



Austausch und Mobilität  
Echanges et mobilité  
Scambi e mobilità  
Exchange and mobility



# CLARIN AND ITS SWISS NODE CLARIN-CH: SUPPORTING RESEARCH BASED ON LANGUAGE RESOURCES

PRESENTATION AT THE 2ND SWISS MULTIPLIER UPKILLS WORKSHOP  
CRISTINA GRISOT, CLARIN-CH SCIENTIFIC COORDINATOR  
UNIVERSITY OF GENEVA, 21 APRIL 2023



# The research infrastructure for language as social and cultural data

CLARIN is a digital infrastructure offering data, tools and services to support research based on language resources.

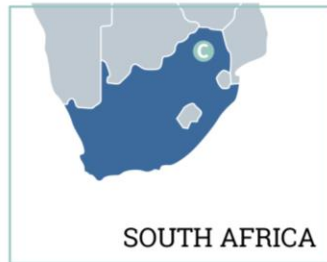
- CLARIN is a pan-European research infrastructure that provides easy and sustainable access to digital language resources to support research in the SSH and beyond
- Fully operational since 2016
- 23 European countries are members of CLARIN
- Switzerland has joined CLARIN ERIC on 1 January 2023 as observer



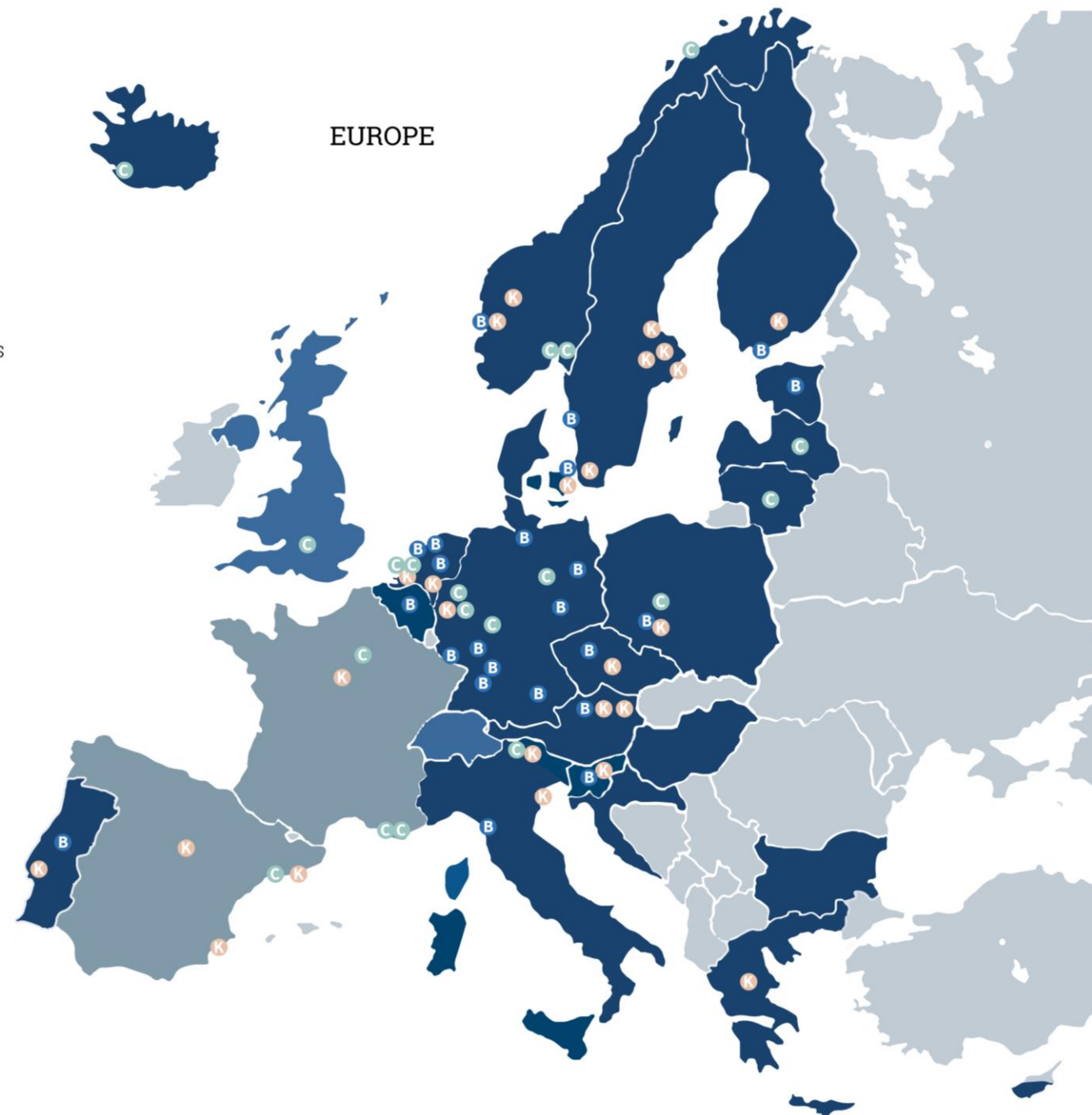
- ERIC members
- Observers
- Countries with participating centres
- B** Centre Providing Data
- C** Centre Providing Metadata
- K** Knowledge Centre



USA



SOUTH AFRICA





### **CLARIN's vision:**

Render accessible all digital language resources and tools from all over Europe through a single sign-on online environment.

### **CLARIN's mission:**

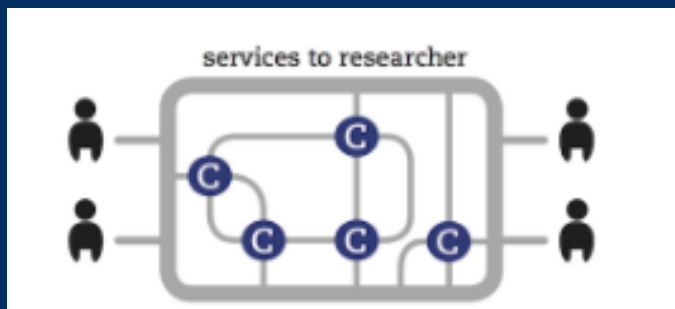
Create and maintain an infrastructure to support the sharing, use and sustainability of language data and tools for research in the SSH.





## CLARIN is a distributed research infrastructure.

Concretely, CLARIN is a networked federation of centers— academic and research institutions that offer the scientific community expertise, resources, tools and services.



Researchers are both service-users and service-providers

**Service Provider  
Centers  
B-centers**

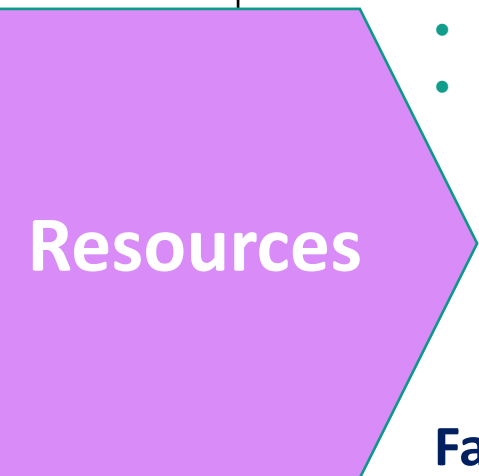
**Knowledge  
Centers  
K-centers**

**Metaknowledge  
centers  
C-centers**



## Families of corpora

- Computer-mediated communication corpora
- Corpora of academic texts
- Historical corpora
- **L2 learner corpora**
- Literary corpora
- Manually annotated corpora
- Multimodal corpora
- Newspaper corpora
- Parallel corpora
- Parliamentary corpora
- Reference corpora
- Spoken corpora



**74**  
in 2023!

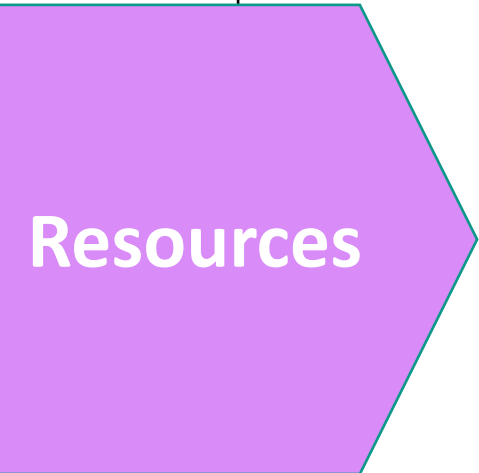
## Families of lexical resources

- Lexica
- Dictionaries
- Conceptual resources
- Glossaries
- Wordlists





# How to find and access CLARIN resources?



1. Manually, by searching in the Resource Families listing

2. CLARIN's unified catalogue: the Virtual Language Observatory VLO

- An easy-to-use interface
- A powerful query syntax



3. The individual CLARIN data repositories

- Located in the various CLARIN centers

4. The SSH Open Marketplace

- A discovery platform for the SSH field





## Services

1. **Services** provided by K-centers, such as:
  - ✓ Online courses Training materials
  - ✓ Best-practice documents
  - ✓ Guidance in getting access to and using data and tools
  - ✓ Hosting of receivers of CLARIN mobility grants
2. **Learning Hub**
  - ✓ Training Materials
  - ✓ Training Events
  - ✓ Teaching Award
  - ✓ DH Course Registry
3. **CLARIN Trainer Network programme**
4. **Data depositing services**
5. **CLARIN Youtube channel**



# What is CLARIN-CH?

# CLARIN-CH is the Swiss branch of CLARIN

In December 2020, a network of six Swiss academic institutions, supported by the Swiss Academy for SSH, founded the CLARIN-CH Consortium.



# Mission



*Join the European CLARIN community and build an active and impactful national network.*

## Objectives:

1. Obtain Switzerland's full membership in CLARIN and connect the Swiss scientific community with the entire CLARIN infrastructure.
2. Foster the sharing of expertise and of resources at the national and European level.
3. Bring together the Swiss community using language resources and create national working groups.
4. Encourage the initiation of national and international collaborations.

# National node

## TECHNICAL CENTER



University of Zurich <sup>UZH</sup>

LiRI - Linguistic Research Infrastructure

LiRI is a national technology platform to serve the needs of quantitative research at UZH, Switzerland and beyond.



## NATIONAL REPOSITORY

Library

University of Zurich

SWISS  base

LaRS - Language Repository of Switzerland

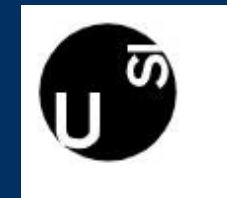
LaRS is a national platform for the publication of linguistic research data.



## TRAINING OFFERS



Doctoral programme in applied linguistics: Argumentation in Professional Practice



Sommerschule der Zürcher Korpuslinguistik und Korpuspragmatik



Universität Zürich <sup>UZH</sup>

Deutsches Seminar - Zukoko

# LIRI AS CLARIN B-CENTER

## 1. Language Technology

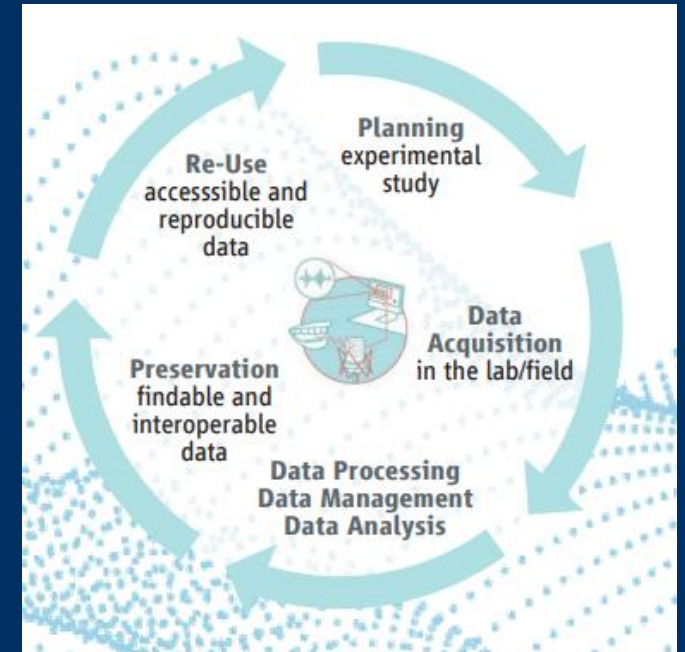
- IT services
- Application development
  - LiRI Corpus Platform (LCP) to handle and query large corpora
- Data processing & NLP
  - Automatic data processing with NLP tools
  - Best practices, coaching, workshops, programming
  - Swissdox@LiRI: a tool that allows retrieving large quantities of Swiss media data for research purposes

## 2. Lab & Equipment

- On site and portable

## 3. Statistical consulting

- Advice or feedback on study design, statistical methods
- Data analysis service



- ✓ Deals with the entire data life-cycle
- ✓ Provides support services, facilities and equipment

# NATIONAL REPOSITORY



LaRS is a FAIR compliant national platform for the publication and long-term archiving of linguistic research data on Swiss servers (SWITCH)

It uses SWISSUbase as its repository system.



<https://info.swissubase.ch>

[swissubase@ub.uzh.ch](mailto:swissubase@ub.uzh.ch)

## KEY FEATURES

FAIR repository recognised by the  
Swiss National Science Foundation

Multi-disciplinary & multilingual

Long-term preservation of your data

Access control options (prior  
approval, embargo, closed contract, etc.)

Free persistent identifier (DOI) for  
all published datasets

Connected with Swiss and European  
key research catalogues

Data curated by archiving experts

Expert support for depositing and  
re-using data

Free for researchers

# CLARIN-CH national working groups



Spring 2023

1. ORD projects for linguistic data
  - i. UpLORD*
  - ii. Swiss-AL: Linguistic ORD Practices for Applied Sciences*
  - iii. Data-sharing skills in corpus-based research on talk-in-interaction*
2. Management of sensitive data and legal aspects for linguistic data in Switzerland
3. Research infrastructures for Argumentation and Rhetoric
4. Multilingual corpora and Second Language Acquisition

OPEN TO  
MEMBERS

# Take home messages



1. CLARIN is a distributed research infrastructure for language resources and language technology.
2. CLARIN-CH is a consortium of scholars targeting the development of an active and impactful Swiss CLARIN community.
3. CLARIN and CLARIN-CH are based on networking and sharing of expertise, language resources, and tools.
4. Researchers are both service-users and service-providers.
5. Conducting research with digital language data can be challenging at the individual level, but much easier at the collective level.





Marianne Hundt  
National Coordinator  
Representative of UZH



Noah Bubenhofer  
NC ad interim  
Director of LiRI



Anita Auer  
President  
Representative of UNIL



Cristina Grisot  
Scientific Coordinator



Martin Luginbühl  
Representative of UNIBAS



Sandrine Zufferey  
Representative of UNIBE



Eric Haeberli  
Representative of UNIGE



Anita Thomas  
Representative of UNIFR



Martin Hilpert  
Representative of UNINE



Andrea Rocci  
Representative of USI



Cerstin Mahlow  
Representative of ZHAW



Beat Immenhauser  
Representative of the ASSH

## Meet the members of the CLARIN-CH consortium

## Home

National Consortium

Governance and Members

CLARIN-CH resources

CLARIN-CH national node

CLARIN-CH co-branding

CLARIN-CH working groups

CLARIN Europe

 News

## Welcome to CLARIN Switzerland



CLARIN – Common Language Resources and Technology Infrastructure – is a pan-European research infrastructure aiming to render accessible language resources and tools from all over Europe through a secure online environment.



Several Swiss academic institutions have manifested their interest to join CLARIN, first as an Observer member and later as Full members. For this, they have founded the consortium CLARIN-CH on 1st October 2020.

The **mission** of CLARIN-CH: **join the European CLARIN community, and build an impactful CLARIN-CH community.**

For this, we formulated four **objectives**:

✓ Obtain Switzerland's CLARIN membership and give Swiss researchers access to the entire CLARIN infrastructure.



## EMAIL

[cristina.grisot@uzh.ch](mailto:cristina.grisot@uzh.ch)

## WEB

<https://clarin-ch.ch>

## SOCIAL MEDIA

<https://www.linkedin.com/company/clarin-ch>

# Thank you!