

# State of the Infrastructure

Franciska de Jong

Executive Director

[f.m.g.dejong@uu.nl](mailto:f.m.g.dejong@uu.nl)

Leipzig, 1 October 2019







# ANNUAL CONFERENCE 2019

## Leipzig, Germany



# Board of Directors



**Franciska de Jong**



**Dieter Van Uytvanck**



**Andreas Witt**



**Darja Fišer**

# CLARIN in data types

- Newspaper archives
- Literary texts
- Social Media data
- Parliamentary records
- Historical letters
- Oral History data
- Disciplinary libraries
- Institutional archival data
- Broadcast archives
- ...

See also the info on the CLARIN Resource Families initiative: <https://www.clarin.eu/resource-families>

# CLARIN's user communities

- Digital Humanities
- Linguistics and Philology
- Translation and Lexicography
- Literary Studies
- History
- Political and Social Sciences
- Media Studies
- Culture, Folklore, Anthropology
- Speech therapy
- Teachers
- General Public

# CLARIN in services



## CLARIN portal

Get an example-based impression of what's currently available



## Depositing services

Store language resources in a sustainable repository at a CLARIN centre



## Virtual Language Observatory

Discover language resources using a faceted browser or a map



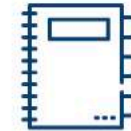
## Easy access to protected resources

Get easy access to protected resources, with your institutional username and password.



## Language Resource Switchboard

Explore and analyze language data with a wide variety of tools



## Virtual Collections

Create your own digital bookmarks, ideal for citing data sets.



## Language Resource Inventory

Submit and access information about language resources relevant to your research.



## Content Search (prototype)

Search different corpora with a single search engine



## Questions & Answers

Searching for a specific data set or application? Wondering how CLARIN can assist your research? Feel free to contact us!

# Strategic Planning 2018-2020

## Four priority areas

- Uptake by researchers
- Technical infrastructure
- Knowledge Sharing Infrastructure
- Sustainability

CLARIN's Strategic Planning webpage:

<https://www.clarin.eu/content/vision-and-strategy>

# Uptake by researchers

## Achievements in 2019 (highlights)

- **CLARIN Resource Families:** again a salient thread in the outreach to new communities. It has inspired work on requirements and on the visibility of data sets, standards and tools
- New programme launched to get access to researcher communities without a strong CLARIN history : **CLARIN Ambassadorship**
- **Tour de CLARIN:** now a flagship activity for promoting the national consortia, their resources, tools, events and researchers within and outside the CLARIN network. Will soon also include K-centres.
- Several awareness raising efforts and a (co)financing instrument for UI events in place and at least one **UI event** has been organised in most of the CLARIN countries, and over 7500 participants were counted
- **Training materials** on how to generate innovative research questions and methods have been developed.



*KPI – Number of views of tutorials and videos*

In September 2018 the number of YouTube views over the previous twelve months was 1470

In September 2019 the number of YouTube views over the previous twelve months was 3002



# Technical infrastructure

## Achievements in 2019 (highlights)

NB. Detailed presentation on Day 3 by Dieter Van Uytvanck

### *Centres, data, services*

- **Regular centre assessments**  
As planned, 2 assessment rounds were organised. Out of 57 registered CLARIN centres, 21 have been certified as B-centre (1 new, 9 recertifications)
- **Access to persistent identifier services**  
Next to its EPIC membership, CLARIN ERIC also became DataCite member, which ensures both technical and policy-wise access to the DOI ecosystem.
- **Closer integration of data and tools**  
More tools were connected to the *Language Resource Switchboard*  
Robustness checks (especially with regards to file size) were done based on datasets from Europeana
- Development of the **infrastructural services**: VLO, FCS, VCR, LRS

### *Workshops*

**ParlaFormat**: towards a standard TEI schema for parliamentary data; May 2019

*KPI – Number of certified B-centres*

In September 2018 21 of CLARIN's B-centres were certified

In September 2019 22 of CLARIN's B-centres were certified

# Knowledge Sharing Infrastructure

## Achievements in 2019 (highlights)

- **Annual conference:** >230 participants, 44 papers, PhD poster session, PhD council session
- A workshop was organised by CLIC on the **impact of the GDPR regulation** on the field
-  **DIGITAL HUMANITIES COURSE REGISTRY**
- The number of **K-centres** has grown: 10 -> 19
- New promotion for **mobility grants**
- **New citable papers** on how CLARIN complies with success criteria such as impact, FAIR and FACT:  
<https://www.clarin.eu/faq/how-can-i-cite-clarin>

*KPI – Average number of website visits per month*

In September 2018 the total number of visits to the CLARIN website over the previous twelve months was 78339

In September 2019 the total number of visits to the CLARIN website over the previous twelve months was 73742

# Sustainability

## Achievements in 2019 (highlights)

### *Extension of the ERIC*

**Cyprus** recently became member of CLARIN ERIC. A workshop is being planned for December 2019/early 2020 for consortia from the countries that recently joined and for representatives of countries that have an interest in joining.

### *Crossing Borders*

#### **Current and new EU projects**

- PARTHENOS, a thematic cluster around European RIs in the humanities
- DSI -4 - Collaboration with Europeana
- EOSC-hub, a pilot in the European Open Science Cloud, that started in January 2018.
- ERIC Forum, collaboration with the other ERICs
- SHHOC, towards a Social Sciences and Humanities Open Cloud; till April 2022. The CLARIN contribution to SSHOC will be realised by the central organisation and six CLARIN nodes. **Room for focus on migration studies, requirements for qualitative data analysis, support for processing interviews and heterogeneous data )**
- TRIPLE (kick-off in October 2019)
- ENRIITC (kick-off in January 2020)



**EOSC-hub**



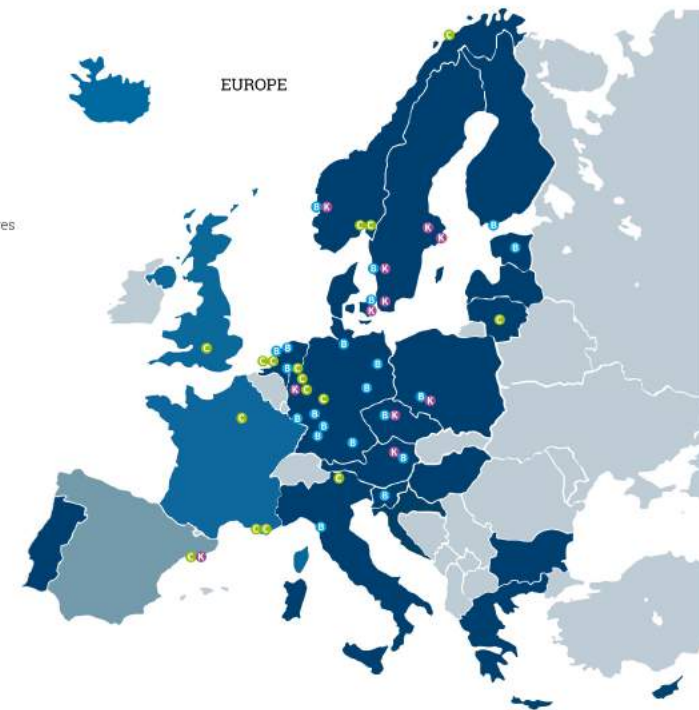
# CLARIN ERIC in members and centres

A consortium of:

- 20 members: AT, BG, CY, CZ, DE, DK, EE, FI, GR, HR, HU, IT, LT, LV, NL, NO, PL, PT, SE, SI
- 4 observers: IS, FR, UK, ZA
- >50 centres



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





# Sustainability

## *Alliances*

- CLARIN maintains strong collaborative ties with sister RIs, in particular with DARIAH-EU and CESSDA
- Active participation in the preparation of recent submissions of COST Action proposals.
- Other collaborative links:
  - ELRA, EGI, EUDAT-CDI, Europeana, LIBER, META-NET/ELG, OpenAIRE, Time Machine

## *Beyond Europe*

- Links with: Australia, Cameroon, Canada, Japan, Nigeria, Singapore, South Africa, USA
- Increased visibility through booth at international event DH2019

### *KPI – Number of members*

In September 2018 CLARIN had 20 members and 2 observers

In September 2019 CLARIN has 20 members and 4 observers

# Other action lines

- Key Performance Indicators (KPIs):
  - Definition and metrics for first set of indicators in place
  - Narratives will remain part of the assessment
  - Future monitoring will call for more focus on how to measure the enabling of scientific excellence
- New governance further implemented (e.g., bylaws for committees)
- Active contribution to the discussion regarding the EU landscape of Research Infrastructures and the establishment of the European Open Science Cloud (EOSC).

# National Coordinators' Forum (NCF)



NCF Chair **Maciej Piasecki**



NCF Vice-chair **Krister Lindén**

# Strategy and Management Board (SAMBA)



**Erhard Hinrichs**



**Maciej Piasecki**



**Krister Lindén**



**Kiril Simov**



**Jan Odijk**



**Kadri Vider**





## **Bente Maegaard**

*former vice executive director*  
2012-2019