



# CLARIN-D

Newsletter

---

Nummer 1, 2011, November

---

## Editorial

---

### Erster CLARIN-D-Newsletter

*Noch gibt es keinen geheimen Handschlag, noch reden wir uns nicht mit PIDs an und noch muss man sich auf den diversen CLARIN-Seiten im Netz mit jeweils eigenem Login anmelden – aber CLARIN-D hat seine Arbeit aufgenommen und bereits die ersten Projektmeilensteine gestemmt. Ein solcher Meilenstein ist die erste Ausgabe dieses Newsletters ...*

---

Herzlich willkommen also zur ersten Ausgabe des CLARIN-D-Newsletters! Dieser Newsletter wird in sechsmona-

tigem Abstand über das CLARIN-D-Projekt berichten, Projektpartner und Facharbeitsgruppen vorstellen, Artikel zu CLARIN-Schwerpunktthemen präsentieren sowie Platz für Gastautoren bieten. Auf diese Weise können wir in CLARIN-D und nach außen über das Projekt informieren.

In dieser ersten Ausgabe stellt der wissenschaftliche Koordinator Prof. Erhard Hinrichs CLARIN-D vor. Aus Leipzig kommt ein Bericht über das Kickoff-Treffen der Facharbeitsgruppen. Die Facharbeitsgruppen sind ein Merkmal, das CLARIN-D von vielen anderen Projekten unterscheidet, da sie eine breite wissenschaftliche Öffentlichkeit einbeziehen. Damit bieten diese Facharbeits-

Alles weitere unter:

[www.clarin-d.org](http://www.clarin-d.org)

gruppen die einmalige Chance, über Kurationsprojekte den CLARIN-D-Gedanken in andere Anwendungsbereiche zu tragen. Die Rubrik „Ein CLARIN-Center stellt sich vor“ wird in jeder Ausgabe des Newsletters erscheinen. Dieses Mal stellt sich das Bayerische Archiv für Sprachsignale mit seinen Sprachkorpora, Softwareentwicklungen sowie seinem CLARIN-D-Projektteam vor. Den Abschluss des Newsletters bildet ein Kurzbericht vom ersten CLARIN-D-Tutorial, das in CLARINs „most western enclave“ Nijmegen stattfand und auf dem in die großen Projektaufgaben Repositories, Authentifizierung sowie Metadaten eingeführt wurde.

Ein Newsletter lebt von seinen Beiträgen. Wir werden daher die CLARIN-D-Partner einerseits gezielt um Beiträge zu

spezifischen Themen bitten, andererseits ermuntern wir Sie zur Eigeninitiative: Schicken Sie uns Ihren Beitrag zu einem projektrelevanten Thema, Photos von offiziellen (und informellen) Projekttreffen, Mitteilungen über personelle Veränderungen, Kurzporträts von Mitarbeitern!

Wir wünschen Ihnen viel Vergnügen beim Lesen des Newsletters!



**Christoph Draxler, Fabian Bross**

**V. i. S. d. P./Impressum:**

Christoph Draxler  
Ludwig-Maximilians-Universität München  
Institut für Phonetik und Sprachverarbeitung  
Schellingstr. 3  
80799 München

Telefon: +49 (0) 89 / 2180 - 2807  
E-Mail: [newsletter@phonetik.uni-muenchen.de](mailto:newsletter@phonetik.uni-muenchen.de)

Für die Inhalte der Artikel sind die jeweiligen Autoren verantwortlich.

Webpräsenz des europäischen Langzeitprojekts:

**[www.clarin.eu](http://www.clarin.eu)**

# CLARIN-Deutschland stellt sich vor

---

## Infrastruktur eines europäischen Langzeitvorhabens

*CLARIN-D ist die deutsche Sektion des europaweiten Langzeitvorhabens CLARIN zur Schaffung einer technischen Forschungsinfrastruktur für die Geistes- und Sozialwissenschaften.*

---

Die enormen Fortschritte in der Informations- und Kommunikationstechnologie wirken sich zunehmend auf die Rahmenbedingungen und die Möglichkeiten des Forschens und Lehrens in fast allen wissenschaftlichen Disziplinen aus. Einzelforschern und Forschungsverbänden können große digitale Datenmengen mit entsprechenden Analysemethoden und -werkzeugen über Internetportale und digitale Archive zur Verfügung gestellt werden. Auf der Basis solcher virtueller Forschungsinfrastrukturen mit integrierten und interoperablen Ressourcen und Diensten lassen sich traditionelle Fragestellungen durch eine breitere Datengrundlage nicht nur empirisch besser absichern, sondern es können gänzlich neue Forschungsparadigmen entwickelt und neuartige wissenschaftliche Erkenntnisse erzielt werden. Mit den Begriffen *e-Science* und *e-Research* haben

[1] <http://www.wissenschaftsrat.de/index.php?id=345>

sich bereits Fachtermini für diese neuen Formen des wissenschaftlichen Arbeitens auch über den angelsächsischen Sprachraum hinaus etabliert.

### VERBESSERUNGEN

In seiner Pressemitteilung vom 31.1.2011 hat der Wissenschaftsrat auf die Bedeutung von virtuellen Forschungsinfrastrukturen besonders für die Geistes- und Sozialwissenschaften hingewiesen [1]. Diese Empfehlung des Wissenschaftsrats schließt sich den Entscheidungen des *European Strategy Forums on Research Infrastructures* (ESFRI) an, das in seine europäische Roadmap für Forschungsinfrastrukturen auch fünf geistes- und sozialwissenschaftliche Projekte aufgenommen hat. Auf der Grundlage dieser Empfehlungen und Entscheidungen fördert das Bundesministerium für Bildung und Forschung seit dem 1.Mai/1. Juni 2011 das Verbundprojekt „CLARIN-D: Web- und zentrenbasierte Forschungsinfrastruktur für Geistes- und Sozialwissenschaftler“ für einen Zeitraum von zunächst drei Jahren. Die CLARIN-D-Infrastruktur soll zur nachhaltigen Verbesserung der Bedingungen für die Forschung in den geistes-, sozial- und kulturwissenschaftlichen Disziplinen führen. Der Nutzen für die adressierten Forschungsbereiche konkretisiert sich durch die folgenden geplanten Maßnahmen und Dienste:

- Kuration und Verfügbarmachung relevanter Text- und Sprachressourcen über das Internet
- die Bereitstellung möglichst webbasierter Werkzeuge und Technologien, die von Forschern benötigt werden, um kollaborativ, disziplinübergreifend und institutionell unabhängig interagieren zu können
- größtmögliche Unterstützung von Forschern und ihren Forschungsprozessen aus allen geistes-, sozial- und kulturwissenschaftlichen Disziplinen bei zunehmend komplexer werdenden Arbeitsabläufen
- einen möglichst freien und ungehinderten Zugriff auf Ressourcen (Daten, Dienste) und auf eine technische Forschungsinfrastruktur (lokal, national, europäisch)
- Ermöglichung einer im internationalen Wettbewerb erfolgreichen Forschung in einem attraktiven Umfeld für die besten Forscher

## ZUSAMMENSETZUNG

Das CLARIN-D-Konsortium besteht aus den CLARIN-D-Zentren an der Berlin-Brandenburgischen Akademie der Wissenschaften, am Institut für Deutsche Sprache, am Max-Planck-Institut für Psycholinguistik in Nijmegen sowie an den Universitäten Hamburg, Leipzig, München, des Saarlandes, Stuttgart und Tübingen (Projektkoordinator). Diese Zentren erhalten neben der BMBF-Förderung flankierende projektbezogene Zusatzmittel von ihren jeweiligen Landesministerien, von ihren Universitäten,

von der Max-Planck-Gesellschaft bzw. von ihrer Akademie.

Der Aufbau der CLARIN-D-Forschungsinfrastruktur soll eng mit den Fachwissenschaften abgestimmt werden, damit sie diejenigen Ressourcen und Dienste zur Verfügung stellt, die für die beteiligten Wissenschaften höchste Priorität haben. Um dies sicherzustellen, sind am CLARIN-D-Projekt Arbeitsgruppen aus sieben geisteswissenschaftlichen Fachdisziplinen prototypisch beteiligt, die von international renommierten Fachwissenschaftlerinnen und Fachwissenschaftlern geleitet werden.

Weiterhin wird eng mit drei Rechenzentren (Forschungszentrum Jülich, Rechenzentrum Garching, Gesellschaft für wissenschaftliche Datenverarbeitung (GWDG) Göttingen) kooperiert, um die Dienste und das Fachwissen von Großrechenzentren beim Aufbau der technischen Infrastruktur gezielt nutzen zu können.

Ansprechpartner und wissenschaftlicher Koordinator des CLARIN-D Verbunds: Prof. Dr. Erhard W. Hinrichs, Seminar für Sprachwissenschaft, Eberhard-Karls-Universität Tübingen, Wilhelmstr. 19, 72074 Tübingen.

Weitere Informationen zum Projekt: [www.clarin-d.de](http://www.clarin-d.de)



**Erhard Hinrichs**  
*Seminar für Sprachwissenschaft, Universität Tübingen*

# Kickoff-Treffen der Facharbeitsgruppen

---

## Ein Bericht der Facharbeitsgruppen aus Leipzig

*„Es ist nicht genug, zu wissen, man muss auch anwenden.“*

*Es ist nicht genug, zu wollen, man muss auch tun.“*

Johann Wolfgang von Goethe, Wilhelm Meisters Wanderjahre

---

Am Freitag den 23.07.2011 fand in Leipzig das Kickoff-Treffen der Facharbeitsgruppen (F-AG) von CLARIN-D statt. Neben allen Leitern der Facharbeitsgruppen waren auch zahlreiche Gäste aus den CLARIN-D-Zentren, sowie Vertreter aus thematisch verwandten Projekten wie DoBeS (Dokumentation bedrohter Sprachen) [1] und CLARIN-NL [2] angereist.

Inhaltlich sollte den Gästen zunächst aufgezeigt werden, welche Ziele im CLARIN-D-Projekt mittel- und langfristig verfolgt werden, welche Schlüsselrolle die Facharbeitsgruppen dabei einnehmen und wie die gesteckten Ziele erreicht werden sollen. Prof. Hinrichs verdeutlichte in kurzer und prägnanter Form noch einmal die CLARIN-D-Vision und verortete die F-AGs und deren

Arbeit in diesem Kontext. Aufbauend auf diese grundlegenden Informationen schloss sich ein Vortrag Prof. Heyers an, welcher zunächst die Struktur und Organisation im Arbeitspaket der Facharbeitsgruppen (CLARIN-D AP4) thematisierte. Anschließend wurde der zu diesem Zeitpunkt aktuelle Stand der Planungen hinsichtlich der Erstellung eines Integrationsplanes für die Kurationsprojekte der F-AGs vorgestellt und die weitere Zeitplanung besprochen.

In der zweiten Phase des Workshops wurden drei Projekte vorgestellt, welche exemplarisch für die in CLARIN-D geplanten Kurationsprojekte stehen können. Ziel dieser Vorträge war die Vermittlung des Charakters solcher Projekte und deren Funktion als „Botschafter“ für bzw. in die jeweilige Geisteswissenschaft bzw. Fachcommunity.

Bei dem ersten Projekt handelte sich um DoBeS, welches die Dokumentation vom Aussterben bedrohter Sprachen zum Gegenstand hat. Im DoBeS-Projekt wurde eine Infrastruktur geschaffen, mit Hilfe derer verschiedene Forscherteams Daten aus der Feldforschung zu bedrohten Sprachen sicher archivieren und allgemein zugänglich machen können. Im Rahmen von DoBeS ergaben sich dabei intensive Diskussionen über den Arbeitsablauf in der Sprachdokumentation sowie über notwendige gemeinsame Standards z.B. bei der Er-

[1] <http://www.mpi.nl/DOBES>

[2] <http://www.clarin.nl>

zeugung von Metadaten. Probleme und Diskussionen welche selbstverständlich auch Gegenstand von CLARIN-D sind und sein werden.

In einer zweiten Präsentation wurde jeweils ein Kurations- und ein Demonstrator-Projekt vorgestellt, welche Teil des niederländischen Pendant CLARIN-NL sind. Beide Projekte sind auf dem Gebiet der Gebärdensprache zu verorten. Im Projekt SignLinC [3] wurden lexikalische Datenbanken des niederländischen „Gebarencentrums“ [4] und annotierte Zeichensprache-Corpora [5] durch Kombination und Erweiterung der Fähigkeiten von ELAN [6], LEXUS [7] und ANNEX [8] miteinander verknüpft und zugänglich gemacht. Das zweite Projekt, IPROSLA, hatte die Integration und Publikation von Informationen zur Erstellung von Ressourcen zu Gebärdensprache zum Gegenstand. Onno Crasborn verdeutlichte in seiner Präsentation, wie schon durch Einsatz relativ kleiner Mittel interessante Projekte umgesetzt werden können, die von hohem Nutzen für eine bestimmte Fachcommunity sind und dementsprechend auf große Resonanz stoßen.

Der dritte und anteilig größte Block der Veranstaltung befasste sich mit den einzelnen Facharbeitsgruppen von CLARIN-D. Ziel war es, einen ersten Überblick über den Zustand der digitalen Ressourcen und einen Eindruck von der Verbreitung und Verwendung in den einzelnen Fachgebieten zu bekommen. Ein weiteres Ziel war der Austausch über die geplante Zusammensetzung der Arbeitsgruppen und möglicher Stra-

tegien zur Gewinnung von geeigneten Mitgliedern. In durchweg engagierten Vorträgen wurde sowohl die Vielfalt der durch die Facharbeitsgruppen adressierten Fachbereiche als auch das Potential von Kurationsprojekten als Botschafter des Infrastrukturprojektes CLARIN-D deutlich.

Im Rahmen aller Veranstaltungsblöcke wurde Zeit für Fragen und Diskussionen reserviert. In einer sehr offenen und anregenden Atmosphäre wurde unter anderem über das weitere Vorgehen in den Facharbeitsgruppen diskutiert. Dabei kristallisierte sich heraus, dass statt einem im Vorhinein festgelegten und zu Beginn der Veranstaltung vorgeschlagenen Verfahren zur Findung von Kurationsprojekten ein individuelles Vorgehen in jeder F-AG benötigt wird. Besonders bezüglich der Zeitplanung zeigten sich Unterschiede in den Bedürfnissen der Facharbeitsgruppen was die Konstituierung aber auch das geplante weitere Vorgehen betrifft. Die Beantwortung starrer Fragenkataloge sollte ebenfalls zu Gunsten eines individuelleren Vorgehens vermieden werden.

Diese und weitere Empfehlungen der F-AG-Leiter wurden inzwischen in die weiteren Planungen des Arbeitspaketes übernommen. Geplant ist nun ein sowohl zeitlich als auch inhaltlich individuelles Vorgehen in den einzelnen Facharbeitsgruppen. Empfehlungen zu den Kurationsprojekten sollen in Form von Vorhabensbeschreibungen durch die F-AGs erarbeitet werden. Eine Empfehlung zur Struktur dieser Vorhabensbeschreibung soll dabei für eine grundlegende

[3] <http://www.ru.nl/sign-lang/projects/signlinc>

[4] <http://www.gebarencentrum.nl>

[5] <http://www.ru.nl/corpusngtuk>

[6] <http://www.lat-mpi.eu/tools/elan>

[7] <http://www.lat-mpi.eu/tools/lexus>

[8] <http://www.lat-mpi.eu/tools/annex>

Vergleichbarkeit zwischen den einzelnen Vorschlägen sorgen, erlaubt jedoch auch individuelle Abweichungen und Ergänzungen. Ein zusätzlich bereitgestellter Fragenkatalog kann als Leitfaden sowohl in fachlichen als auch in technischen und lizenzrechtlichen Belangen herangezogen werden. Dieses Vorgehen fand die Zustimmung des Konsortiums-kreises und wird in den nächsten Wochen mit den Planungen des Lenkungskreises und der einzelnen F-AG-Leitern abgestimmt.

Ein weiteres wichtiges Ergebnis des ersten Treffens der Leiter aller Facharbeitsgruppen stellt die Wahl von Prof. Dr. Thomas Gloning [9] zum Vertreter der Facharbeitsgruppen dar. Herr Prof. Dr. Gloning wird daher von August an für die Dauer von einem Jahr die Interessen der Facharbeitsgruppen im Lenkungskreis von CLARIN-D vertreten.

Die Folien aller Vorträge können auf den Seiten [10] der ASV heruntergeladen werden und sind selbstverständlich auch auf der frisch erstellten Webseite [11] des CLARIN-D-Projektes verlinkt. Im Rahmen der Nachbereitung des Treffens angefordertes Informationsmaterial

zu CLARIN-D kann in Form eines Foliensatzes ebenfalls auf der CLARIN-D-Projektwebseite bezogen werden.

Das Leipziger CLARIN-D-Team möchte sich an dieser Stelle noch einmal recht herzlich bei allen Teilnehmern für die engagierte Beteiligung, die gehaltenen Vorträge und die aufschlussreichen Anregungen und Diskussionen bedanken. Wir sind überzeugt, dass der im Rahmen des Treffens spürbare Geist der Kooperation und konstruktiven Zusammenarbeit auch in Zukunft in CLARIN-D und insbesondere in den Facharbeitsgruppen Bestand hat und freuen uns auf die weitere Zusammenarbeit.



**Volker Boehlke**  
*Institut für Informatik, Universität Leipzig*

Weitere Infos auch im Wiki:

<http://de.clarin.eu/mwiki>

[9] <http://www.zmi.uni-giessen.de/home/profil-tgloning.html>

[10] <http://asv.informatik.uni-leipzig.de/events/13>

[11] <http://de.clarin.eu>

# Ein CLARIN-Center stellt sich vor: das Bayerische Archiv für Sprachsignale

---

## Von Visionen, Technikern und Spracharchiven

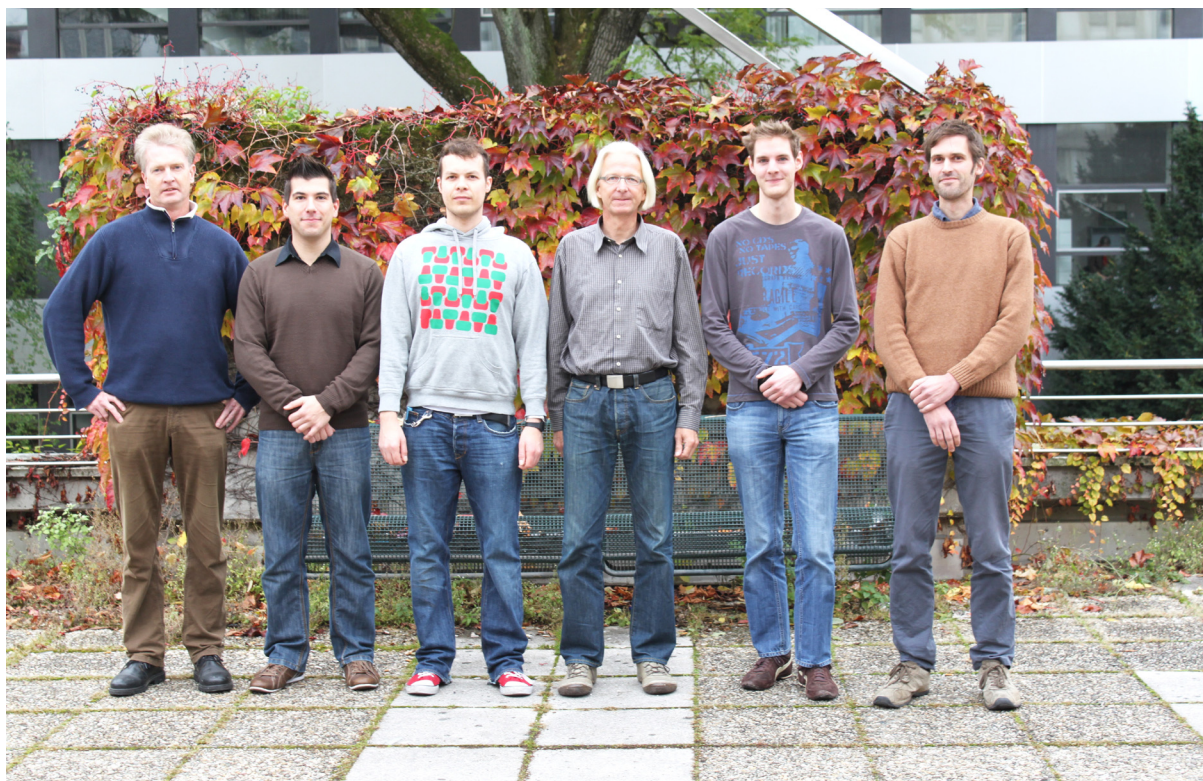
*Das ‚Bayerische Archiv für Sprachsignale‘ entstand 1996 aus einer Idee heraus: Die Verbindung von Geisteswissenschaft und Technik. Heute stellt es eines der größten Spracharchive für gesprochenes Deutsch dar.*

---

Anfang der Neunziger des letzten Jahrhunderts hatte Hans Tillmann, der damalige Direktor des Instituts für Phonetik in München, eine Vision: Er war davon überzeugt, dass die gerade aufkeimende Sprachtechnologie nur durch die Zusammenarbeit von geisteswissenschaftlichen und technischen Spezialisten zu tragfähigen Ergebnissen führen würde. Aber wie sollte man Ingenieure und Geisteswissenschaftler zusammenbringen? Er wusste, dass sich die Kooperation solcher unterschiedlicher Wissenschaftskulturen nicht ‚von oben‘ verordnen ließ. Was also könnte das verbindende Glied zwischen Technik und Grundlagenwissenschaft sein? Tillmanns Antwort war: die Sprachdaten selber, um die es ja beiden Seiten ging, das Rohmaterial, die empirische Grundlage von Wissenschaft und

technischer Anwendung. Hans Tillmann reiste nach Berkely, California, und besuchte einen jungen Post-Doc namens Florian Schiel, den er im Jahr zuvor in Banff auf der ICSLP kennengelernt hatte. Er überredete ihn, nach Beendigung seines Postdoc-Programms wieder nach München an sein Institut zu kommen und dort ein Zentrum für Sprach-Ressourcen aufzubauen. 1996 wurde das ‚Bayerische Archiv für Sprachsignale‘ (BAS) mit einer Anschubfinanzierung des Staates Bayern für 5 Jahre offiziell aus der Taufe gehoben. Die ersten Korpora des BAS stammten aus den BMBF-Projekten PhonDat und ASL. Aber bald sprach es sich herum, dass in München nun ein Archiv existierte, das Sprachkorpora aufnahm, deren Urheber-Projekte abgelaufen waren, und die andernfalls – wie so oft – verloren gegangen wären. Gleichzeitig wurde die Industrie aufmerksam und beauftragte das BAS mit der Produktion von neuen Sprachkorpora. Mit dem Aufkommen des WWW Mitte der Neunziger entstand somit der BAS-Katalog, der bis heute alle in BAS erhältlichen Ressourcen auflistet [1]. FTP-Server für frei verfügbare symbolische Daten wurden aufgesetzt, während die Signaldaten selber auf CD-R und später auf DVD-R und Hard Discs vertrieben wurden.

[1] [www.bas.uni-muenchen.de/Bas](http://www.bas.uni-muenchen.de/Bas)



CLARIN-D-Team des BAS München: v.l.n.r.: F. Schiel, T. Kiesler, F. Bross, C. Draxler, B. Jackl, U. Reichel

## WEITERENTWICKLUNG

Parallel zur Archivarbeit forschen Mitglieder des BAS auch in den technischen Bereichen Spracherkennung, Sprachsynthese und Sprecherverifikation. Für Experimentalarbeiten im Bereich Spracherkennung unterhält das BAS ein Spracherkennungssystem für spontanes Deutsch (auf der Basis von HTK), für Synthese wurde am BAS aus den selbst produzierten BITS-Korpora 4 deutsche Synthese-Systeme entwickelt (in Kooperation mit Marc Schröder, DFKI), und für Experimente im Bereich Sprecherverifikation verfügt das BAS über ein experimentelles Demo-System zur Zugangskontrolle.

In den letzten 10 Jahren hat sich das BAS vermehrt auch in der Entwicklung von Werkzeugen engagiert. Christoph Draxler entwickelte die universelle Aufnahme-Software SpeechRecorder,

die web-basierte Annotationsplattform WebTranscribe sowie PERCY, ein System zur Durchführung von Perzeptionsexperimenten über das Internet. Florian Schiel schuf in den späten Neunzigern die theoretischen Grundlagen für die hypothesen-basierte, statistische Segmentation von Sprache und implementierte auf dieser Basis, das MAUS-System. Uwe Reichel ist verantwortlich für das deutsche text-to-phonem-System BALLOON. Zur Zeit werden diese standalone Werkzeuge nach und nach in das integrierte Korpus-Produktions- und Analyse-Paket WikiSpeech eingebunden, welches den gesamten Prozess der Korpuserhebung, -annotation und -analyse abbildet und unterstützt.

## ZAHLEN UND FAKTEN

Getreu seines Auftrags, technisch und wissenschaftlich verwertbare Sprach-

## Mitmachen!

Liebe Leser des CLARIN-D-Newsletters, wenn ihr Ideen für einen kurzen Beitrag zu diesem Newsletter habt oder dringend einen Gedanken loswerden wollt, schickt euren kurzen Artikel samt Bild an [newsletter@phonetik.uni-muenchen.de](mailto:newsletter@phonetik.uni-muenchen.de). Hinweise zur Beitragsgestaltung findet ihr im Wiki.

Ressourcen auf unbegrenzte Zeit zu archivieren, weiterzuentwickeln und allen interessierten Anwendern zur Verfügung zu stellen, hat sich das BAS bis heute zu einem der größten Sprach-Archive der deutschen gesprochenen Sprache entwickelt. Derzeit werden 34 Sprachkorpora und 5 multi-modale Korpora vom BAS betreut. Von diesen stammen 28 Korpora aus eigener Produktion, 11 wurden von anderen Einrichtung an das BAS übergeben. Seit seinem Bestehen hat das BAS insgesamt 826 Sprachkorpora auf 2.984 Datenträgern an die Wissenschaftsgemeinde und Industrie distribuiert. Auf seinen Servern stehen zu jedem Korpus Dokumentation, Annotationen und Metadaten (IMDI) frei zur Verfügung. Desgleichen können die meisten vom BAS entwickelten Tools zur unbeschränkten Nutzung heruntergeladen werden. Das BAS bietet seit 15 Jahren kontinuierlichen Support per Telefon oder Email. Im Durchschnitt beantworten die Experten des BAS im Jahr

25-35 Anfragen aus Wissenschaft und Industrie. Häufig erstrecken sich solche Konsultationen über mehrere Tage und münden auch oft in konkreten Kooperationen. Mit der ‚BAS Services Schiel‘ steht dem BAS seit 2004 auch ein privatwirtschaftliches Spinoff-Unternehmen zur Abwicklung von Industrie-Aufträgen zur Verfügung [2].



**Florian Schiel**  
*Institut für Phonetik und Sprachverarbeitung, LMU München*

[2] [www.bas-services.de](http://www.bas-services.de)

# First CLARIN-D tutorial

---

## Impressions of the first CLARIN-D tutorial

*The very first CLARIN-D tutorial brought together about 25 technical delegates between 7 en 9 September in CLARIN-D's most western enclave, Nijmegen.*

The first day was completely devoted to repository systems, with a general introduction by Peter Wittenburg and Saskia Woutersen-Windhouwer, who also focused on the social context that is needed to promote the preservation and publication of research data. Then 4 specific repositories (D-space, Fedora, SLDR, Imdil/Lamus) were presented and discussed.

The day afterwards opened with an introduction on setting up a Shibboleth Service Provider. All participants welcomed the practical nature of the presentations. During the discussion parts it became clear that especially attribute harmonization and privacy policies of-

ten clash with providing an easy-to-use infrastructure.

During the afternoon the use of Persistent Identifiers was the main topic. After presenting some key questions about this topic (think of granularity and versioning) it was time to dive into the details of the EPIC API.

The tutorial was concluded by a day of Component Metadata: relevant software concepts as ISOcat, the Component Registry, the VLO, the CMDI browser and Arbil were introduced. Thorsten Trippel gave some explanation about how CMDI was used for the purpose of the NaLiDa-project.

All presentations:

<http://www.clarin.eu/events/3443>



**Dieter Van Uytvanck**  
*Max-Planck-Institut für Psycholinguistik,  
Nijmegen*



Photo: Thomas Zastrow