



HANNA HEDELAND, TOMMI ANTERO PIRINEN, HEIDEMARIE SAMBALE

DAS HAMBURGER ZENTRUM FÜR SPRACHKORPORA

DAS FORSCHUNGSDATEN- ZENTRUM

Das Hamburger Zentrum für Sprachkorpora (HZSK) ist ein Forschungsdatenzentrum an der Universität Hamburg, das sich insbesondere **gesprochenen, mehrsprachigen und multimodalen Daten** sowie **Daten aus der linguistischen Feldforschung** widmet. Als mehrfach zertifiziertes CLARIN-Zentrum steht das HZSK Wissenschaftlern, die mit solchen Daten arbeiten, in allen Fragen des Forschungsdatenmanagements zur Seite – von der **Datenerhebung** über die **Auswertung** bis zur **Publikation** und der **geregelten Nachnutzung** der Projektdaten. In regelmäßigen Abständen werden allgemeine **Schulungen** und nach Möglichkeit **projektspezifische Workshops** angeboten. Zudem betreibt das Zentrum den **CLARIN-D-Helpdesk**, über den Wissenschaftler Unterstützung in vielen weiteren Fragen der digitalen sprachbasierten Forschung bekommen.

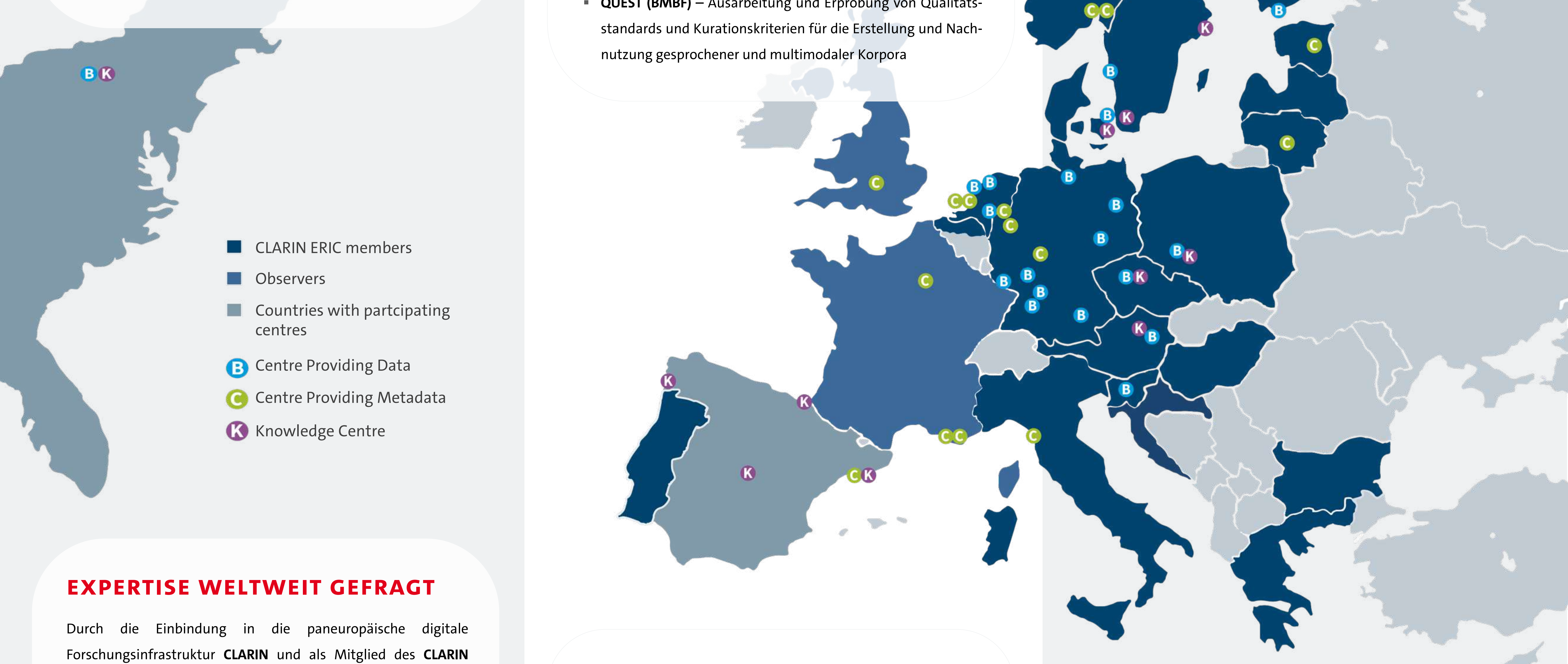
WEITVERZWEIGTES NETZWERK

Am HZSK entwickeln und erweitern verschiedene Projekte eine gemeinsame technische Infrastruktur für die **Erstellung, Auswertung** und kontrollierte **Weitergabe** hauptsächlich audiovisueller annotierter Sprachdaten. Die Zusammenarbeit bezieht sich auch auf entsprechende Standards und **Best Practices** des Forschungsdatenmanagements. Aktuell sind folgende Projekte daran beteiligt:

- **CLARIN-D (BMBF)** – Aufbau einer digitalen Forschungsinfrastruktur für die Geistes- und Sozialwissenschaften
- **CLARIAH-DE (BMBF)** – Zusammenführung der beiden digitalen Forschungsinfrastrukturen CLARIN-D und DARIAH-DE
- **ZuMult (DFG)** – zielgruppenspezifische Zugänglichmachung gesprochener Korpora
- **INEL (Akademienprogramm)** – Aufarbeitung sprachlicher Ressourcen indigener Sprachen und Varietäten Nordeuropas
- **QUEST (BMBF)** – Ausarbeitung und Erprobung von Qualitätsstandards und Kurationskriterien für die Erstellung und Nachnutzung gesprochener und multimodaler Korpora

AKTUELLER, INTUITIVER, VIELSEITIGER

Das HZSK entwickelt zusammen mit dem Institut für Deutsche Sprache (Mannheim) die **EXMARaLDA-Software** zur Erstellung, Verwaltung, Analyse und Publikation von Korpora gesprochener Sprache. Im Rahmen einer Kooperation mit dem **Language Technology Group** an der Universität Hamburg wurde kürzlich auch das webbasierte Annotationswerkzeug **WebAnno** für transkribierte Audio-/Videodaten implementiert. Ferner ist das HZSK auch an der **Entwicklung datentechnischer Standards** für Forschungsdaten der angewandten Linguistik beteiligt, um die **Interoperabilität von Daten aus verschiedenen Kontexten** (Projekten, Forschungsfragen, theoretischen Rahmen) technisch zu ermöglichen.



EXPERTISE WELTWEIT GEFRAGT

Durch die Einbindung in die paneuropäische digitale Forschungsinfrastruktur **CLARIN** und als Mitglied des **CLARIN Knowledge Centre for Linguistic Diversity and Language Documentation** bietet das HZSK seine Expertise auch über die Grenzen Deutschlands hinaus an. Darüber hinaus mussten für die Zertifizierung als CLARIN-Servicezentrum hohe Anforderungen an die lokale technische Infrastruktur erfüllt werden. So ist das Forschungsdatenrepositorium als vertrauenswürdige Archiv zertifiziert (**Core Trust Seal**) und auch die Authentifizierung mit Kennungen vieler europäischer akademischer Institutionen (**Single Sign-on**) erlaubt den direkten Zugriff auf Ressourcen, die der akademischen Öffentlichkeit zur Verfügung stehen.

FÜR UND MIT FORSCHERN

Das HZSK übernimmt Korpora aus drittmittelgeförderten Projekten und anderen Kontexten um sie der **wissenschaftlichen Öffentlichkeit für Forschung und Lehre zur Verfügung zu stellen**. Der Zugriff auf die Daten kann für die Nutzer durch unterschiedliche Zugangsrestriktionen beschränkt werden.

Es gibt Kooperationen in verschiedenen Formen mit **über 50 Forschungsvorhaben und Institutionen**, die ihre Daten dem HZSK übergeben haben und/oder auf seine Expertise und Dienstleistungen zurückgreifen.

HAMBURGER ZENTRUM FÜR SPRACHKORPORA
CORPORA.UNI-HAMBURG.DE

hzsk hamburger zentrum
für sprachkorpora

 **EXMARaLDA.org**