

## Tour de CLARIN

[Tour de CLARIN](#) is the name of the successful initiative that highlights prominent user involvement activities of CLARIN national consortia and K-centres. In 2020, outstanding resources, tools, events and researchers from the following countries and K-centres were featured:

**Countries:** Norway, UK, Germany, Iceland, France, South Africa.

**K-centres:** IMPACT-CKC (Spain), the K-centre for Polish Language Technology (Poland), the Phonogrammarchiv of the Austrian Academy of Science (Austria), the K-centre for Atypical Communication Expertise (ACE) (The Netherlands), the Lund University of Humanities Lab (Finland), and the Spanish CLARIN K-centre (Spain).

## CLARIN Café

In May 2020, the [CLARIN Café](#) meeting format was launched: an informal and interactive space for discussion where researchers, lecturers, students and experts can meet to share experiences and insights that are relevant to the CLARIN community. Four more Cafés were held over the course of 2020. The CLARIN Café series started as a response to the lack of interactive spaces during the COVID-19 outbreak, and it soon became very popular with a broad range of audiences. Thanks to its informal setting, the Café sessions can easily be attended by both newcomers and audiences that are already familiar with CLARIN. CLARIN Cafés have become a regular element in the CLARIN events calendar.

## CLARIN Resource Families

In 2020, increasing emphasis was put on activities that stimulate the uptake of the research infrastructure in the relevant research communities. Continuing the work that was taken up in the previous years, special attention was given to improving the visibility and further extending CLARIN's [Resource Families](#), which has yielded comprehensive overviews of 3 types of resources: families of corpora (10), families of lexical resources (5), and families of tools (3).

## CLARIN Responses to COVID-19

Over the course of 2020, CLARIN initiated several action lines in response to the COVID-19 outbreak, ranging from support for online events, to the enabling of COVID research in the social sciences and humanities domain, which aimed at capturing the social and cultural dynamics of the pandemic.

CLARIN funded the project [ParlaMint](#), which focuses on extending the existing curated collections of parliamentary data with debates on the Corona dynamics. The enhanced datasets have become the basis for comparative research into the ways in which public bodies across Europe have responded to the crisis.

With the organisation of - inter alia - the [CLARIN Annual Conference](#) and the [CLARIN Café](#) series as virtual events, the CLARIN Office staff have become deeply familiar with the best practices of arranging online events and the use of platforms for virtual meetings. This experience has been put to the service of the community by setting up a support scheme for [virtual events](#), for organisers affiliated with CLARIN.

Although the working conditions in 2020 were challenging for many people, the pandemic has also brought the insight that virtual events can be more inclusive and reach larger audiences.

## CLARIN in a Nutshell

Name	CLARIN is short for "Common Language Resources and Technology Infrastructure".
Vision	All digital language resources and tools from all over Europe and beyond are accessible through a single sign-on online environment, to support researchers in the humanities and social sciences.
Mission	Create and maintain an infrastructure to support the sharing, use and sustainability of language data and tools for research in the humanities and social sciences.

### Value Proposition

CLARIN is a networked federation of language data repositories, service centres and centres of expertise. CLARIN makes digital language resources available to scholars, researchers, students, and citizen-scientists from all disciplines, especially in the Social Sciences and Humanities (SSH), through single sign-on access. CLARIN offers long-term solutions and technology services for deploying, connecting, analysing and sustaining digital language data and tools. CLARIN supports scholars who want to engage in cutting-edge, data-driven research, contributing to a truly multilingual European Research Area.

### Governance

The **General Assembly** represents the members of CLARIN ERIC and is the highest decision-making body of CLARIN ERIC. It is assisted by an international **Scientific Advisory Board**. The day-to-day management is in the hands of the **Board of Directors** chaired by the **executive director** Prof. Franciska de Jong and supported by the **CLARIN Office**.

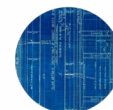
The largest effort towards the further integration of data, tools and expertise stems from the activities in the national consortia. The **National Coordinators' Forum** is responsible for the coordination of the collaboration across countries.

CLARIN ERIC  
c/o Utrecht University  
Drift 10  
3512 BS Utrecht  
The Netherlands

[www.clarin.eu](http://www.clarin.eu)  
[clarin@clarin.eu](mailto:clarin@clarin.eu)

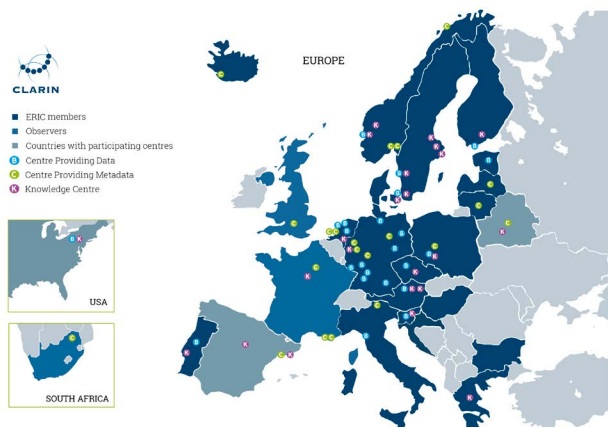


# ANNUAL REPORT 2020



## CLARIN ERIC Members

In 2020, Iceland joined CLARIN ERIC as a member. At the end of 2020, the consortium composition consisted of **21 members**: AT, BG, CY, CZ, DE, DK, EE, FI, GR, HR, HU, IS, IT, LT, LV, NL, NO, PL, PT, SE, SI; **3 observers**: UK, FR, SA; and **1 third party**: Carnegie Mellon University (USA).



## Annual Conference 2020

5 - 7 October 2020, virtual event

The [Annual conference](#) took place in October 2020 as a virtual event. As in previous years, the majority of the conference programme was dedicated to (peer-reviewed) papers and posters that allowed presenters to show their achievements and to collect valuable feedback from the community. New elements in the 2020 programme were the session on the use of CLARIN resources in the classroom and a topical panel session. The virtual format enabled CLARIN to share quality content with almost 500 registered participants, including both attendants of previous events as well as new participants with an interest in finding out what CLARIN is about.

## CLARIN in EU-Funded Projects

In 2020, CLARIN ERIC continued its active participation in several European and international initiatives. These projects bring intensified collaboration with other European research infrastructures and non-academic parties, including companies and cultural heritage institutes. In 2020, CLARIN was active in the following collaborative infrastructural projects: [ENRIITC](#), [EOSC-hub](#), [ERIC Forum](#), [EUROPEANA-DSI4](#), [SSHOC](#), and [TRIPLE](#).

## CLARIN and EOSC

CLARIN ERIC has joined the [EOSC](#) Association as a member. EOSC is a federation of existing infrastructures that aims to offer a virtual environment with services for the storage, management, analysis and re-use of research data, across borders and scientific disciplines. For the CLARIN contribution as a thematic service provider to the EOSC: see this [video](#). It illustrates how several language processing services and thousands of language resources have been made interoperable and accessible through the EOSC portal.

## Events

Due to the COVID-19 pandemic, most of the [events](#) organized in 2020 took place in a virtual format. A total of 264 events, such as workshops, tutorials, seminars and masterclasses, were offered with support from the CLARIN network, attracting more than 8925 participants. The highlights include:



[Twin Talks II and III: Understanding and Facilitating Collaboration in DH](#). These workshops were the second and third in a series where the main objective was a better insight into the dynamics on the Digital Humanities work floor, where humanities scholars and digital experts meet and work together to solve humanities research questions.

[ParlaCLARIN II: Creating, Using and Linking Parliamentary Corpora with Other Types of Political Discourse](#). The second ParlaCLARIN workshop brought together developers, curators and researchers of regional, national and international parliamentary debates that are suitable for research in disciplines in the Humanities and Social Sciences.

## CLARIN Publications

Among the publication output of 2020, the following ones are worth highlighting:

- Popel, M., Tomkova, M., Tomek, J. et al. [Transforming machine translation: a deep learning system reaches news translation quality comparable to human professionals](#). In: Nature Communications 11, 4381 (2020).
- De Jong, F., Maegaard, B., Fišer, D., Van Uytvanck, D., & Witt, A. (2020). [Interoperability in an Infrastructure Enabling Multidisciplinary Research: The case of CLARIN](#). In: Proceedings of LREC 2020, 12th International Conference on Language Resources and Evaluation, ELRA.
- Fišer, D., Pahor de Maiti, K. (2020): [Voices of the Parliament](#). In: Modern Languages Open, (1), p.46.

## Technical Developments

### CLARIN CENTRES

In 2020, one B-centre was certified and one centre was re-certified. By the end of 2020, there were 25 certified CLARIN B-centres (two more than in 2019) and the total number of registered centres grew from 57 to 69.

### FEDERATED IDENTITY

23 out of 24 member/observers connected to the Service Provider Federation, bringing the total number of organisations that can login to 1960 (previously 1870).

### VIRTUAL LANGUAGE OBSERVATORY

Two new versions were launched, resulting in improved record visualization and schema.org-compliant embedded metadata, compatible with Google's Dataset Search.

### LANGUAGE RESOURCE SWITCHBOARD

The file format profiler was improved and now supports TEI, TFC, and CONNLU. Both the tool inventory as well as the documentation underwent a major overhaul.

## Knowledge Infrastructure

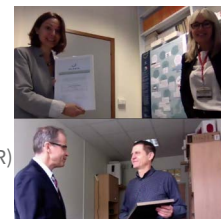
The number of certified nodes in the CLARIN [Knowledge Infrastructure](#) (K-centres) has gone up from 20 to 23 in 2020. In 2020, the [first workshop for K-centres](#) was held. The purpose of the workshop was to explore the options for: creating more external visibility of the K-centre network and their individual service offer reinforcing the collaboration between the K-centres and their central support.

In the context of the H2020 [SSH Open Cloud project](#), CLARIN organized three training webinars on CLARIN tools, resources and methods: [Hands-on Tutorial on Transcribing Interview Data: Quantify with Ease: Combining Quantitative and Qualitative Corpus Analysis](#); [Sharing Datasets of Pathological Speech](#)

## Awards

At the virtual CLARIN Annual Conference, the Steven Krauwer Awards for CLARIN Achievements in 2020 were handed out to:

[Francesca Frontini](#) (Istituto di Linguistica Computazionale "A. Zampolli" / CLARIN-IT & Université Paul-Valéry Montpellier 3 / CLARIN-FR) and [Tomasz Walkowiak](#) (Wrocław University of Science and Technology / CLARIN-PL).



## Finances

BALANCE 31 December	2020	2019
<i>Assets</i>		
Fixed assets		
equipment	€14.352	€10.260
Current assets		
receivables	€805.036	€848.289
bank accounts	€2.232.269	€1.636.607
<b>Total Assets</b>	<b>€3.051.657</b>	<b>€2.495.156</b>
<i>Equity and liabilities</i>		
capital and reserves	€1.870.347	€1.579.384
current liabilities	€1.181.310	€915.772
<b>Total Equity and liabilities</b>	<b>€3.051.657</b>	<b>€2.495.156</b>
PROFIT and LOSS	2020	2019
<i>Income</i>		
membership fees	€1.170.309	€1.151.048
projects/other	-	€93.932
<b>Total Income</b>	<b>€1.170.309</b>	<b>€1.244.980</b>
<i>Expenditures</i>		
personnel and travel costs	€817.823	€821.100
other	€387.054	€373.919
<b>Total Expenditures</b>	<b>€1.209.560</b>	<b>€1.195.020</b>
<b>RESULT</b>	<b>€290.964</b>	<b>€49.960</b>

The accounts for the participation by CLARIN ERIC and other CLARIN nodes in projects funded by EU programmes are not included in this overview.