

Tour de CLARIN

[Tour de CLARIN](#) is the name of the successful initiative that highlights prominent user involvement activities of CLARIN's national consortia as well as K-centres and B-centres. In 2021, outstanding resources, tools, events and researchers from the following countries and centres were featured: **Countries:** Portugal. **K-centres:** SAFMORIL, the Knowledge Centre for Systems and Frameworks for Morphologically Rich Languages; CORLI, the French Knowledge Centre for Corpora, Languages, and Interaction; CLASSLA, the Knowledge Centre for South Slavic Languages; NLP:EL, the Knowledge Centre for Greek. **B-centres:** ARCHE, the Austrian B-Centre for Digital Humanities and Cultural Heritage; The CLARIN-PL B-Centre.

CLARIN Cafés

The [CLARIN Café](#) format, which was introduced in 2020, has developed into a mature platform facilitating informal exchange within the CLARIN community. In 2021, 10 CLARIN Cafés were (co-) organised by the CLARIN Office, with a total of 572 participants. Examples of the wide range of topics covered include the [CLARIN Café - ParlaMint Unleashed](#) (June 2021) and the [CLARIN Café on Text and Data Mining Exceptions in the Directive on Copyright in the Digital Single Market](#) (October 2021). The majority of participants came from Europe (16.7% Italy, 11% Netherlands, 9% Poland, 6% Germany, 5.7% France, 5.4% Czechia).

CLARIN Impact Stories

This new series showcases high-quality and innovative research that uses CLARIN tools and resources. The [Impact Stories](#) illustrate the huge variety of disciplines that use the CLARIN infrastructure, highlight the excellent research linked to it, and demonstrate the wider impact that CLARIN and the social sciences and humanities have on broader societal issues. The initiative was launched in the autumn of 2021. The first Impact Stories focused on the datasets of the CLARIN flagship project [ParlaMint](#) during the Helsinki Humanities Hackathon 2021 and the 'Stories in Motion' project at the Erasmus University Rotterdam, which has developed a model for archiving and (re)using oral histories in order to extend the impact for such collections.

CLARIN Funding

With mobility and face-to-face events on hold due to the COVID-19 pandemic, the [funding](#) for CLARIN activities concentrated on support for virtual events organised by the nodes in the network and support for the further development of resources and collaborations within the infrastructure. Two main instruments were utilised for this, the [CLARIN Resource Families project call](#) and the [Support for Virtual Events call](#).

CLARIN Ambassadors

The aim of the [CLARIN Ambassadors Programme](#) is to raise awareness of and encourage participation in CLARIN ERIC in disciplines and communities that are not yet fully integrated in CLARIN. In 2021, three experts were appointed CLARIN Ambassadors: **Eva Soroli** is Associate Professor for Psycholinguistics at the University of Lille, France; **Paul Rayson** is a Professor for Computer Science at Lancaster University, UK, and Director of the UCREL; **Satu Saalasti** is a University lecturer in Speech and Language Pathology at the University of Helsinki, Finland.

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, 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**.

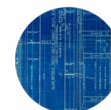
The largest effort towards the further integration of data, tools and expertise stems from the activities in the national consortia. The **National Coordinators' Forum** is responsible for the coordination of the collaboration across countries.

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

www.clarin.eu
clarin@clarin.eu

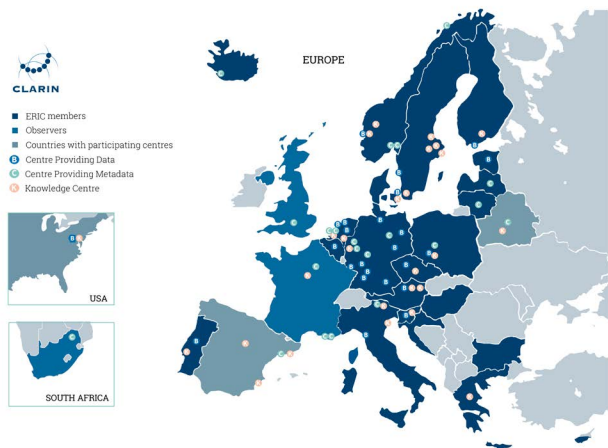


ANNUAL REPORT 2021



CLARIN ERIC Members

In 2021, Belgium joined CLARIN ERIC as a member. At the end of 2021, the consortium consisted of **22 members**: AT, BE, 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 2021

27 - 29 September 2021, virtual event

The [CLARIN Annual Conference](#) took place in September 2021 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. One of the highlights was the panel on [‘The Role of Corpora for the Study of Language Use and Mental Health Conditions’](#). The virtual format made it possible for almost 500 registered participants to attend the event, a significantly higher rate of attendance than at previous editions of the Annual Conference.

CLARIN in EU-Funded Projects

In 2021, CLARIN ERIC continued its active participation in several European and international initiatives. Such projects bring intensified collaboration with other European research infrastructures and non-academic parties, including companies and cultural heritage institutions. In 2021, CLARIN was active in the following collaborative infrastructural projects: [ELE](#), [EOSC Future](#), [ENRIITC](#), [ERIC Forum](#), [EUROPEANA-DSI4](#), [SSHOC](#), [TRIPLE](#) and in the Erasmus+ project [UPSKILLS](#).

The project activities significantly strengthen CLARIN’s multi-disciplinary perspective and its collaboration with other initiatives in the European and wider ecosystem of research infrastructures. Reinforced synergy in particular was realised in the cluster of research infrastructures from the domain of the social sciences and humanities that are collaborating in the SSH Open Cloud (SSHOC) project.

In the context of the H2020 SSH Open Cloud project, CLARIN organised two training webinars on CLARIN tools, resources and methods:

- SSHOC Workshop: [Data Citation in Practice](#)
- SSHOC Workshop: [Showcase your Survey for Free on the EMM Survey Registry](#).

Events

Due to the COVID-19 pandemic, most of the [events](#) organised in 2021 took place virtually. A total of 253 events, including workshops, tutorials, seminars and masterclasses, were offered with support from the CLARIN network, attracting more than 11.000 participants. The highlights include:

- The first CLARIN:EL Summer School ‘Language Technology and Social Sciences & Humanities: a new research methodology’
- The Helsinki Digital Humanities Hackathon #DHH21, where the [ParlaMint](#) dataset was used to analyse the response to the COVID-19 pandemic in four different European countries
- The webinar [SSHOC’ing Drama in the Cloud](#), with the topic of encoding theatrical text collections and the added value of SSHOC and CLARIN services at the [LIBER](#) conference
- The [DELAD workshop](#) on sharing corpora of Speech with Communication Disorders (CSD).



CLARIN Publications

Among the publication output of 2021, the following ones are worth highlighting:

- Fitsilis, F., & Mikros, G. (2021). Development and Validation of a Corpus of Written Parliamentary Questions in the Hellenic Parliament. *Journal of Open Humanities Data*, 7, 18. [DOI](#)
- Gruszczyski, W., Adamiec, D., Bronikowska, R. et al. The Electronic Corpus of 17th- and 18th-century Polish Texts. *Lang Resources & Evaluation* 56, 309–332 (2022). [DOI](#)
- Cinelli, M., Pelicon, A., Mozetič, I. et al. Dynamics of online hate and misinformation. *Sci Rep* 11, 22083 (2021). [DOI](#)

Technical Developments

CLARIN CENTRES

In 2021, two B-centres were re-certified. By the end of 2021, there were 24 certified CLARIN B-centres. The total number of registered CLARIN centres is 68.

VIRTUAL LANGUAGE OBSERVATORY

A new version was released, with new features such as a temporal coverage facet. In addition, the VLO has been migrated to a significantly faster server.

LANGUAGE RESOURCE SWITCHBOARD

A new version of the Switchboard was released, bringing improvements to the user interface as well as several changes to the file format recognition. More tools were connected to the Language Resource Switchboard. All connected tools are now visible in the SSH Open Marketplace.

Knowledge Infrastructure

The number of certified nodes in the [CLARIN Knowledge Infrastructure](#) (K-centres) went up from 23 to 25 in 2021. The K-centre workshop that was held in mid-December 2021 received a positive response overall and stimulated ideas for future collaboration between the various K-centres.

Awards

Steven Krauwer Award

The eighth award ceremony for the Steven Krauwer Award took place during the CLARIN Annual Conference 2021. The Steven Krauwer Award for CLARIN Achievements in 2021 was awarded to [Tomaž Erjavec](#) (Jožef Stefan Institute/[CLARIN.SI](#)) for his outstanding contributions towards CLARIN goals.



Teaching with CLARIN Award

The [Teaching with CLARIN Award](#) is given to lecturers or other educational professionals who have successfully integrated CLARIN resources into the courses they teach and/or for training material they developed and submitted in response to the [Teaching CLARIN in Times of Corona call](#). The awards were handed out during the CLARIN Annual Conference to Darja Fišer & Kristina Pahor de Maiti, and Mieta Lennes.

Finances

BALANCE 31 December	2021	2020
<i>Assets</i>		
Fixed assets		
Equipment	€9.218	€14.353
Current assets		
Receivables	€349.594	€805.036
Bank accounts	€2.433.250	€2.232.269
Total Assets	€2.792.062	€3.051.658
<i>Equity and liabilities</i>		
Capital and reserves	€1.882.744	€1.870.347
Current liabilities	€909.318	€1.181.311
Total Equity and Liabilities	€2.792.062	€3.051.658

PROFIT and LOSS	2021	2020
<i>Income</i>		
Membership fees	€1.302.307	€1.170.309
Projects/other	€213.368	€330.686
Total Income	€1.515.675	€1.500.995
<i>Expenditures</i>		
Personnel and travel costs	€978.710	€817.823
Other	€524.568	€392.209
Total Expenditures	€1.503.278	€1.210.032
RESULT	€12.397	€290.964

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