the change of language in the GDR press. And if we can observe such a change, how does it develop? Example: the vanishing of the ideological peace paradigm, trends of militarization and the rediscovery of the Prussians in history politics.

2. **The end of utopias:**

What happened to the “Year 2000” and other ideological drafts describing the future of Socialism. How have future and time concepts developed since the 1970ies.
3. Transfer of language:

As a third approach the project also wants to identify possible transfers of language from the west to the east of Germany and vice versa. There are mutual processes of adaptation in the focus. At what time distinct new concepts like “weiße Flecken” and “Bürgerähe” are emerging? When are they vanishing again, and why (e.g. the concept “Westblock”)? Background to this transfer stands a hypothesis, that the change of 1989 was also foreshadowed by a change and transfer of language in the GDR.

4. Language as a political agent:

The particular linguistic strategies of the GDR Agitprop language will also be part of an analysis of the GDR press. This includes euphemisms (“unzerbrüchliche Freundschaft mit der Sowjetunion”), obloquies (“Boykottheitze”, “Abweichler”, “Bummelanten”), and taboos. Taboos means a conscious avoidance of particular topics and concepts (e.g. the term “Western Union”).

**Benefits of CLARIN-D for the scientific community addressed by the working group**

Since the first months of the F-AG 10 just passed, it is still to be evaluated in the ongoing work of the group, what might be valuable benefits of CLARIN-D for the scientific community of historians. The main hope lies in the adaptation and implementation of computer linguistic tools into the regular work of historians. We also hope to find a “small solution” for the implementation and conversion of already existing digital corpora (with a high relevance for the historical sciences) into the CLARIN-D infrastructure. A solution which simplifies the procedures beyond the curation projects.

**Publications**

There are no publications, yet.