

Disseminating and Teaching CLARIN in the time of Corona. Challenges and opportunities

Francesca Frontini
CLARIN ERIC & ILC-CNR

Latvian Language Digital Resources And Tools For Education

June 2, 2021



June 2nd 1946

Outline

- What is CLARIN
- What is the CLARIN Knowledge infrastructure
- What it offers to researchers ...
- ... and teachers, lecturers and educators

CLARIN in eight bullets

- **CLARIN** is the Common Language Resources and Technology Infrastructure
- **ESFRI** ERIC status since 2012, Landmark since 2016
- that provides easy and sustainable access for scholars in the **humanities and social sciences** and beyond
- to **digital language data** (in written, spoken, video or multimodal form)
- and **advanced tools** to discover, explore, exploit, annotate, analyse or combine them, wherever they are located
- through a **single sign-on** environment
- that serves as an ecosystem for **knowledge exchange**
- and: ready for **integration in EOSC** (European Open Science Cloud; [link](#))

CLARIN Value Proposition: <https://www.clarin.eu/content/value-proposition> (link to [pdf](#))

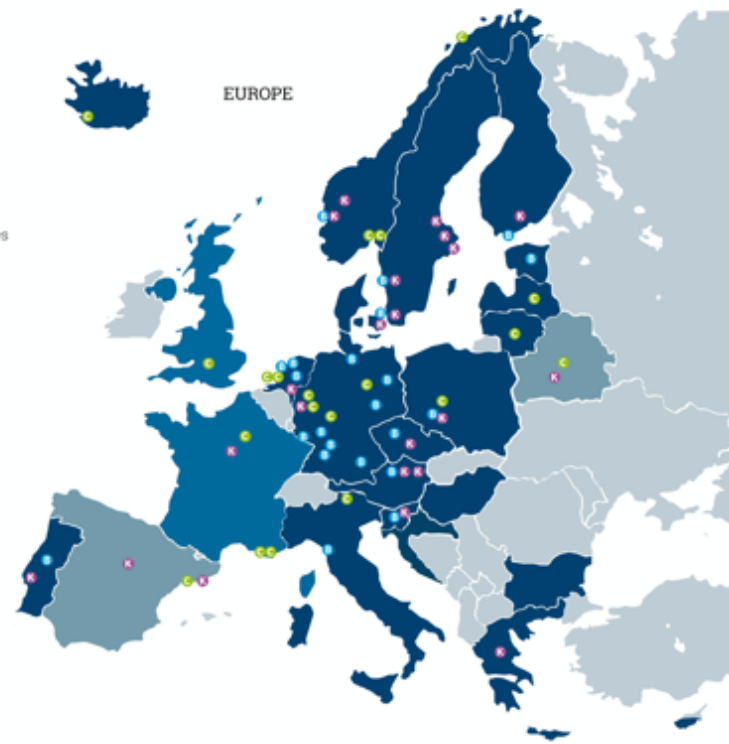
CLARIN ERIC in members and centres

A consortium 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: FR, UK, ZA
- 1 linked party: CMU
- >60 centres
(incl. 24 CTS certified data centres)
- Support for linguistic diversity
 - Data covering more than **1500** languages
 - Tools for many languages
 - Language resources in all modalities



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

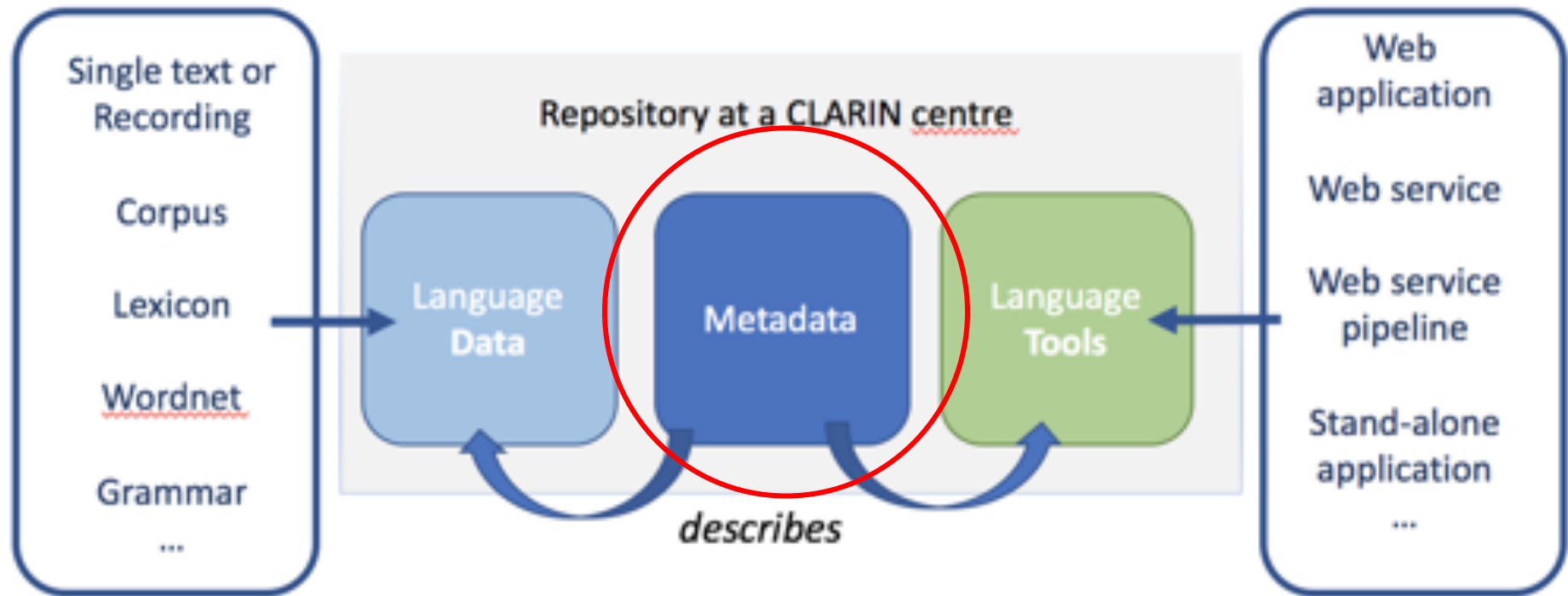


CLARIN data types and communities of use in SSH and beyond

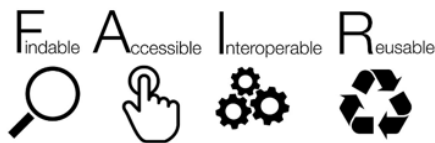
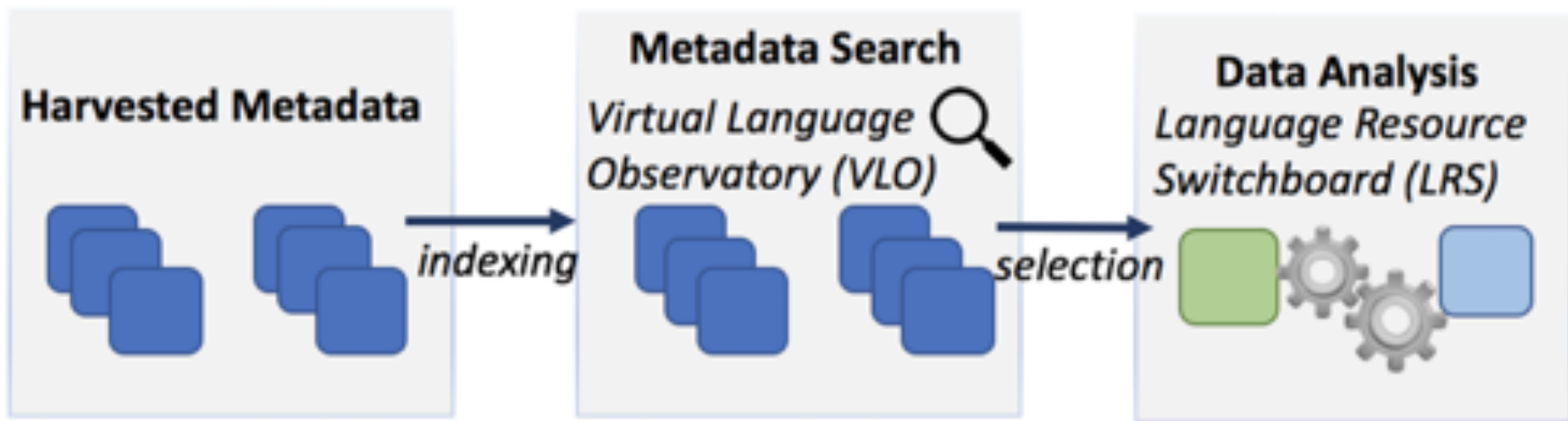
- Newspaper archives
- Parliamentary records
- Literary texts
- Historical letters
- Broadcast archives
- Oral History data
- Social Media data
- L-2 Learner Resources
- Survey data
- Patient recordings
- Excavation reports

- Digital humanities
- Linguistics and Philology
- Data Science /AI
- Translation and Lexicography
- Literary Studies
- History
- Political and Social Sciences
- Media Studies
- Culture, Folklore, Anthropology
- Speech therapy
- General Public

The CLARIN data architecture: *a network of distributed centre repositories*



The CLARIN data architecture: *central processing of metadata*



clarin.eu/fair



vlo.clarin.eu



switchboard.clarin.eu

Search through 1,678 records

Showing 1 to 10 of 1,614 results within selection for Latvian x No x

Use the categories below to limit the search results to those matching the selected value(s).

Language A

latvian

Latvian x

more...

Collection W

Resource type W

Modality W

Format W

Multilingual A

1 2 3 4 5 6 7 8 9 10

Latvian Treebank v2.5

(Part of Language resources and tools of Ailab IMCS UL)

Latvian Treebank (LVTB) is in development since 2010. Corpus is SemT1-Kamola hybrid dependency-constituency grammar model.1 data used for deriving LVTB version 2.5.

Latvian

Landing page for this record

The Corpus of Early Written Latvian

(Part of Language resources and tools of Ailab IMCS UL)

The Corpus of early written Latvian "SENIE"(16th-18th cc.) provk Latvian of the 16th-18th century, and its aim is to facilitate studies lexis, morphology and syntax of the texts) and to serve as the basis Latvia...

Latvian

Landing page for this record



treebank

Search

Advanced Search

Limit your search

Author

Subject

Showing 1 through 2 out of 2 results

1

Corpus

CLARIN Centre Of Latvian Language Resources And Tools

Latvian Treebank v2.5

(Ailab IMCS UL / 2020-02-01)

Author(s):

Rituma, Laura ; Pretkalniņa, Lauma ; Saulīte, Baiba ; Nešpore-Bērzkalne, Gunta ; Grūziņš, Normunds

This item contains 1 file (16.44 MB)

Publicly Available

vlo.clarin.eu

Resources

latvian-0287.txt 360.21 KiB 

Mediatype

text/plain

Language

Latvian

Matching Tools

Group by task ☒

Search for tool

Dependency Parsing



Open 

UDPipe

Distant Reading



Open 

Voyant Tools

Text Analytics



Open 

WebLicht Advanced Mode 

switchboard.clarin.eu

CLARIN as ecosystem for knowledge exchange



CLARIN Knowledge Centres (abbreviated K-centres) are a cornerstone of the CLARIN knowledge infrastructure. K-centres are institutions that have agreed to share their knowledge and expertise on one or more aspects of the domain covered by the CLARIN infrastructure.

CLARIN as ecosystem for knowledge exchange



- [Individual languages](#) (e.g. Danish, Czech, Portuguese), language families (e.g. South Slavic) or groups of languages (e.g. morphologically rich languages, the languages of Sweden)
- [Written text and modalities other than written text](#) (e.g. spoken language, sign language)
- [Linguistic topics](#) (e.g. language diversity, language learning, diachronic studies)
- [Language processing topics](#) (e.g. speech analysis, building treebanks, machine translation)
- [Data types other than corpora](#) (e.g. lexical data, word nets, terminology banks)
- [Using or processing families of language data](#) that will exist for most languages (e.g. newspapers, parliamentary records, oral history)
- [Generic methods and issues](#) (e.g. data management, ethics, IPR, OCR)

CLARIN as ecosystem for knowledge exchange

A dark blue rectangular banner with a background of faint, light blue binary code (0s and 1s). The text "Digital Humanities Course Registry" is written in a bold, white, sans-serif font, centered horizontally.

Digital Humanities Course Registry

In collaboration with DARIAH ERIC.
The DH Course Registry contains a selection
of DH courses offered by European academic
organizations.

DIGITAL HUMANITIES COURSE REGISTRY

[FILTER](#)
[SORTING](#)

NAME	UNIVERSITY	PLACE	DATE	TYPE
Digital Humanities	Universität Basel	Basel, Switzerland	20 Sept 2 years	Master Programme
Data Science with Minor in Digital Humanities	Friedrich-Alexander-Universität Erlangen-Nürnberg	Erlangen, Germany	18 Oct 6 semester	Bachelor Programme
Master Mittelalter und Frühe Neuzeit, Vertiefung Digital Humanities	Friedrich-Alexander-Universität Erlangen-Nürnberg	Erlangen, Germany	18 Oct 4 semester	Master Programme
Digital Humanities	Friedrich-Alexander-Universität Erlangen-Nürnberg	Erlangen, Germany	15 Oct 4 semester	Master Programme
Digitale Geistes- und Sozialwissenschaften	Friedrich-Alexander-Universität Erlangen-Nürnberg	Erlangen, Germany	1 Oct 6 semester	Bachelor Programme
Linguistic and Literary Computing	Technische Universität Darmstadt	Darmstadt, Germany	1 Oct 4 semester	Master Programme
Introduction to Computational Linguistics (DHS)	University of Latvia	Riga, Latvia	5 Sept 2021 1 semester	Course
Cultural Data Studies	Philipps-Universität Marburg	Marburg, Germany	1 Oct 4 semester	Master Programme
Digital Humanities and Social Sciences	University of Latvia	Riga, Latvia	1 Sept 2021 1 semester	Course
Fakultätsweites Basismodul: Digitale Geisteswissenschaften	Karl-Franzens-Universität Graz	Graz, Austria	3 Oct 1 semester	Course
Doktoratsstudium "Digitale		Graz,	1 Oct	


[STORIES](#)
[INFO](#)
[LOGIN](#)


Imprint

©2016-2021 CC-BY 4.0

CLARIN as ecosystem for knowledge exchange

VideoLectures

HOME • BROWSE LECTURES • PEOPLE • CONFERENCES • ACADEMIC ORGANISATIONS • EU SUPPORTED • ABOUT US

Project: Academic Organisations • CLARIN - Common Language Resources and Technology Infrastructure

View order

- Overview
- Hot
- Popular
- Just published
- Recent
- Top Voted

ParlaCLARIN Workshop: Creating and Using Parliamentary Corpora, Miyazaki 2018

ParlaCLARIN Workshop 2018 - Miyazaki

CLARIN - COMMON LANGUAGE RESOURCES AND TECHNOLOGY INFRASTRUCTURE

RECENT EVENTS MORE...

CLARIN Annual Conference 2019 - Leipzig

8th CLARIN Annual Conference, Leipzig 2019

The CLARIN Annual Conference is the main annual event for those working on the construction and operation of CLARIN across ...

Essentials of Coding and Encoding

Baltic Summer School of Digital Humanities

Essentials of Coding and Encoding

Watching Now

This page - CLARIN - Common Language Resources and Technology Infrastructure

Category: Computers

Voted: 100% NET

International Center for Advanced

Course on Information Theory, Pattern

PhD Thesis Defenses: Dynamics of large

CLARIN stands for "Common Language Resources and Technology Infrastructure".

It is a research infrastructure that was initiated from the vision that all digital language resources and tools from all over Europe and beyond are accessible through a single sign-on online environment for the support of researchers in the humanities and social sciences.

In 2012 CLARIN ERIC was established and took up the mission to 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. Currently CLARIN provides easy and sustainable access to **digital language data** (in written, spoken, or multimodal form) for scholars in the **social sciences and humanities**, and beyond. CLARIN also offers advanced tools to discover, explore, exploit, annotate, analyse or combine such data sets, wherever they are located. This is enabled through a networked federation of centres: language data repositories, service centres and knowledge centres, with single sign-on access for all members of the academic community in all participating countries. Tools and data from different centres are interoperable, so that data collections can be combined and tools from different sources

CLARIN as ecosystem for knowledge exchange



Get to know the different types of resources in different languages CLARIN makes available for researchers from digital humanities, social sciences and human language technologies.

<https://www.clarin.eu/resource-families>

Dictionaries in the CLARIN infrastructure

Size: 83,000 entries		download and for online browsing through the Language Bank of Finland.	Browse Download
Licence: CC-BY 4.0			
Dictionary of Modern Latvian (MDW)	Latvian	The resource is available for online browsing.	Browse
Size: 45,877 entries			
Licence: browsable			
Dictionary of Standard Latvian Language (LSW)	Latvian	This is a dictionary that contains examples of the standard Latvian language from the 1870s to the 1990s. The resource is available for online browsing.	Browse
Size: 64,200 entries			
Licence: browsable			
Historical Dictionary of Latvian (16th-17th Centuries) (DWW)	Latvian	This is a dictionary of Latvian from the period between the 16th and 17th centuries with a reference to the "Historical Corpus of the Latvian Language". The resource is available for online browsing.	Browse
Size: 1,800 entries			
Licence: Terms of Use		For a related publication, see Andronova et al. (2016).	
Rendering Dictionary of Personal Names	Latvian	This dictionary contains person names in four languages (Czech, French, Italian and German) and their renderings in Latvian in accordance with the rendering rules of proper nouns. The resource is	Browse
Size: 4,000 entries			
Licence: Terms of Use			

Table of Contents

[Introduction](#)

[Dictionaries in the CLARIN infrastructure](#)

[Monolingual resources](#)

[Multilingual resources](#)

[Publications](#)

CLARIN as ecosystem for knowledge exchange



CLARIN Going Virtual

- 11 CLARIN cafés
- Support for virtual events
- 1 virtual conference (with another coming up)
 - Largest attendance ever!



<https://www.clarin.eu/content/clarin-responses-covid-19>

<https://www.clarin.eu/events>



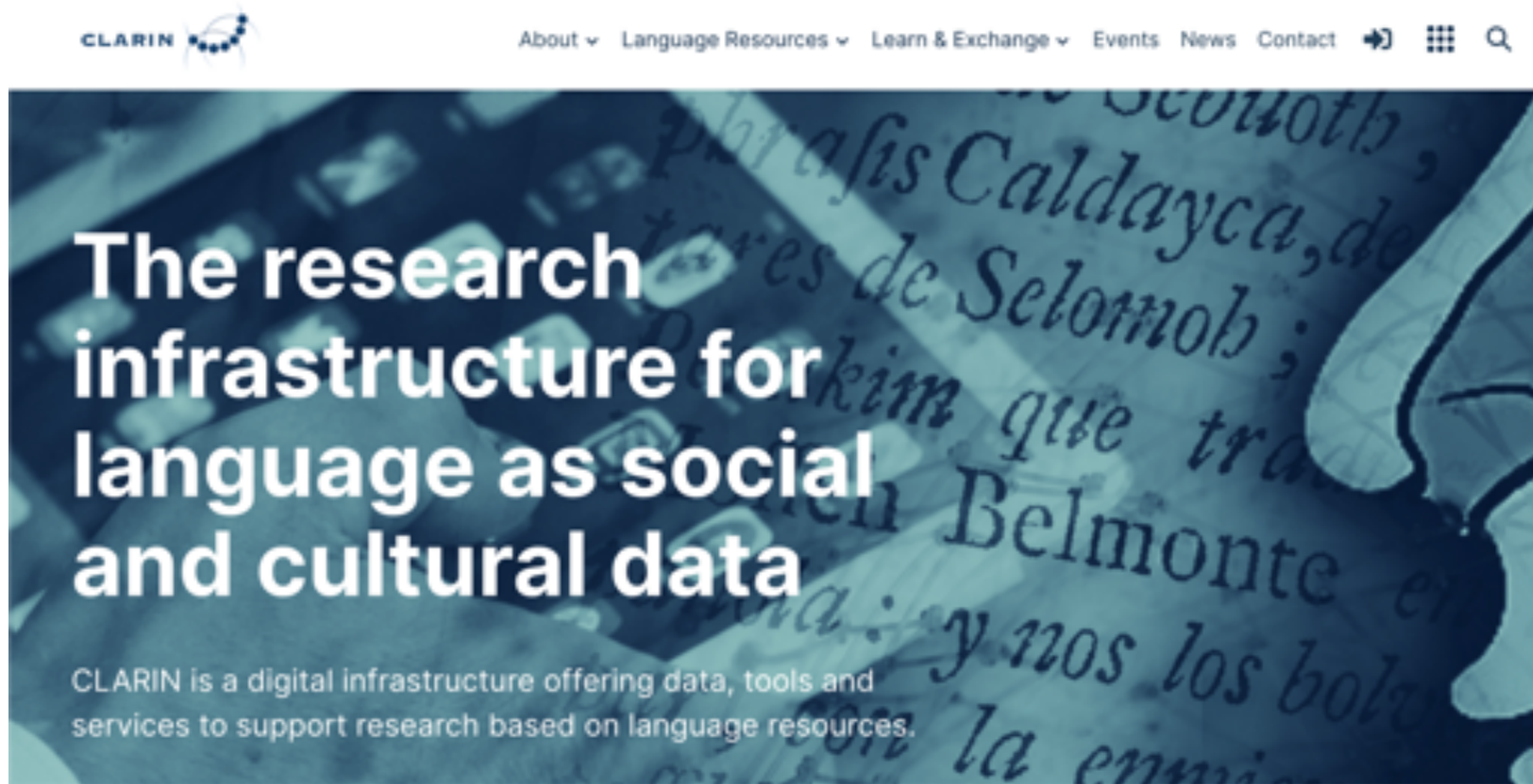
CLARIN  **CLARIN2020 Annual Conference**

Reaching more communities and researchers

- More virtual events and asynchronous access (recordings)
- More opportunities to engage with CLARIN (no limitations due to travel, lack of funding)
 - Lecturers
 - Early career researchers
 - Students



New website, tomorrow!



CLARIN in the classroom

In order for CLARIN to have a lasting impact, it is necessary for both methodological and infrastructural issues to become integrated in university curricula.



CLARIN in the classroom

Survey 22 member and observer countries
(2008-2020)

CLARIN used in 55 individual courses, already
taught at 31 different universities in 20
countries



What teachers think of CLARIN - Advantages

- Accessibility
 - *"The interfaces are usable, reliable, state of the art; students like the idea that they can access services that are hosted in other countries with their university identifier."*
- Findability of resources
 - *"Search the VLO for concrete corpora and to show the power of well curated metadata."*
- Reuse of resources
 - *"Students use CLARIN corpora and tools in their theses, projects and internships with companies."*
- Archiving of teaching materials
 - *"The integration of CLARIN content in teaching motivated some of the lecturers to archive their materials and make them accessible to the research community and the general public."*
- Support for research activities
 - *"A student who explored Fin-CLARIN said that it is paradise for researchers."*

Reusable, integrated online services from the Switchboard

The screenshot displays the Weblicht interface, which is used for linguistic analysis. At the top, there are tabs for 'Main Page', 'Chain 1', and '+ New Chain'. Below the tabs, the 'Available Annotations for Plain Text French' section lists 'POS Tags/Lemmas'. The 'Query' section contains a search bar with the placeholder text 'Enter either a TIGERSearch query, or simply a word in quotation marks.' and buttons for 'History' and 'Query Language Help'. The 'Visualization' section shows a sentence: 'Ce qui s' est passé au cours de ces dernières années a attiré notre attention sur les terribles problèmes auxquels sont confrontés non seulement les producteurs d' aliments , mais également les consommateurs .' with a 'Sentence' label and 'out of 16'. Below the sentence, there is a visualization of the sentence structure with various linguistic annotations. At the bottom, the 'Input and Chain Selection' section shows a table with columns for 'Example Food/chain 1', 'SIS to TIG Converter', 'SIS to TIG', and 'SIS to TIG/Tagger'. The table contains data for 'Document Type: TIG', 'TIG version: 5', 'Text', and 'Language: French'. There are also buttons for 'Run Tools', 'Clear Results', and 'Download chain'.

Weblicht, a lifeline when courses moved online.

Single sign on, no need to install software

 Sign in via the CLARIN Service Provider Federation



Select your home organisation below. This is usually the organisation where you work or study. Signing in here will allow you to access certain CLARIN resources and services which are only available to users who have logged in. If you cannot find your organisation in the list below, please select the clarin.eu website account and use your CLARIN website credentials. If you don't have such credentials you can register an account [here](#). For questions please contact spf@clarin.eu.

Previously chosen home organisation

Université de Montpellier 3
 France

Home organisation list



All countries



clarin.eu website account



European Union

AAU@EduHr Single Sign-On Service



Croatia



Aalborg University



Denmark

Key issues

- Lack of resources and/or tools for a specific language
 - *"Lack of resources for French, in particular of corpora and NLP modules."*
- Lack of standardization
 - *"Lack of standardization that hinders reuse of resources and tools."*
- Lack of teacher expertise
 - *"There is a potential to use lots of other tools that I am insufficiently familiar with. There is insufficient documentation of available resources and no tutorial on how to effectively use them in the classroom."*
- Lack of training materials
 - *"Lack of time and other resources to locate or create materials for all potentially relevant technologies."*
- Too difficult for students
 - *"Many corpora are in TEI, and there are not many tools which allow graphical exploration or analysis of TEI corpora."*

Working towards a CLARIN Training Suite



First workshop – November 2019

[Home](#) > [Workshop: CLARIN@universities](#)

Workshop: CLARIN@universities

Monday, November 4, 2019 - 12:00 to Tuesday, November 5, 2019 - 13:30

The [CLARIN@universities](#) workshop will take place on 4 and 5 November 2019 in Utrecht (lunch-to-lunch).

This workshop on the experience with the integration of CLARIN content into university programmes is organised by the [Knowledge Infrastructure Committee](#).

[Background](#) | [Approach](#) | [Workshop participants](#) | [Workshop programme](#) | [Workshop materials](#) | [Getting there](#)

Important note for those travelling from Schiphol to Utrecht between 1 and 3 November: Between Friday 1 November and Sunday 3 November, the railway company (NS) is working on the tracks, and the direct train connection from Schiphol to Utrecht Central Station cannot be taken. Instead you need to take a train from Schiphol to Amsterdam Central Station and from Amsterdam Central Station you can take the intercity train to Utrecht Central Station.

CLARIN café on higher education, June 2020

FIN-CLARIN Online courses

Corpus Linguistics and Statistical Methods (5 ECTS)

Introduction to Speech Analysis (5 ECTS)

Data Clinic (5 ECTS)

(in 2020)

<https://www.kielipankki.fi/support/training/>

  **DIGITAL HUMANITIES**
COURSE REGISTRY

29

CLARIN in the Classroom (October 2020)

CLARIN IN THE CLASSROOM

CLARIN ANNUAL CONFERENCE
07-10-2020

Background and Current Activities

COMMUNITY & TOOLS

1 Corpora & Linguistics

ONLINE COURSES
on
DIGITAL RESEARCH METHODS

EX3 PEDAGOGICAL APPLICATIONS of ORVELIT CORPUS

- search resources
- plan projects
- gain knowledge
- conduct research
- report findings
- HELP STUDENTS FIND THEIR WAY

EX1 DATA CLINIC

help students with tools

EX2 Corpus Literacy in German Linguistics

learn to use corpora

EX4 LABLASS

SYSTEM BUILT TO CREATE DICTIONARIES

2 Resources for the Humanities

EX1 MAKING A SPARL TUTORIAL

show potential of special data type

EX2 CLARIN-IT

open science
Fair data
part of community
students are only consumers

EX3 MAKE STUDENTS PRODUCTIVE

show students coding systems

EX4 LEARN TO WORK WITH CLARIN TOOLS

2 GROUPS of students
ELLDO ↔ Computer Linguistics

3 CLASSES:

- 1 CORPUS LINGUISTICS
- 2 DOCUMENTARY LINGUISTICS
- 3 EXPERIMENTAL PHONETICS

3

we to work

THINK IN A RELATED DISCI

UNBELIEV

LEARNING CONTENT

integrating computational into humanities

DIGITAL HUMANITIES HACKATHON in Helsinki

real-life

INTERACTION IN MULTIDISCIPLINARY MODE

johny@antichamber.be

- Teaching of the Computational Linguistics course started in 2003 at Liepāja University:
 - course for the master students in linguistics
 - course for the master students in computer science
 - later Computational Linguistics was included in a course for doctoral students *Novel approaches to Linguistics* at Liepāja University
- Latvia joined CLARIN ERIC in 2016
- CLARIN-LV repository has been registered in March, 2020
- In 2017 new Computational Linguistics course for Master Students in English philology was started
- In Autumn, 2020 Computational Linguistics course started for Master Students in Baltic Philology

Training officer

Introducing Iulianna van der Lek: New Training and Education Officer CLARIN ERIC

CLARIN ERIC is pleased to welcome Iulianna van der Lek as the new Training and Education Officer.

Iulianna will work closely with the CLARIN community to develop the CLARIN Training Suite and the Trainer Network. She will collaborate with university lecturers from the humanities and social sciences disciplines to accelerate the further integration of CLARIN into university curricula. In addition, she is coordinating CLARIN's activities in the UPSKILLS project, an ERASMUS+ initiative that aims at upgrading the skills of Linguistics and Language Students.

[Read more](#)



The Upskills project

<https://upskillsproject.eu/>



Need analysis seeks to develop a detailed professional profile for graduates of linguistics and language-related courses, with the associated learning objectives (knowledge, skills and competences), typical tasks and responsibilities

More comparable resources

Parliamentary corpora in the CLARIN infrastructure

Corpus	Language	Description	Availability
Multilingual comparable corpora of parliamentary debates ParlaMint 2.0 Size: 3.3 million utterances, 462 million words Annotation: tokenised, MSD-tagged, syntactically parsed (Universal Dependencies) Licence: CC BY 4.0	Bulgarian, Croatian, Czech, Danish, Dutch, English, French, Hungarian, Icelandic, Italian, Latvian , Lithuanian, Polish, Slovenian, Spanish, Turkish	<p>ParlaMint is a multilingual set of comparable corpora containing parliamentary debates mostly starting at the end of 2015 and extending to mid-2020, with each corpus being about 20 million words in size. The sessions in the corpora are marked as belonging to the COVID-19 period (after October 2019), or being "reference" (before that date).</p> <p>The corpora have extensive meta-data about the speakers (name, gender, party affiliation, MP status), are structured into time-stamped terms, sessions and meetings, with each speech being marked by its speaker and their role (chair, regular speaker). The speeches also contain marked-up transcriber comments, such as gaps in the transcription, interruptions, applause, etc.</p> <p>The corpus is available for download from the CLARIN SI repository and through the concordancer noSketchEngine.</p>	Concordancer Download
Croatian parliamentary corpus ParlaMeter-hr 1.0 Size: 14.1 million tokens Annotation: tokenised, MSD-tagged, lemmatised, named	Croatian	<p>The corpus contains minutes of the National Assembly of the Republic of Croatia and currently covers its Vth mandate from 15 November 2016 to 21 November 2018. The corpus contains speaker metadata (gender, age, education, party affiliation).</p> <p>The corpus is available for download from the CLARIN SI repository and through the concordancers KonText and</p>	Concordancer Download

Table of Contents

- Introduction**
- Parliamentary corpora in the CLARIN infrastructure
- Other parliamentary corpora
- CLARIN-funded project: ParlaMint
- Additional materials
- Publications on the parliamentary corpora

Parliamentary data are a for comparative study of public debate.

<https://www.clarin.eu/resource-families/parliamentary-corpora>

More comparable resources

- Fill gaps in Resource Families

Parliamentary corpora in the CLARIN infrastructure

Corpus	Language	Description	Availability
Multilingual comparable corpora of parliamentary debates ParlaMint 2.0 Size: 3.3 million utterances, 462 million words Annotation: tokenised, MSD-tagged, syntactically parsed (Universal Dependencies) Licence: CC BY 4.0	Bulgarian, Croatian, Czech, Danish, Dutch, English, French, Hungarian, Icelandic, Italian, Latvian , Lithuanian, Polish, Slovenian, Spanish, Turkish	<p>ParlaMint is a multilingual set of comparable corpora containing parliamentary debates mostly starting at the end of 2015 and extending to mid-2020, with each corpus being about 20 million words in size. The sessions in the corpora are marked as belonging to the COVID-19 period (after October 2019), or being "reference" (before that date).</p> <p>The corpora have extensive meta-data about the speakers (name, gender, party affiliation, MP status), are structured into time-stamped terms, sessions and meetings, with each speech being marked by its speaker and their role (chair, regular speaker). The speeches also contain marked-up transcriber comments, such as gaps in the transcription, interruptions, applause, etc.</p> <p>The corpus is available for download from the CLARIN.SI repository and through the concordancer <i>noSketchEngine</i>.</p>	Concordancer Download
Croatian parliamentary corpus ParlaMeter-hr 1.0 Size: 14.1 million tokens Annotation: tokenised, MSD-tagged, lemmatised, named	Croatian	<p>The corpus contains minutes of the National Assembly of the Republic of Croatia and currently covers its Vth mandate from 15 November 2016 to 21 November 2018. The corpus contains speaker metadata (gender, age, education, party affiliation).</p> <p>The corpus is available for download from the CLARIN.SI repository and through the concordancers <i>KonText</i> and</p>	Concordancer Download

Table of Contents

Introduction

Parliamentary corpora in the CLARIN infrastructure

Other parliamentary corpora

CLARIN-funded project: ParlaMint

Additional materials

Publications on the parliamentary corpora



<https://www.clarin.eu/content/parlamint-towards-comparable-parliamentary-corpora>

NoSketchyEngine

▼ 🔍 ParlaMint-LV 2.0 (Latvian parliament)

Home
Search
Word list
Corpus info
My jobs
User guide 📖

ParlaMint-LV 2.0 (Latvian parliament) ⓘ

Latvian parliamentary corpus ParlaMint-LV, 2014-2021 v2.0

Counts	General info	Lexicon sizes
Tokens 8,295,174	Corpus description Document	word 135,984
Words 6,342,984	Language Latvian	lc ⓘ 123,123
Sentences 674,272	Encoding UTF-8	norm ⓘ 135,307
Paragraphs 122,136	Compiled 05/12/2021 16:23:07	lemma 42,993
Documents 419	Tagset Description	lemma_lc ⓘ 42,114
		pos ⓘ 18
		feats ⓘ 946
		id ⓘ 248
		den ⓘ 18

Query **NOUN, ADJ** 68,432 (8,272.52 per million) ⓘ

Page 1 of 3,432 Go Next Last

Structures and attributes
name 231,561 ⓘ
text 419 ⓘ
note 122,136 ⓘ
p 122,136 ⓘ
s 674,272 ⓘ
speech 122,136 ⓘ

Menu position

Query NOUN, ADJ 68,632 (8,272.32 per million) EN		
Page 1 of 3,432 Go Next Last		
2016-06-16, VRJonišRaimonds	ierīnāt jaunus darbus. Šodien [ORG: Saime] / lems par pievienošanos <small>/pɛi'vɛnoʃanos/</small> [ORG: Ekonomiskās sadarbības] / un attīstības organizācijai.	
2016-06-16, VRJonišRaimonds	organizācijai, iespēja pievienoties 34 pasauļes <small>/pasau'ļes/</small> attīstībā <small>/attɪ'stɪbā/</small>	
2016-06-16, VRJonišRaimonds	gada, uzrunājot šīs pirmo reizi pēc Valsts prezidenta <small>/prezɪ'dɛnta/</small> svēgīgā <small>/svɛ'gɪgā/</small>	
2016-06-16, VRJonišRaimonds	ir ne vien Latvijā, bet visa Baltijas reģiona <small>/rɛ'gɪona/</small> enerģētiskās <small>/ɛnɛ'r'gɛtɪskās/</small>	
2016-06-16, VRJonišRaimonds	darbu Turpināš Nacionālajiem bruņotajiem spēkiem <small>/spɛ'kɛim/</small> negaidīta <small>/nɛ'gaidɪ'ta/</small>	
2016-06-16, VRJonišRaimonds	organizācijas dalībvalstīs ir jāpārda savas sarīstības <small>/sarɪ'stɪbɪ's/</small> kolēktīvā <small>/kolɛktɪ'vā/</small>	
2016-06-16, VRJonišRaimonds	ģeopolitiski izsalcinājumi. Turpinā Krimas <small>/krɪ'mas/</small> nelikumīgā <small>/nɛ'lɪkumɪ'gā/</small>	
2016-06-16, VRJonišRaimonds	konfliktā, kā arī ekonomiskā un politiskā nestabilitātē <small>/nɛ'stɪlɪ'tatē/</small> vairākos <small>/vair'ko's/</small>	
2016-06-16, VRJonišRaimonds	Āfrikas valstīs. Šajos apstākļos mūsu drošībai <small>/drɔ'ʃɪ'bɪ'b/</small> būtiska <small>/bʊtɪ'skā/</small>	
2016-06-16, VRJonišRaimonds	esmu norādījis, ka man nav pieņemama [ORG: Saimes] / prakse <small>/prak'se/</small> trešajā <small>/trɛ'sajā/</small>	
2016-06-16, VRJonišRaimonds	apspriest. Tīrs reizes šajā gadā esmu lūdzis likuma <small>/lɪkum'a/</small> atbilstošā <small>/atbɪ'stoʃā/</small>	
2016-06-16, VRJonišRaimonds	likumprojektu virzībā. Par sarežģītiem ierosinājumiem <small>/ɪɛrɔ'sɪnājumɛim/</small> nepieciešamas <small>/nɛ'pɪcɪɛ'šamas/</small>	
2016-06-16, VRJonišRaimonds	vairāk jāromdina [ORG: Latvijas Zinātņu akadēmijas] / zinātnieku <small>/zɪ'nātɲɛikʊ/</small> potenciāls <small>/pɔtɛncɪ'ʃals/</small>	
2016-06-16, VRJonišRaimonds	apkarotāni, lai gūtu vairāk ienākumu valsts budžetā <small>/budʒɛ'tā/</small> sabiedriskā <small>/sabɪɛdrɪ'skā/</small>	
2016-06-16, VRJonišRaimonds	pilnveidošanā. Nedrīkstam aizmirst par gauskiem <small>/gaus'kɛim/</small> demogrāfiskās <small>/dɛ'mɔ'grāfɪ'skās/</small>	
2016-06-16, MārnieceIvāna	" Par [ORG: Saimes] / deputātes [PER: Laimdota Straujuma] / apstiprināšanu <small>/apstɪ'prɪnɪ'shānu/</small> [ORG: Centrālās]	
2016-06-16, MārnieceIvāna	darba kārtību un tajā iekļautu likumprojektu " Grozījumi <small>/grozɪ'jumi/</small> Alkoholiskā <small>/alkɔ'hɔlɪ'skā/</small>	
2016-06-16, MārnieceIvāna	projektu [ORG:] " Par Saimes / deputāte [PER: Ineša Buķa] / apstiprināšanu <small>/apstɪ'prɪnɪ'shānu/</small> [ORG: Centrālās]	
2016-06-16, MārnieceIvāna	" Par Saimes piekrišanu likumā " Par valsts budžetu <small>/budʒɛ'tu/</small> 2016. <small>/2016/</small>	
2016-06-16, MārnieceIvāna	" Par Saimes piekrišanu likumā " Par valsts budžetu <small>/budʒɛ'tu/</small> 2016. <small>/2016/</small>	

Use Case 1

Gender in Parliamentary Discourse

Background

While more women than ever are being elected to parliaments around the world, equality is still a long way off, and current progress is far too slow. Most parliaments are still heavily male-dominated, and some have no women members of parliament at all. Even where women are present in greater numbers, glass ceilings often remain firmly in place. (Source: [Women in Parliament](#) by the Inter-Parliamentary Union)

CLARIN Resources and services used

- CLARIN Resource Families – parliamentary corpora
- the CLARIN.Si repository
- the noSketch Engine concordancer

Citizen scientist

Do female speakers in the Slovenian and Croatian parliaments speak more or less than their male counterparts?



Student

Is the language of female parliamentary speakers similar in Slovenia and Croatia?



Step-by-step guide

- 1 Search the Parliamentary CLARIN Resource Family for relevant Slovenian and Croatian corpora. In this walkthrough, we'll use the Croatian and Slovenian ParlaMETER corpora, since they are roughly comparable in terms of time span, linguistic annotation and speaker metadata, but you can also use any of the other parliamentary corpora.

Croatian parliamentary corpus ParlaMETER 1.0	Croatian	The corpus contains minutes of the National Assembly of the Republic of Croatia and currently covers the 1995-2016 period from 18 November 2010 to 11 November 2016. The corpus contains speaker metadata (gender, age, education, party affiliation). The corpus is available for download from the CLARIN.Si repository and through the concordancers KorText and noSketch engines, as well as through a dedicated webpage .	CLARIN Download
Slovenian parliamentary corpus ParlaMETER 1.0	Slovenian	The corpus contains Slovenian parliamentary debates from 1990 to 2016. It differs from the ParlaMETER 1.0 corpus listed below in that it contains only basic meta-data about the speakers, a typology of sessions and structural and editorial annotations. The corpus is available for download from the CLARIN.Si repository and through the concordancers KorText and noSketch engines.	CLARIN Download



Yankelevich, Tanya, Fiser, Darja, Lenardic, Jakob, Gorgaini, Elisa, & Braukmann, Ricarda. (2020, June). **LIBER 2020 - Workshop: SSHOC Train-the-Trainer Bootcamp for Librarians**. Zenodo. <http://doi.org/10.5281/zenodo.3970799>

Slovenian, Croatian



Italian, English, Polish

Helsinki DH Hackathon 2021

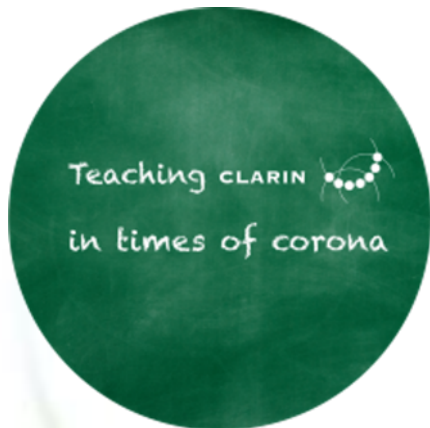
<https://dhhackathon.wordpress.com/2021/05/28/parliamentary-debates-in-the-covid-times/>

Frontini, Francesca, 2021.

Presentation on CLAIRN at **Pisa summer school**.

"Teaching CLARIN in Times of Corona"

Deadline ~~01/06/2021~~ **15/06/2021**



All relevant submissions will be showcased on the CLARIN website, for the purpose of sharing best practices and examples suitable for reuse. In addition, three submissions will be selected for the Teaching CLARIN award, and will receive the sum of **350 euros each**. The award ceremony will be hosted during [CLARIN Annual Conference 2021](#).

Getting involved in CLARIN

- Join our NewsFlash
 - <https://www.clarin.eu/content/newsflash>
- Check out our events
 - <https://www.clarin.eu/events>
- Open calls
 - <https://www.clarin.eu/content/funding-opportunities>
- Follow us on Twitter @CLARINERIC
- And stay tuned for the next cafés
 - <https://www.clarin.eu/content/clarin-cafe>
 - **ParlaMint Café - 28th June 14:00-16:00 CEST**

www.clarin.eu

or

francesca.frontini@ilc.cnr.it