CLARIN: AN OVERVIEW

CLARIN AND LIBRARIES WORKSHOP
ROYAL LIBRARY OF THE NETHERLANDS, THE HAGUE
2022, MAY 9TH & 10TH

Andreas Witt, National Coordinator CLARIN-D

Leibniz Institute for the German Language
Mannheim
Mannheim University
(& Heidelberg University & University of Cologne)

Main Goals of CLARIN

- Construction of an integrated, interoperable, and scalable research infrastructure via a network of centers
- Deployment of linguistic data, tools, and services
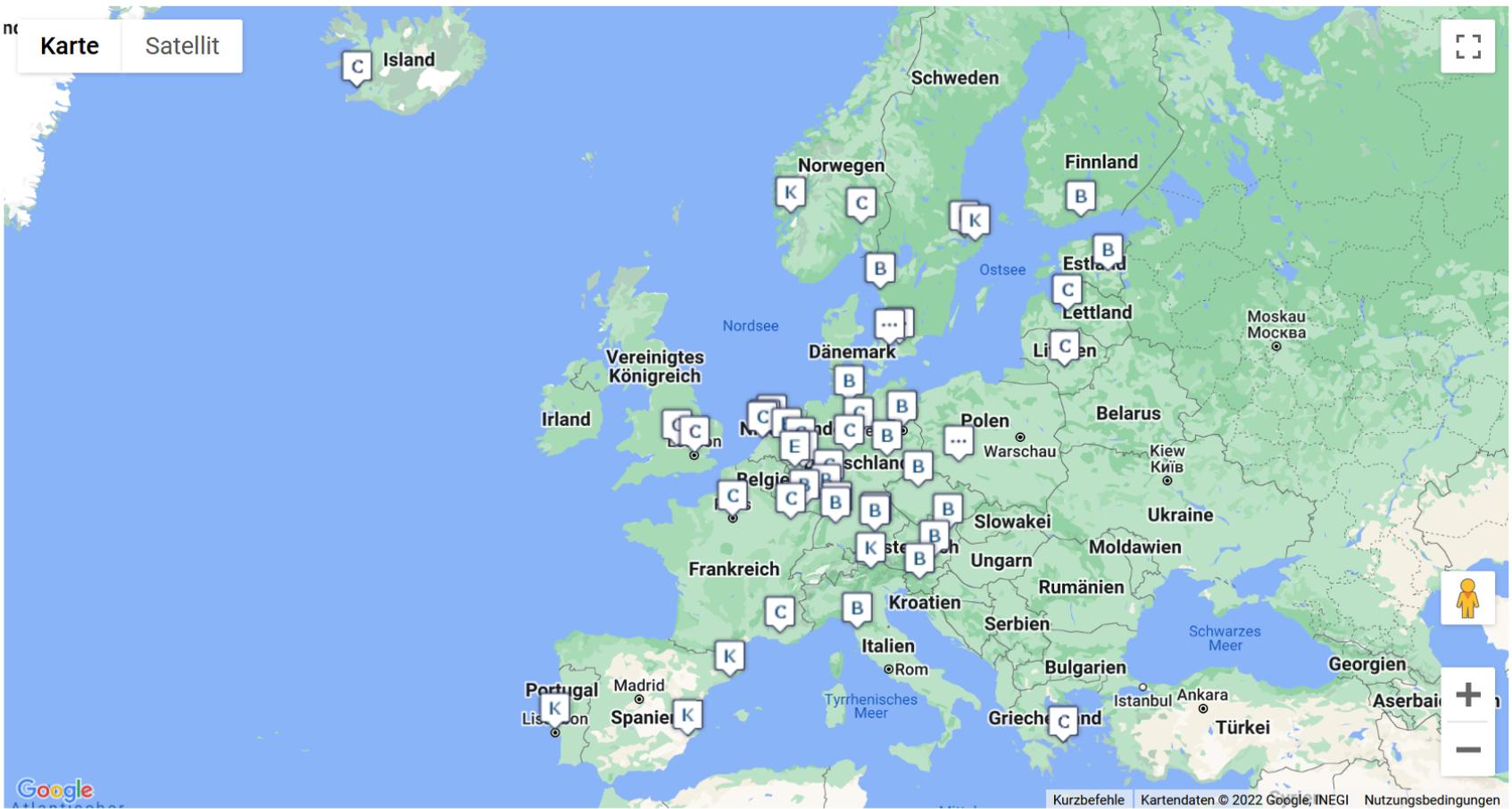- Target group: researchers from the social sciences and humanities

Central structures of CLARIN

- CLARIN Centres: hosting the infrastructure
  - Central services (PID, search tools, …)
  - Hosting data
  - Training and dissemination
  - Supporting community

CLARIN

CLARIN – organizational structure

- Legal status
  - ERIC: European Research Infrastructure Consortium
  - The Members of an ERIC are countries
- Governance
  - General Assembly (representatives from the ministries of the member countries)
  - Scientific Advisory Board
  - Board of Directors
  - Several committees

CLARIN - Government structure - BoD

Board of Directors

The Board of Directors (*BoD*) is responsible for the day-to-day management of CLARIN *ERIC* and for the implementation of the strategies and policies set out by the General Assembly.

	Franciska de Jong	Executive Director
	Dieter Van Uytvanck	Vice Executive Director and Technical Director
	Francesca Frontini	Member of the Board of Directors
	Antal van den Bosch	Member of the Board of Directors

CLARIN in Europe / Members

Members	National Consortium (NC)	Leading NC Partner	National Coordinator
Austria	CLARIAH-AT	ACDH-CH	Karlheinz Mörth
Belgium	CLARIN BE	Instituut voor de Nederlandse Taal (INT)	Vincent Vandeghinste
Bulgaria	CLaDA-BG	Bulgarian Academy of Sciences	Kiril Simov
Croatia	HR-CLARIN	University of Zagreb	Marko Tadić
Cyprus	CLARIN-CY	Digital Heritage Research Lab (Cyprus University of Technology)	Marinos Ioannides
Czech Republic	LINDAT/CLARIAH-CZ	Charles University Prague	Eva Hajičová
Denmark	CLARIN-DK	University of Copenhagen Center of	Costanza Navarretta

CLARIN in Europe / Members

Members	National Consortium (NC)	Leading NC Partner	National Coordinator
Finland	FIN-CLARIN	University of Helsinki	Krister Lindén
Germany	CLARIN-D	Institut fuer Deutsche Sprache, University of Mannheim	Andreas Witt
Greece	clarin:el	ILSP-ATHENA Research Center	Stelios Piperidis
Hungary	HunCLARIN	Research Institute for Linguistics, Hungarian Academy of Sciences	Tamás Váradi
Iceland	CLARIN Iceland	The Árni Magnússon Institute for Icelandic Studies Institute for	Starkaður Barkarson

CLARIN in Europe / Members

Members	National Consortium (NC)	Leading NC Partner	National Coordinator
Latvia	CLARIN-LV	Institute of Mathematics and Computer Science, University of Latvia	Inguna Skadiņa
Lithuania	CLARIN-LT	Vytautas Magnus University	Jurgita Vaičenonienė
The Netherlands	CLARIAH-NL	Utrecht University	Jan Odijk
Norway	CLARINO	University of Bergen	Koenraad De Smedt
Poland	CLARIN PL	Wroclaw University of Technology	Maciej Piasecki
Portugal	PORTULAN CLARIN	University of Lisbon	António Branco
Slovenia	CLARIN.SI	Jožef Stefan Institute	Tomaž Erjavec
Sweden	SWE-CLARIN	Språkbanken	Lars Borin

CLARIN – Observers and 3ʳᵈ party partners

Observer	National Consortium (NC)	Leading Partner NC	National Coordinator
South Africa	SADiLaR	North-West University	Langa Khumalo
United Kingdom	CLARIN-UK	Oxford University	Martin Wynne

Third Party	National Consortium (NC)	Leading Partner NC	Coordinator
Carnegie Mellon University (USA)	TalkBank	Carnegie Mellon University (USA)	Brian MacWhinney

Common Language Resources and Technology Infrastructure

Certificate of Recognition

CMU Talkbank

The Executive Director of CLARIN ERIC hereby acknowledges that the aforementioned institution has successfully completed the CLARIN Centre Assessment Procedure and is officially recognized as a

CLARIN B Centre

The aforementioned institution has permission to use the following logo in its official communications:

CLARIN B CENTRE

Franciska de Jong
Executive Director
CLARIN ERIC
Issued on: 12 October 2021
Valid till: 12 October 2024

PID: http://hdl.handle.net/11372/DOC-109

Display K-centres

Display K-centres

Full name	IMPACT centre of competence - CLARIN K-centre in digitisation
Short name	IMPACT-CKC
URL	https://www.digitisation.eu
Hosted by	(1) Universidad de Alicante (UA), Alicante, Spain (2) Instituut voor de Nederlandse Taal (INT), Leiden, The Netherlands (3) Koninklijke Bibliotheek (KB), The Hague, The Netherlands (4) Poznan Supercomputing and Networking Center (PSNC), Poznan, Poland (5) The British Library (BL), London, United Kingdom (6) Bibliothèque nationale de France (BnF), Paris, France (7) Universidad de Guadalajara Mexico (UgG), Guadalajara, Mexico
City of main hub	Alicante
Country of main hub	ES
Area of competence	IMPACT-CKC (IMPACT centre of

VLO

<< | < | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | > | >>

Use the categories below to limit the search results to those matching the selected value (s).

- Language
- Collection
- Resource type
- Modality
- Format
- Keyword

- Availability
- Search options

WordNet

plWordNet is a lexico-semantic network which reflects the lexical system of the Polish language. There are at present ca. 144,000 nouns, verbs and adjectives in plWordNet, ca. 203,000 word senses and ca. 500,000 relations. It is already the second-largest wordnet in the world, and it keeps growing.

*

Finnish WordNet

The language resource is available in Kielipankki - the Language Bank of Finland at http://urn.fi/urn:nbn:fi:lb-2016042205; download: http://urn.fi/urn:nbn:fi:lb-2014052713. The Finnish WordNet is a lexical database for Finnish. It is a part of the FIN-CLARIN infrastructure project. FinnWordNet is lic…

Czech WordNet

Multilingual Lexicons; The Czech WordNet was developed by the Centre of Natural Language Processing at the Faculty of Informatics, Masaryk University, Czech Republic. The Czech WordNet captures nouns, verbs, adjectives, and partly adverbs, and contains 28,201 word senses (synsets). Every synset encodes the equivalenc…

Bulgarian WordNet

Multilingual Lexicons; The Bulgarian WordNet is a network of lexical-semantic

Supporting five phases of e-Humanities projects

Finding, acquisition and preprocessing of data

Data mining, data analysis

Visualizing data

Providing access to data

Data archiving
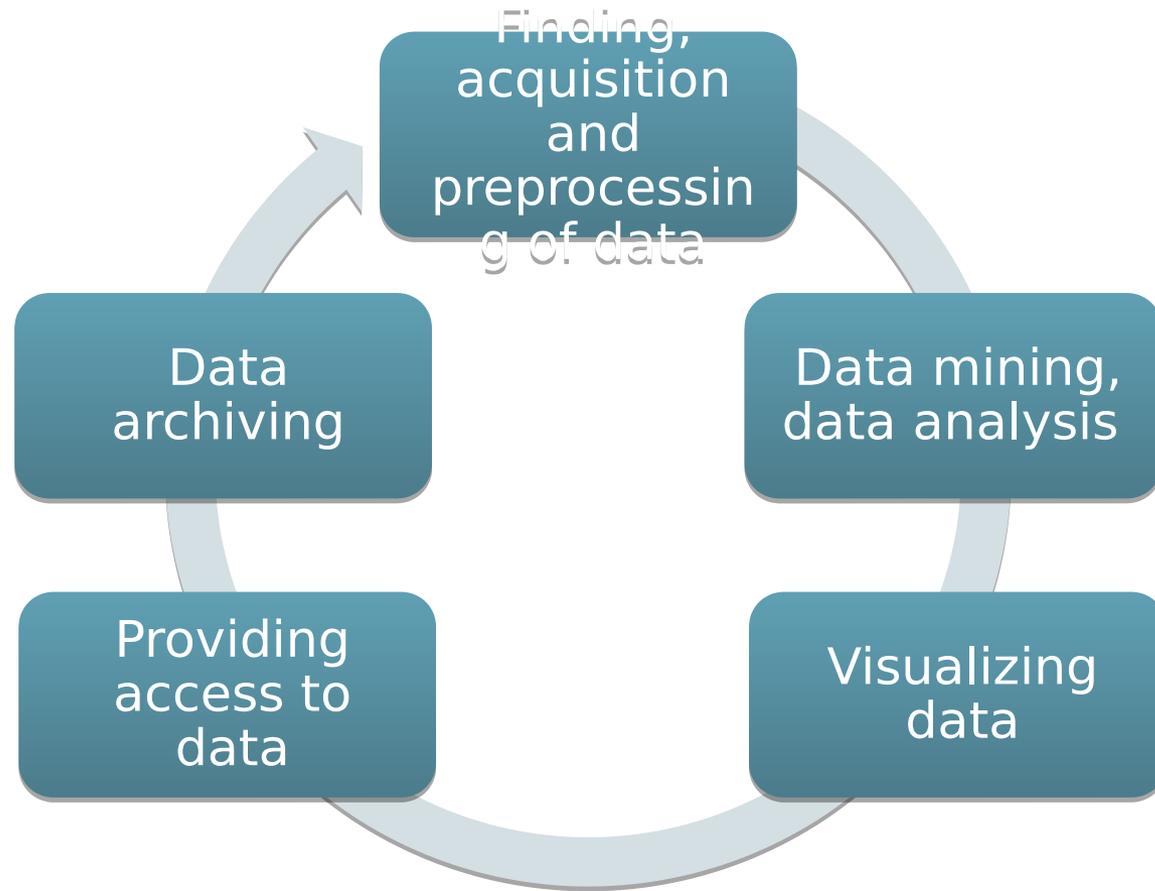
Referencing resources

- Resources referenced via persistent identifiers (PIDs)
- PIDs resolve to a repository/archive
- If repository moves, resolvers change
- No requirement for "last visited" or similar

---

[7]All the sentence materials, instructions and a digital version of the questionnaire are freely available online: http://hdl.handle.net/11022/0000-0000-1CCA-E (Experiment 1), http://hdl.handle.net/11022/0000-0000-1CCB-D (Experiment 2), and http://hdl.handle.net/11022/0000-0000-1CCC-C (Experiment 3).

Searching in resources

Federated Content Search (FCS)

- Distributed search in a common interface
- No central index
  - Respecting IPR
  - Using existing search environments
- Webservice based
  - Rest-based webservices
  - Chains of tools with integrated AAI (no individual login required)

16

CLARIN and Standards

CLARIN and standards

Why are standards important for CLARIN?

- The project partners and centers is heterogeneous in both topics and organizational structure
- The resources available are heterogeneous regarding topics, technical formats, and copyright
- Standards may help to establish interchange

Standards and CLARIN

- In CLARIN, everyone agrees that the use of standards is desirable to interoperate with each other
- But: many researchers still lack knowledge to use 'the right standards'
- Even if information about a specific standard is available, information about its relations to other standards is not

The CLARIN Standards Committee

The CLARIN Standards Committee (CSC)

- Members can be anyone involved in CLARIN
- Chair: Piotr Bański, Germany
- Vice-Chair: Karlheinz Mörth, Austria
- See https://www.clarin.eu/governance/standards-committee
  for a list of current members

Legal & ethical issues in CLARIN

Legal frameworks concerning language resources

- Copyright
- *sui generis* database right
- Data protection

CLIC – CLARIN Legal Issues Committee

- Legal experts from every national consortium
- **Chair**: Pawel Kamocki – here in the room
- Regular meetings
- reports to the Board of Directors
- coordination of CLARIN ERIC's lobbying efforts at the EU level
- evaluation of relevant bills at the EU level

CLARIN's 'legal offer' for researchers

- CLARIN Legal Infrastructure & standard licence terms
- CLARIN Legal Information Platform
- CLARIN Public License Selector & Calculator
- CLARIN-D Legal Helpdesk
- Lobbying & Networking

Clarin Annual Conferences

Clarin Annual Conferences

- CLARIN 2021, 27-29 September 2021, Virtual event
- CLARIN 2020, 5-7 October 2020, Virtual Event
- CLARIN 2019, 30 September - 2 October 2019, Leipzig, Germany
- CLARIN 2018, 8-10 October 2018, Pisa, Italy
- CLARIN 2017, 18-20 September 2017 , Budapest, Hungary
- CLARIN 2016, 26-28 September 2016, Aix-en-Provence, France
- CLARIN 2015, 15-17 October 2015, Wroclaw, Poland
- CLARIN 2014, 23-25 October 2014, Soesterberg, The Netherlands
- CLARIN 2013, 21-22 October 2013, Prague, Czech Republic
- CLARIN 2012, 24-27 October 2012, Sofia, Bulgaria

CLARIN Annual Conference 2022

10 October 2022 - 12 October 2022

St. Vitus Cathedral, Prague Castle, Prague, Czechia

About the CLARIN Annual Conference

The CLARIN Annual Conference is the main annual event for those working on the construction and operation of CLARIN across Europe, as well as for representatives of the communities of use in the humanities, and social sciences.

In 2020 and 2021, the CLARIN Annual Conference took place virtually. In 2022, we hope to at least partially return to the traditional format, with Prague, Czechia, as the venue. The event will be organised by CLARIN ERIC in collaboration with LINDAT/CLARIAH-CZ.

CLARIN2022 is organised for the wider humanities and social sciences communities in order to exchange

CLARIN: AN OVERVIEW

Andreas Witt

Leibniz Institute for the German Language
Mannheim
Mannheim University
(& Heidelberg University & University of Cologne)