CLARIN-CH: SUPPORTING RESEARCH BASED ON LANGUAGE RESOURCES

Katharina Eggenberger & Cristina Grisot

CLARIN GA - Lisbon, 19 April 2023

POLITICAL FRAMEWORK

Switzerland not Member of the European Union

Horizon Europe: national funding available for Swiss partners in collaborative projects

Specific parliamentary approval for Swiss membership in ERICs in December 2022

Recognised importance of internationally distributed research infrastructures

ROADMAP FOR RI

Periodic elaboration of a Swiss Roadmap for Research Infrastructures

Linguistic Research Infrastructure (LiRI) hosted by University of Zurich, included in Roadmap 2019 edition

CLARIN ERIC included in Roadmap 2023 edition

CLARIN ERIC: observer as of 1st January 2023/ membership planned for 2025

Challenges ahead: ensure sustainable funding for research infrastructures hosted by universities which act as national nodes in ERICs

CLARIN-CH CONSORTIUM

CLARIN-CH

December 2020

USI

zhaw

UNIVERSITÄT BERN

SAGW ASSH

University of Zurich[UZH]

UNIL | Université de Lausanne

unine Université de Neuchâtel

Universität Basel

UNIVERSITÉ DE GENÈVE

UNI FR

UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

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

CLARIN-CH

TECHNICAL CENTER

University of Zurich[UZH]

LiRI - Linguistic Research Infrastructure

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

https://www.liri.uzh.ch/

NATIONAL REPOSITORY

Library
University of Zurich

SWISSbase

LaRS – Language Repository of Switzerland

LaRS is a national platform for the publication of linguistic research data. It uses SWISSUbase as its repository system.

https://www.swissubase.ch/en/

LIRI AS TECHNICAL CENTER

- Provides support services, facilities and equipment
- Enables collaborative research
- Deals with big data
- Engages with Open Science
- Enables Data Reproducibility
- Deals with the entire data life-cycle

Lab & Equipment

- On site and portable

Statistical consulting

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

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

SOS **Charged services**

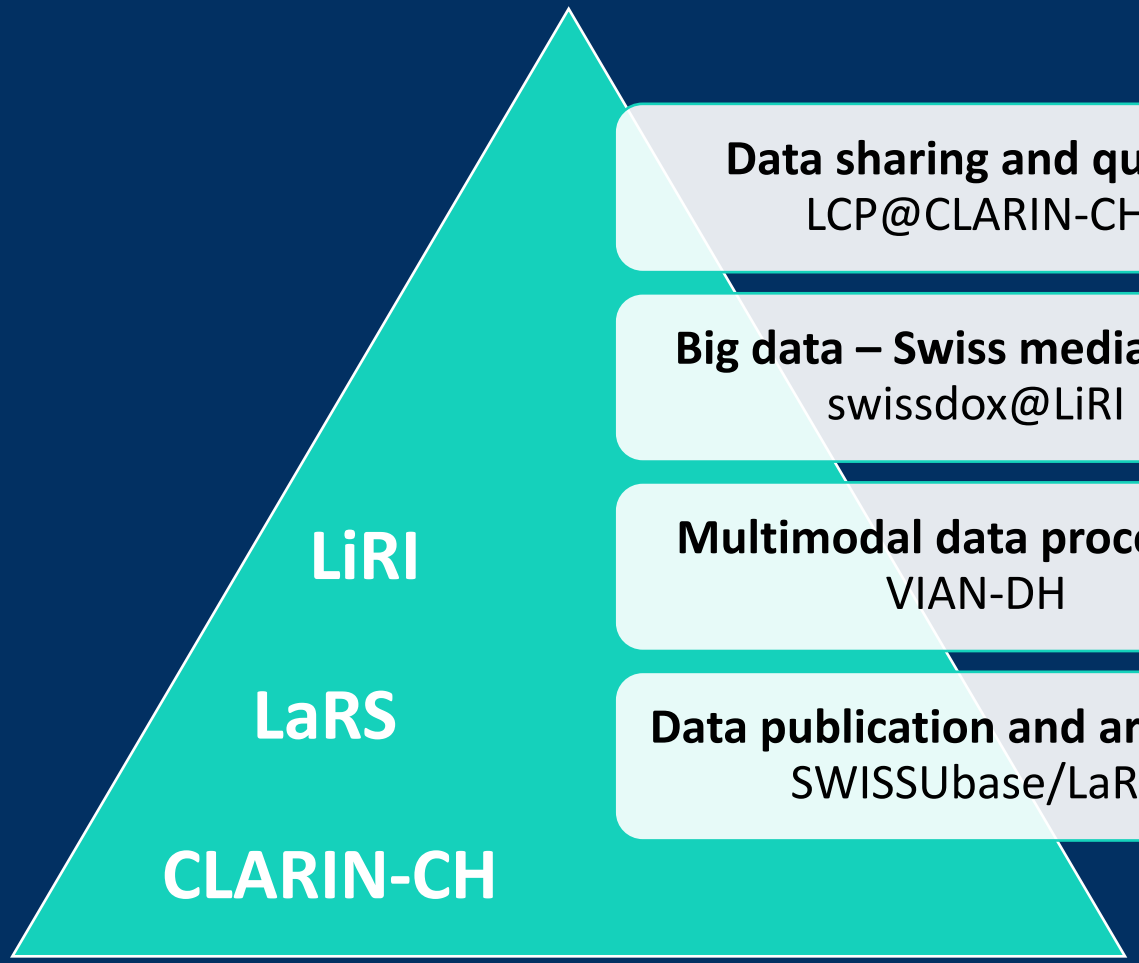
NEW **To be added in the budget of research projects**

UPGRADING THE SWISS LINGUISTIC ORD-ECOSYSTEM – *UPLORD* PROJECT

CLARIN-CH

Data Life-cycle

- Data generating
- Data processing
- Data analyzing
- Data sharing and archiving



Data sharing and query:
LCP@CLARIN-CH

Big data – Swiss media texts:
swissdox@LiRI

Multimodal data processing:
VIAN-DH

Data publication and archiving:
SWISSUbase/LaRS

FAIR resources
Standardization
Interoperability

Annotation tool for multimodal data

Sensitive data
Legal issues
Metadata

Training
Documentation
Best practices

WORKING GROUPS

- CLARIN-CH WG are groups of researchers (CLARIN-CH members, other national and European scholars) who are interested in language-, resource- and infrastructure-related topics.

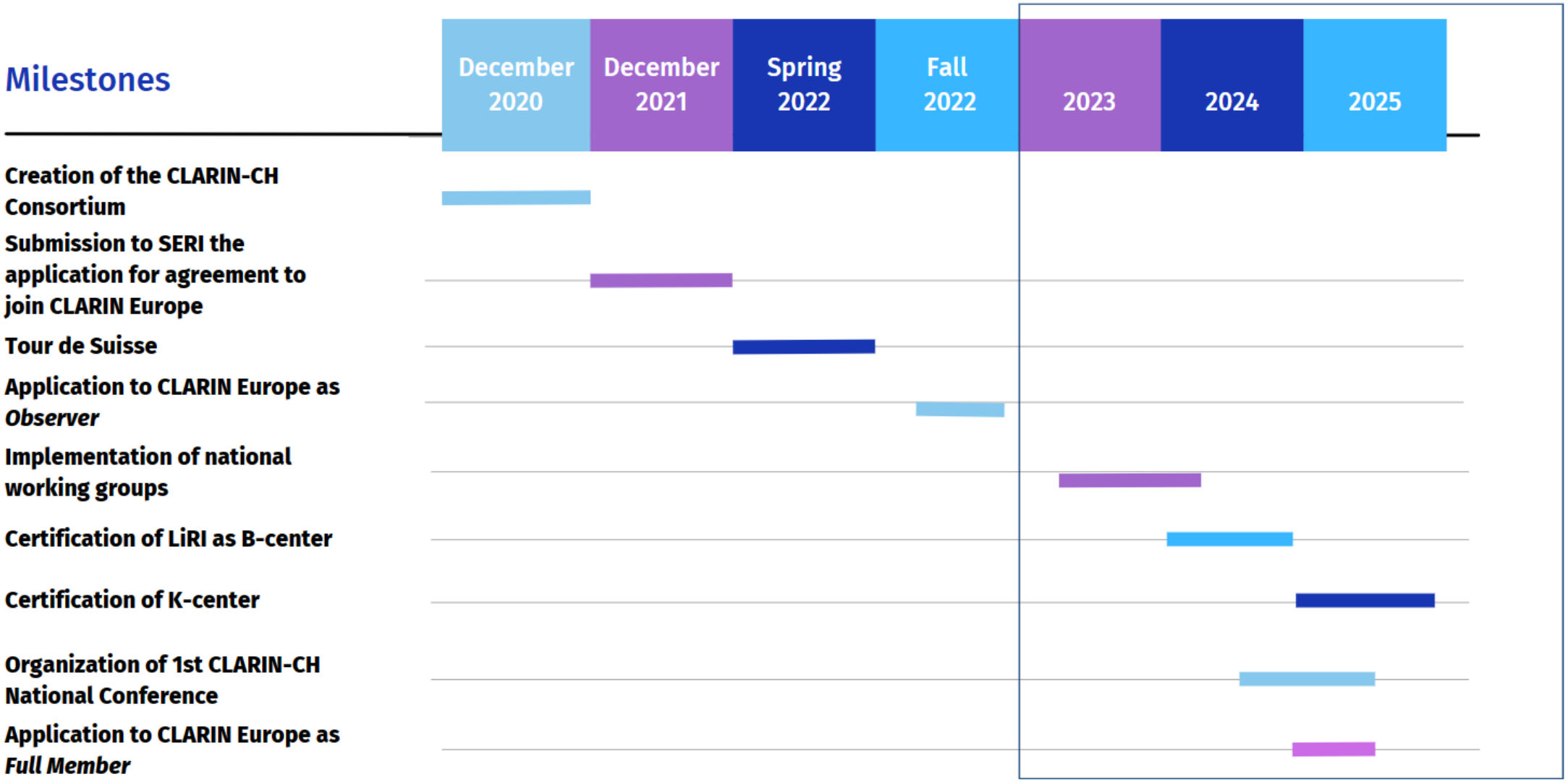
- **Advantage**: work together with peers in a formalized and sustainable environment.
- **Purpose**: bring together expertise on a specific topic, prepare joint research projects and to serve the community, which can be scholarly and technical.
- **Long-term goal**: build and to extend the CLARIN and CLARIN-CH infrastructure.

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

ROADMAP UNTIL 2025



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

Katharina Eggenberger
National Representative

Martin Luginbühl
Representative of UNIBAS

Sandrine Zufferey
Representative of UNIBE

Eric Haeberli
Representative of UNIGE

Anita Thomas
Representative of UNIFR

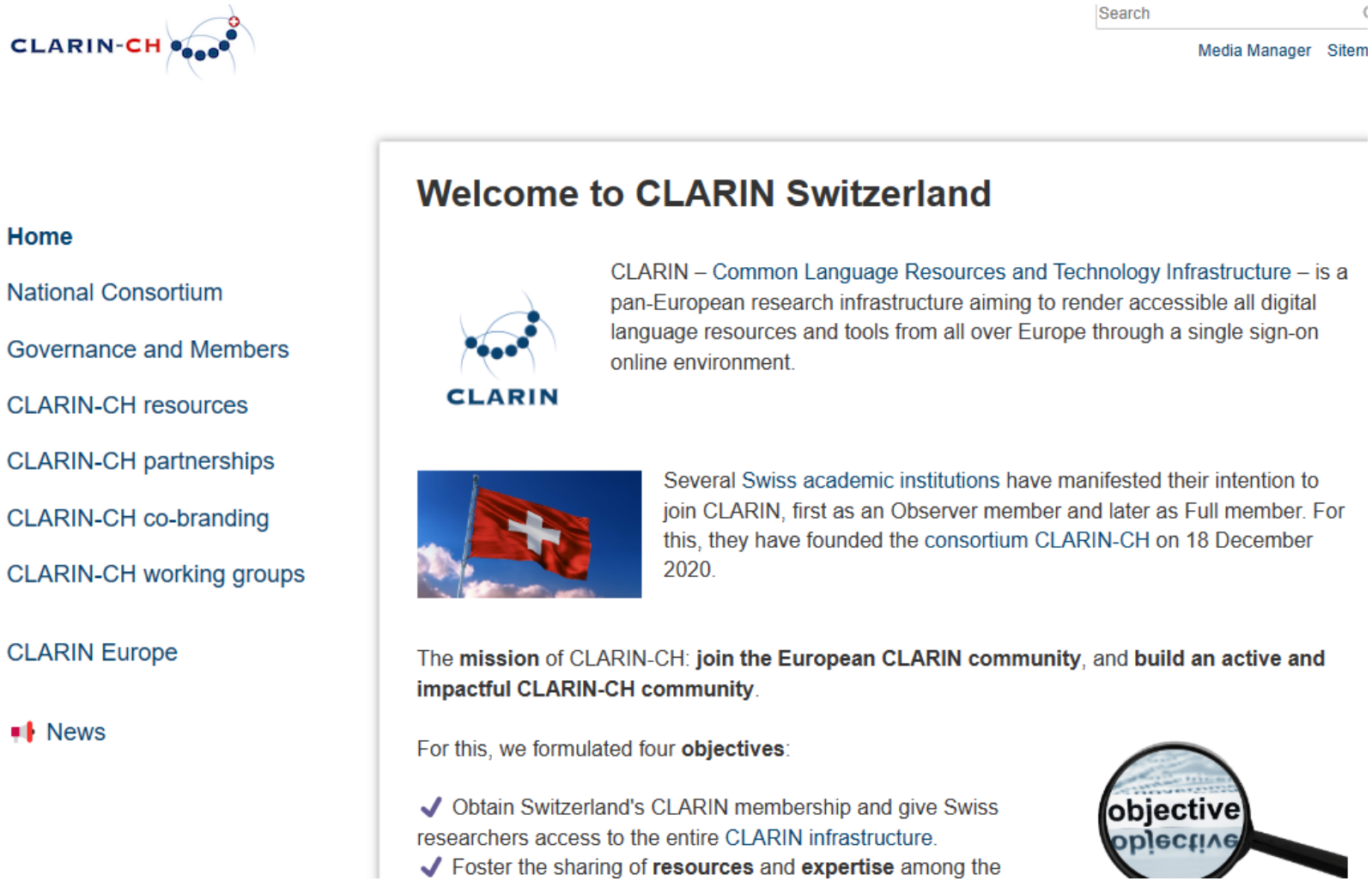
Martin Hilpert
Representative of UNINE

Andrea Rocci
Representative of USI

Cerstin Mahlow
Representative of ZHAW

Beat Immenhauser
Representative of the ASSH

Meet the representatives of Switzerland and the members of the CLARIN-CH consortium



Thank you!

EMAIL

cristina.grisot@uzh.ch

WEB

https://clarin-ch.ch

SOCIAL MEDIA

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