

CLARIN ERIC: Common Language Resources and Technology Infrastructure

Mietta Lennes

FIN-CLARIN & The Language Bank of Finland

University of Helsinki

14 April 2022



The research infrastructure for language as social and cultural data

CLARIN is a digital infrastructure offering data, tools and services to support research based on language resources.

Easy and sustainable access

Single sign-on

Advanced tools & technologies

Interoperability

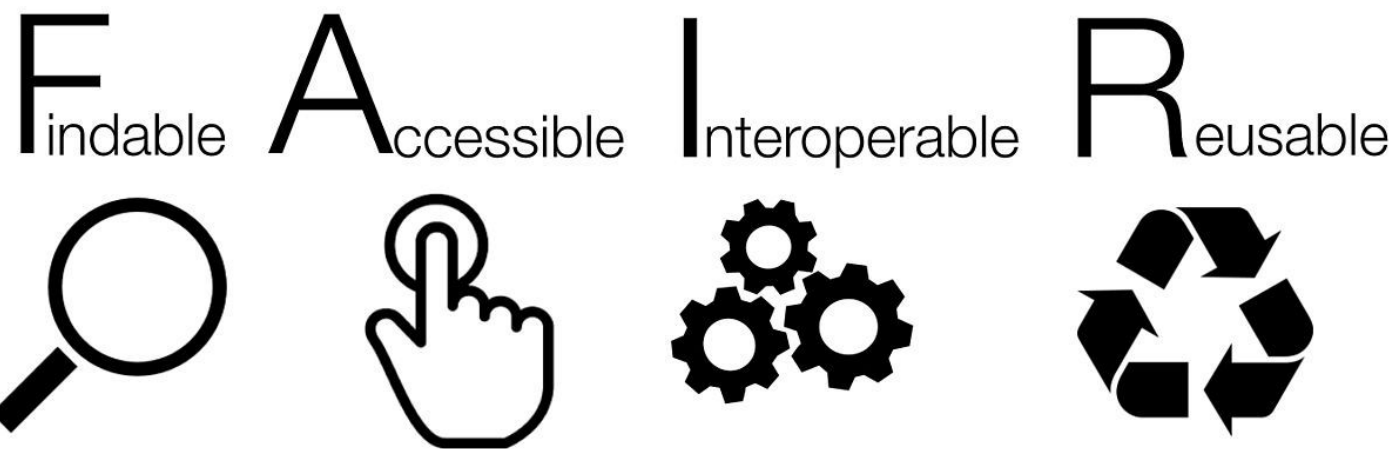
Training

Knowledge exchange

Support for linguistic diversity



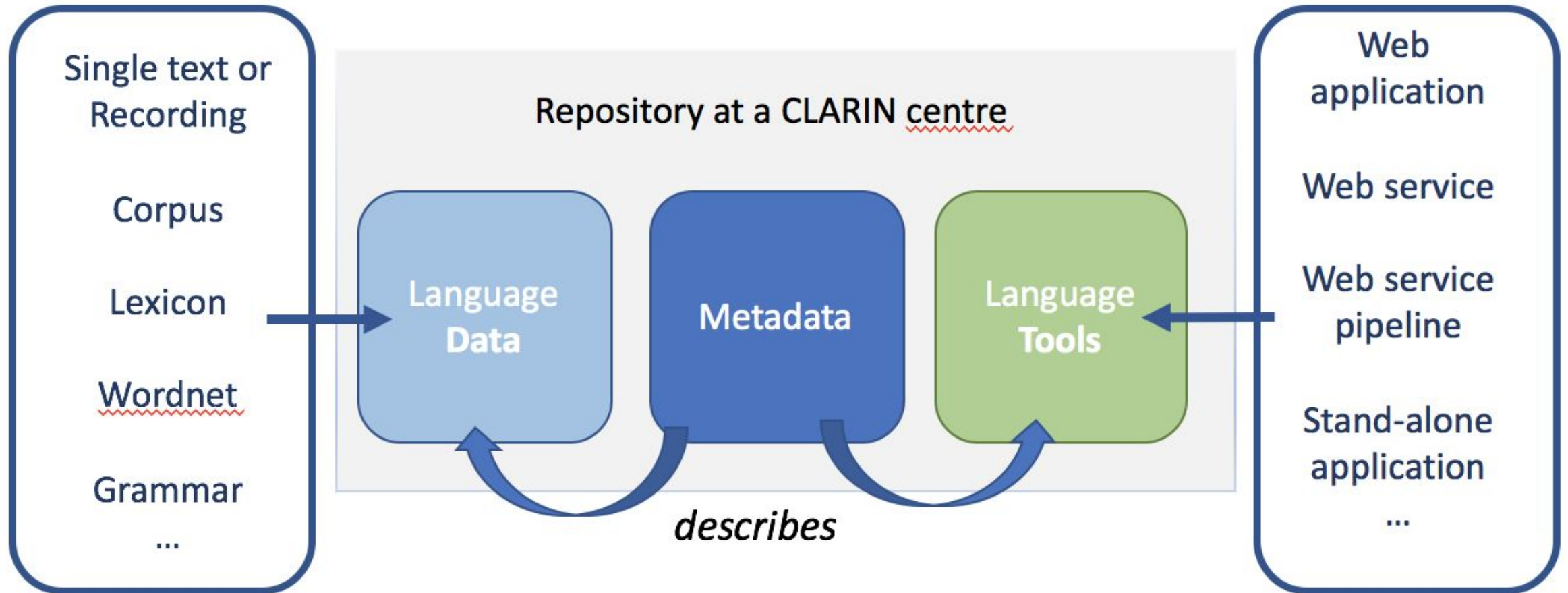
CLARIN makes data



clarin.eu/fair

The CLARIN data architecture:

Network of distributed centre repositories





CLARIN centres

Member countries (22):

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech Republic
- Denmark
- Estonia
- Finland
- Germany
- Greece
- Hungary
- Iceland
- Italy
- Latvia
- Lithuania
- The Netherlands**
- Norway
- Poland
- Portugal
- Slovenia
- Sweden

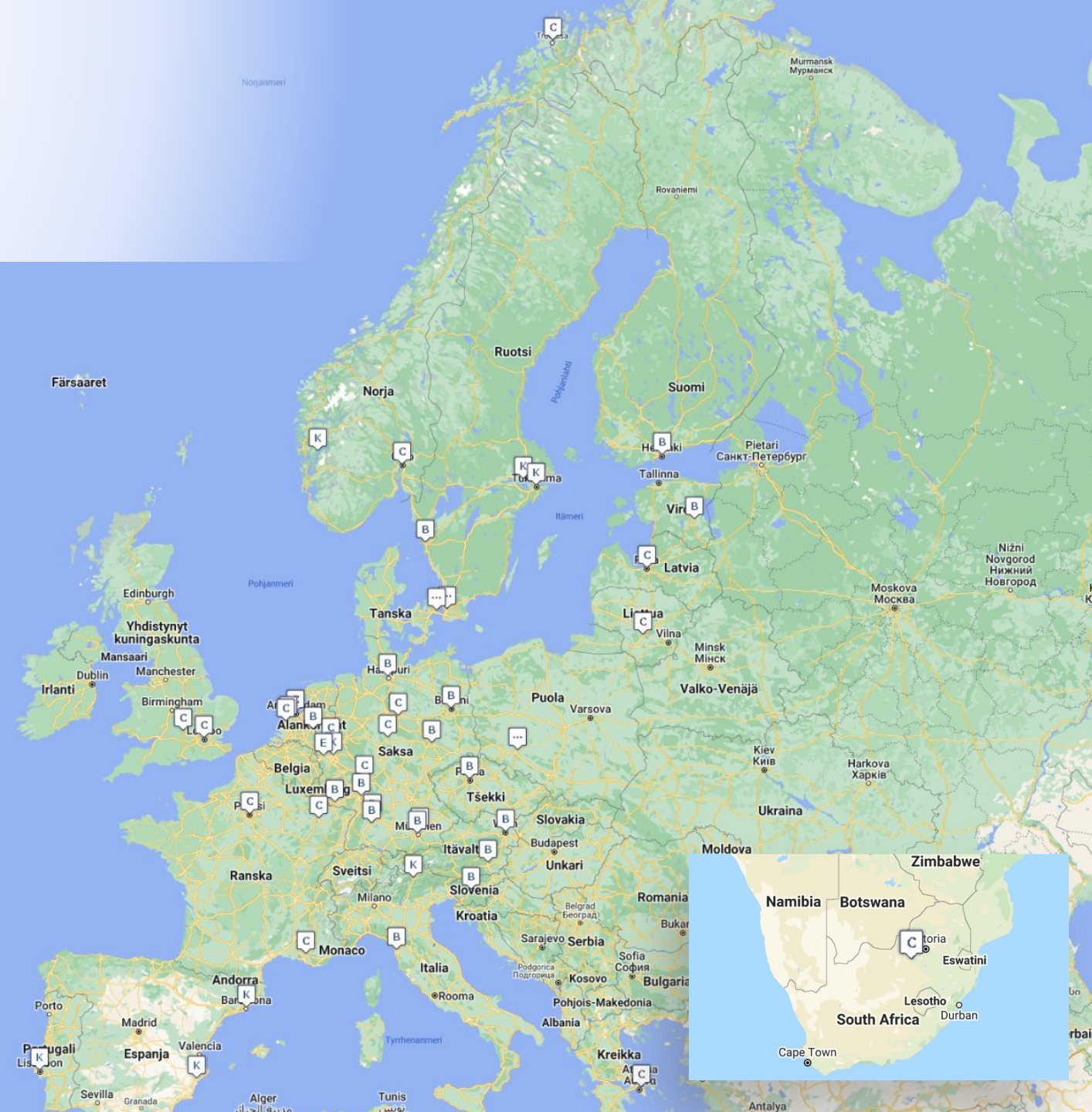
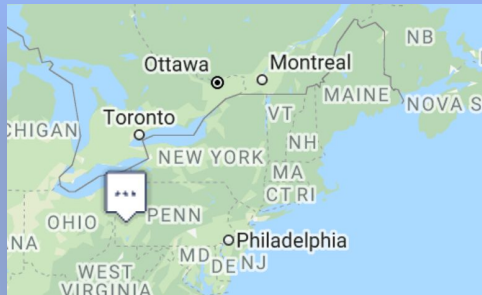
Observers (2):

- South Africa
- United Kingdom

Third party (1):

- CMU (USA)

Updated: 14.4.2022





CLARIN centres

Member countries (22):

- Austria
- Belgium
- Bulgaria
- Croatia
- Cyprus
- Czech Republic
- Denmark
- Estonia
- Finland
- Germany
- Greece
- Hungary
- Iceland
- Italy
- Latvia
- Lithuania
- The Netherlands**
- Norway
- Poland
- Portugal
- Slovenia
- Sweden

Observers (2):

- South Africa
- United Kingdom

Third party (1):

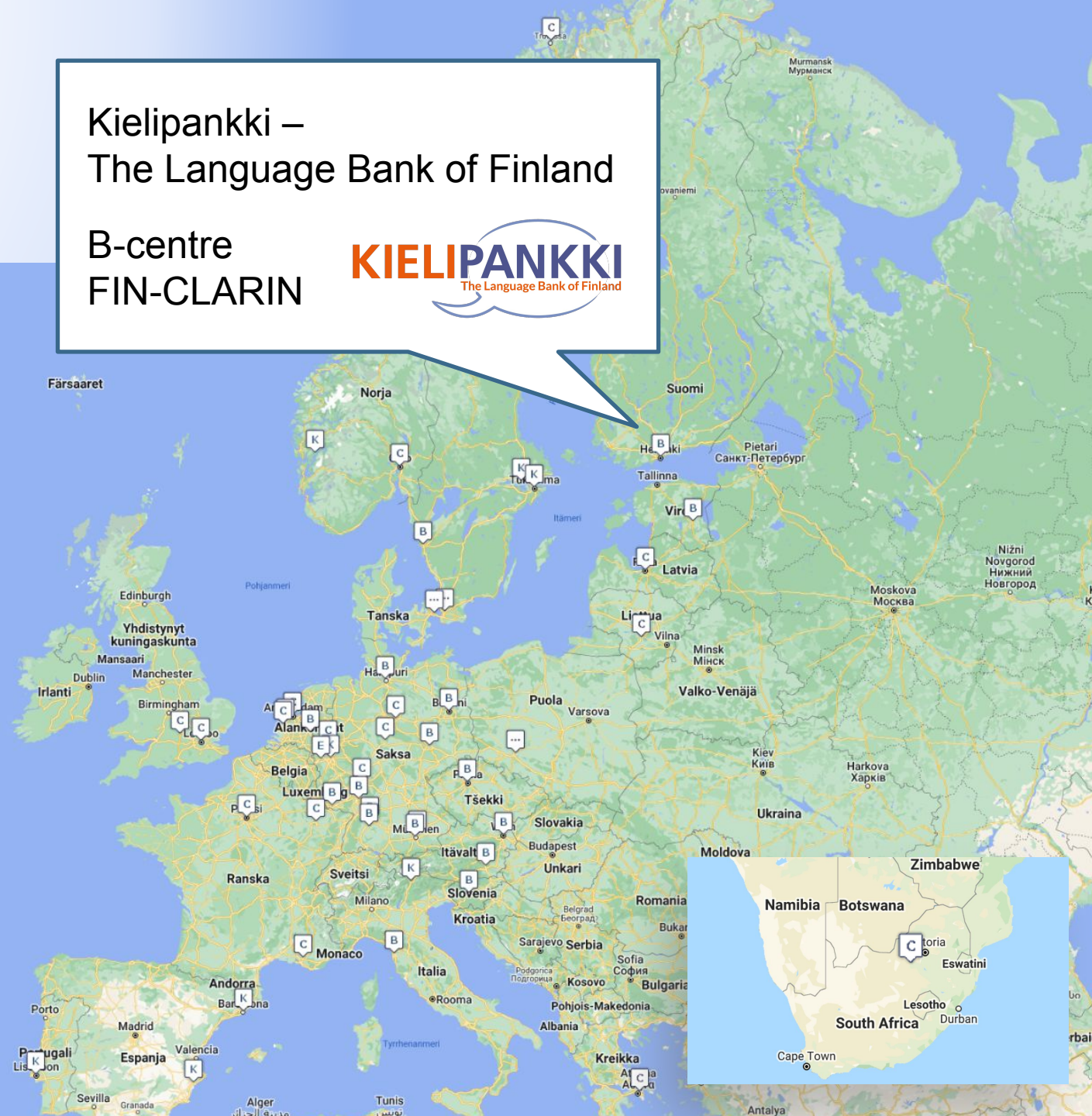
- CMU (USA)

Updated: 14.4.2022



Kielipankki –
The Language Bank of Finland

B-centre
FIN-CLARIN

FIN-CLARIN

KIELIPANKKI
The Language Bank of Finland

University of Oulu

University of Vaasa

University of Jyväskylä

University of Eastern Finland

Tampere University

University of Turku

Institute for the Languages in Finland
CSC -IT Center for Science (Technical center)
University of Helsinki (National coordination)

Sweden

Finland

Russia

Estonia

KIELIPANKKI

The Language Bank of Finland

LANGUAGE BANK ACCESS CORPORA TOOLS ORGANIZATION SUPPORT

SUOMEKSI PÅ SVENSKA

Access



Apply for rights to use our language resources.

Corpora



Browse our corpora.

Tools



Try our tools.

Organization



Who are the Language Bank?

Support



Help and instructions.

Search the Language Bank Portal:

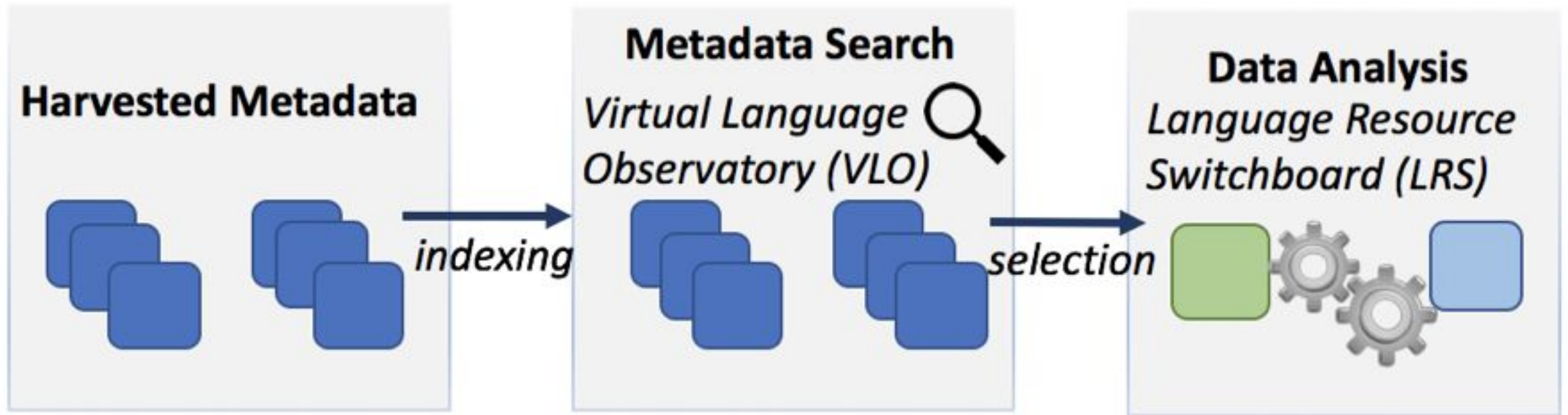
Haku ...



Researcher of the Month: Ari Huhta

The CLARIN data architecture:
Metadata services



clarin.eu/fair

vlo.clarin.eu

switchboard.clarin.eu



voice disguise



Showing 1 result within selection for voice disguise x Finnish x

Results per page: 10

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

Language

Type to filter or search for more

Finnish x

Bodo (brx) (1)

English (1)

Resource type

Format

Temporal Coverage

Availability

Search options

Corpus of Age-related Voice Disguise

This corpus includes normal and age-related disguised speech uttered by 60 native Finnish speakers (31 females and 29 males). The speakers were asked to read the same text fragments several times, in their modal voice and in two disguised voices, first pretending to be an elderly speaker and then pretending to be a chi...

Finnish English



Virtual Language Observatory
vlo.clarin.eu

List of corpora at the Language Bank of Finland

www.kielipankki.fi/corpora




































Cite

Corpora-specific reference instructions and citations in other papers are available by clicking the quote ” link.

If you are looking for a corpus not listed here, please have a look in [META-SHARE](#) or [CLARIN Virtual Language Observatory \(VLO\)](#).





Please find an overview of all our resources sorted by resource families in the [resource families of FIN-CLARIN](#).

Etsi:

| Abbreviation | Name and metadata | License | Apply | Location | Service level | Help | Cite |
|--------------------|---|---|---|----------|---------------|---|---|
| acquis-ftb3 | The Finnish Sub-corpus of the JRC-Acquis Multilingual Parallel Corpus |  | | Korp | B |  |  |
| acquis-ftb3-dl | Finnish Sub-corpus of the JRC-Acquis Multilingual Parallel Corpus, Downloadable Version |  | | Download | B |  |  |
| agricola-v1-1-korp | The Morpho-Syntactic Database of Mikael Agricola's Works version 1.1, Korp |  | | Korp | B |  |  |
| ai2d-rst-v1-1 | AI2D-RST: A multimodal corpus of 1000 primary school science diagrams version 1.1 |  | | Download | B |  |  |
| aku-egg-dl | Speech and EGG (Electroglottography) Simultaneous Recordings, downloadable version |  | | Download | B |  |  |
| amph | amph-Corpus |  |  | Download | B |  |  |
| ArkiSyn-korp | ArkiSyn Database of Finnish Conversational Discourse, Helsinki Korp Version |  | | Korp | B |  |  |
| AVOID | Corpus of Age-related Voice Disguise (AVOID) |  |  | Download | B |  |  |
| BeserCorp | The Corpus of Beserman Udmurt |  | | Korp | B |  |  |
| ccmh-src | Corpus Cyrillo-Methodianum Helsingiense: Corpus of Old Church Slavonic Texts, source |  | | Download | B |  |  |
| ceal-dl | The Downloadable Version of Classics of English and American Literature in Finnish |  | | Download | A |  |  |

Citation instructions



| | | | | | | |
|--------------|---|---|---|----------|---|---|
| ArkiSyn-korp | ArkiSyn Database of Finnish Conversational Discourse, Helsinki Korp Version |  | | Korp | B | ” |
| AVOID | Corpus of Age-related Voice Disguise (AVOID) |  |  | Download | B | ” |
| BeserCorp | The Corpus of Beserman Udmurt |  | | Korp | B | ” |

[suomeksi] [in English]

Reference instructions: AVOID

Please cite the language resource as follows:








Kinnunen, T., Hautamäki, R. G., Sahidullah, M., Hautamäki, V., Werner, S., & Bentz, M.. *Corpus of Age-related Voice Disguise (AVOID)* [speech corpus]. Kielipankki. Retrieved from <http://urn.fi/urn:nbn:fi:lb-2018060621>

Show: [Bibtex] [Zotero]

[Search for references to the language resource in Google Scholar](#)

List of tools available via the Language Bank of Finland

Etsi:

| Start | Name | Description | Instructions | Install | Administrator | Service level |
|---|--|--|------------------------------|---|---|---------------|
|  | Korp | A web-based concordance tool that can be used for corpus queries based on morphosyntactic analysis and various other features. | Instructions | |  | A |
| Download | Download service | Download certain corpora. | | |  | A |
|  | META-SHARE | Metadata repository of all the language resources at the Language Bank of Finland. | | |  | A |
| Mylly | Mylly | Versatile data analysis platform with interactive visualizations and workflows. | Instructions | |  | C |
| Sanat | Sanat | A platform for publishing lexica and word lists. | | |  | B |
| FinTag | Finnish Tagtools | A part-of-speech and morphology tagger and a named entity recogniser for Finnish. | | Install |  | A |
| Demo | Demo tools at the Language Bank of Finland | Demos of tools that are in development at the Language Bank of Finland: FinTag and FINER, FinSentiment, FinnWordNet, HFST POS taggers, HFST morphological analyzers, Lemmamatch, etc. (In Finnish) | | |  | C |
|  | WebAnno | Text annotation tool. | User Guide | Standalone installation |  | A |

Tool: korp.csc.fi

Finnish | Swedish | Other languages | Parallel

Log in Suomi | Svenska | English 0 Cite Korp MENU



179 of 989 corpora selected — 5.25G of 14.10G tokens



Simple Extended Advanced Compare 1

Filter: Add publication

in order and also as initial part compound_middle final part and case-insensitive

KWIC: hits per page: 25 sort within corpora: not sorted Statistics: compile based on: word Show statistics Show word picture

Tool: korp.csc.fi

Finnish | Swedish | Other languages | Parallel

Log in Suomi | Svenska | English 0 Cite Korp MENU



Simple Extended Advanced Compare 1

Filter: Add publication

Search

in order and also as initial part compound_mid

KWIC: hits per page: 25 sort within corpora: not so

179 of 989 corpora selected — 5.25G of 14.10G tokens

1600 1700 1800 1900 2000

Select all
 Select none

- Tietokonevälikkeistä viestintää (33)
- Akateemisia tekstejä (24)
- Historiallisia aineistoja (115)
- Oppijansuomen aineistoja (4)
- Kirjallisuusaineistoja (26)
- Käsien annotoituja aineistoja (1)
- Lehti- ja uutisaineistoja (622)
- Parlamenttiaineistoja (2)
- Puheaineistoja (138)
- Juridisia aineistoja (7)
- Selkokieliaineistoja (4)
- Muita aineistoja (13)

399,701,316 sentences in the selected corpora

cs Show word picture



Eduskunnan täysistunnot selected — 22.46M of 12.93G tokens



Simple Extended Advanced Compare

Plenary Sessions of the Parliament of Finland: *maahanmuuttaja* 'immigrant' (all forms)

base form is
 maahanmuuttaja Aa
 or
 +

Search within sentence

KWIC: hits per page: 25 sort within corpora: not sorted Statistics: compile based on: word Show word picture Show map

KWIC Statistics Word picture Map

Results: 1,618

« < 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 ... > » Show context

EDUSKUNNAN TÄYSISTUNNOT (does not support extended context)

in jonain päivänä näkeväni myös kuntoutusammattilaisia ja maahanmuuttajapalveluissa lisää kielitaitoisia **maahanmuuttajia** kuten esimerkiksi Lappeenrannassa on jo tällä hetkellä.
 Sivustusbudjetissa uudet suomalaiset eli **maahanmuuttajat** on noteerattu yhtenä painopistealueena.
 Lisävaroja tulee **maahanmuuttajien** perusopetukseen valmistavaan koulutukseen sekä maahanmuuttajataustaisten opettajien koulutta
 Väestömme ikääntyessä **maahanmuuttajia** tarvitaan Suomeen yhä enemmän.
 Myös **maahanmuuttajien** tarpeet ammatillisessa koulutuksessa on huomioitava nykyistä paremmin.
Maahanmuuttajien ammatilliseen koulutukseen ohjataan lisää rahaa.
 Merkittävä toinen Pop panostus on **maahanmuuttajien** valmistavan koulutuksen keston pidentäminen puolesta vuodesta vuoteen.
 Nuorten syrjäytymistä vastaan on tehty paljon täällä toimenpiteitä esimerkiksi **maahanmuuttajien** osalta erityisopetuksen osalta ammatillisia aloituspaikkoja on systemaattisesti lisätty ja lisätään ed
 tuista varoista 15 miljoonaa kohdennettiin uudelleen siten että ne menevät ammatilliseen koulutukseen **maahanmuuttajien** koulutukseen ja nuorten syrjäytymisen ehkäisemiseen ja työllisyyden edistämiseen.
 Erillinen määräraha **maahanmuuttajien** kielikoulutukseen olisi välttämätön jotta siihen ei jouduttaisi käyttämään muutenkin riittämättömiä
 Me tarvitsemme **maahanmuuttajia** .
 Sisäasiainministeriön vuoden 2009INE talousarviossa on tavoitteena että **maahanmuuttajien** koko koulutusjärjestelmä vastaa sekä laadullisesti että määrällisesti yksilöllisiä tarpeita.
 Tavoitteena on myöskin että **maahanmuuttajien** työttömyys puolittuu vuoteen 2012INE mennessä.
 e merkittäville maahanmuuttoalueille laaditaan yhdessä valtion ja seutujen kuntien kanssa pilottiohjelma **maahanmuuttajien** kotouttamisen ja työllistymisen edistämiseksi.
 Samoin on koko yhteiskunnan etu myös että **maahanmuuttajat** työllistyvät mahdollisimman hyvin ja saavat tarvitsemansa koulutuksen.
 Sinne tuli sekä paikallisia että näitä **maahanmuuttajia** pienen ajan päästä tuli valtavasti kumpienkin vanhempia myös mukaan.
 On hieno asia se että **maahanmuuttajia** ja turvapaikanhakijoita on otettu Suomeen aiempaa enemmän.
 ää että yhdistys ja järjestötoiminta on erittäin keskeisessä asemassa ja tiedämme myöskin miten tärkeitä **maahanmuuttajat** ovat meidän jalkapallollemme jos nyt vain viitataan Saksa otteluun.
 eman kuin sosiaalisia yrityksiä jotka esimerkiksi olisivat kuin vuokratyöyrityksiä ja tarjoaisivat ja etsisivät **maahanmuuttajille** nimenomaan turvapaikkamenettelyä tulleille ihmisille ehkä sitten tällä tavalla työpaikkoja.
 Miten tullaan varmistamaan se että suomen kielen opetusta on tarjolla ja että sitä saa silloin kun **maahanmuuttaja** on valmis sitoutumaan kielen opiskeluun.
 niin tämä on todella paljon ja samaan aikaan meillä sanotaan että maahan tarvitaan niin paljon **maahanmuuttajia** ja niin paljon jotta tämä kansa voi selvitä tulevaisuudessa.

Corpus

Eduskunnan täysistunnot

Persistent Identifier (metadata):

urn:nbn:fi:lb-2017020202

Cite corpus

Licence: CC BY-NC-ND (CLARIN PUB)

Link to corpus in Korp

Text attributes

file name: 2008-09-18.vrt

date: 2008-09-19

text_time: 06:12:09

speaker_name: Outi Alanko-Kahiluoto

speaker_parl_group: parlgrou_pvihr

speaker_role: [empty]

speech_type: [empty]

speaker: Outi Alanko-Kahiluoto /vihr

utterance begin time: 4797368

utterance duration: 20960

utterance end time: 4818328

utterance number: 9577

Word attributes

Plenary Sessions of the Parliament of Finland: *maahanmuuttaja* 'immigrant'

Sisäasiainministeriön vuoden 2009INE talousarviossa on tavoitteena että **maahanmuuttajien** koko koulutusjärjestelmä vastaa sekä laadullisesti että määrällisesti yksilöllisiä tarpeita. Tavoitteena on myöskin että **maahanmuuttajien** työttömyys puolittuu vuoteen 2012INE mennessä. nmuuttoalueille laaditaan yhdessä valtion ja seutujen kuntien kanssa pilottiohjelma **maahanmuuttajien** kotouttamisen ja työllistymisen edistämiseksi. Samoin on koko yhteiskunnan etu myös että **maahanmuuttajat** työllistyvät mahdollisimman hyvin ja saavat tarvitsemansa koulutuksen. Sinne tuli sekä paikallisia että näitä **maahanmuuttajia** pienen ajan päästä tuli valtavasti kumpienkin vanhempia myös mukaan. On hieno asia se että **maahanmuuttajia** ja turvapaikanhakijoita on otettu Suomeen aiempaa enemmän. **maahanmuuttajat** ovat meidän jalkapallolemme jos nyt vain viitataan Saksa otteluun. **maahanmuuttajille** nimenomaan turvapaikkamenettelystä tulleille ihmisille ehkä sitten tällä tavalla työpöytä **maahanmuuttajia** on valmis sitoutumaan kielen opiskeluun. **maahanmuuttajia** ja niin paljon jotta tämä kansa voi selvitä tulevaisuudessa. **maahanmuuttajia** . **maahanmuuttajia** valmistavaa opetusta jota esitetään pidennettäväksi puolesta vuodesta vuoden pituisen **maahanmuuttajien** valmistavan koulutuksen lisäämistä ja suomi tai ruotsi toisena kielenä opetusta. Lausunnossa valiokunta teki yksitoista toimenpide ehdotusta joista yksi koski **maahanmuuttajien** valmistavan opetuksen ja kielikoulutuksen laajentamiseksi. tässä salissa lausuma jossa eduskunta edellytti hallitusta ryhtymään toimenpiteisiin **maahanmuuttajien** valmistavan opetuksen ja kielikoulutuksen laajentamiseksi.

« < 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 ... > »

Download KWIC

as Annotations

Excel (XLS)

JSON

publication time: 10.00.00
session duration (ms): 10178760
[original video](#)
[original transcript](#)
name of speaker: Maria Guzenina-Richardson
parliamentary group: parlgrou_sd
speaker role: [empty]
type of speech: [empty]
speaker: Maria Guzenina-Richardson /sd
utterance begin time (ms): 600928
utterance duration (ms): 11357
utterance end time (ms): 612285
utterance number: 17661
session duration: 2:49:38.760
utterance begin time: 0:10:00.928
utterance end time: 0:10:12.285
utterance duration: 0:00:11.357

Word attributes

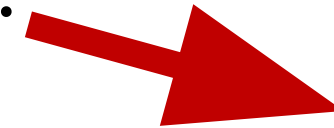
base form: maahanmuuttaja
baseform (compound boundaries):
maahanmuuttaja
part of speech: noun
msd: NUM_PI|CASE_Par
dependency relation: direct object
word position in a name: outside (O)

[Show Dependency Tree](#)

[Show video](#)

[Show video in LAT service](#)

Click on link to show video!



Plenary Sessions of the Parliament of Finland, Korp version 1.5: Link to video



Play/Pause

-30 s

-5 s

+5 s

+30 s

Play
selection

Stop video at selection end:

Media source:

https://down0-80133-down0.dna.qbrick.com/80133-down0/Syysistuntokausi_2008/97._torstaina_23._lokakuuta_2008.mp4

Corpus



Eduskunnan täysistunnot

Persistent Identifier (metadata):

urn:nbn:fi:lb-2019101621

Licence (audio and video): CLARIN PUB +BY +PRIV +ND +OTHER 1.0

Licence (text): CLARIN PUB +BY +PRIV 1.0

[Link to search in Korp](#)

Text attributes

publication date: 2008-10-23

publication time: 18:06:38

Please note that the video feature on this page is not supported by all browsers. The page is most likely to work in a Firefox browser.

The original transcript has been aligned with the media file and annotated by using automatic methods. The alignment or the annotations may thus contain errors.

Utterance: Sitten äskettäin Espoossa pidetyssä vaalitulaisuudessa ministeri Väyrynen nostatti pelkotiloja espoolaisissa todeten että metro tuo kaupunkiin maahanmuuttajia .

name of speaker: Maria Guzenina-Richardson

parliamentary group: parlgroup_sd

speaker role: speakerrole_

type of speech: speechtype_

utterance begin time: 0:10:00.928

utterance end time: 0:10:12.285

utterance duration: 0:00:11.357

utterance begin time (ms): 600928

utterance end time (ms): 612285

utterance duration (ms): 11357

Switchboard – supporting interoperable tools

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 data types and communities of use

- 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

Resource families

Introduction

The aim of the CLARIN Resource Families initiative is to provide a user-friendly overview per data type of the available language resources in the CLARIN infrastructure aimed at the needs of researchers from digital humanities, social sciences and human language technologies. The overviews are meant to facilitate comparative research and the listings are sorted by language.

The listings for each family include the most important metadata and brief descriptions, such as resource size, text sources, time periods, annotations and licences as well as links to download pages and concordancers, whenever available. In addition to the resources found in the CLARIN infrastructure an overview is provided of other existing valuable language resources which have not yet been integrated in the infrastructure.

The listings also provide hyperlinks to other relevant materials such as the thematic CLARIN workshops and tutorials and their accompanying videolectures, as well as a list of key publications on the resources surveyed.

Currently overviews are available of 12 corpora families, 5 families of lexical resources, and 4 tool families. See below. For the possibility to apply for funding for small projects that can help extending the scope of the initiative, see <https://www.clarin.eu/content/clarin-resource-families-project-funding>.

Corpora

- [Computer-mediated communication corpora](#)
- [Corpora of academic texts](#)
- [Historical corpora](#)
- [L2 learner corpora](#)
- [Literary corpora](#)
- [Manually annotated corpora](#)
- [Multimodal corpora](#)
- [Newspaper corpora](#)
- [Parallel corpora](#)
- [Parliamentary corpora](#)
- [Reference corpora](#)
- [Spoken corpora](#)

Lexical Resources

- [Lexica](#)
- [Dictionaries](#)
- [Conceptual Resources](#)
- [Glossaries](#)
- [Wordlists](#)

Tools

- [Normalization](#)
- [Named entity recognition](#)
- [Part-of-speech tagging and lemmatization](#)
- [Tools for sentiment analysis](#)

Resource families

[CMC corpora](#)

[Historical corpora](#)

[L2 corpora](#)

[Manually annotated corpora](#)

[Multimodal corpora](#)

[Newspaper corpora](#)

[Parallel corpora](#)

[Parliamentary corpora](#)

[Reference corpora](#)

[Spoken corpora](#)

[Lexica](#)

[Dictionaries](#)

[Conceptual resources](#)

[Glossaries](#)

[Wordlists](#)

[Tools for normalization](#)

[Tools for named entity recognition](#)

[Part-of-speech taggers and lemmatizers](#)

[Tools for sentiment analysis](#)

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

CLARIN licence categories



Publicly available



Available for academic, logged in users



Personal permission is required for access

More detailed license conditions



+BY author must be cited
+NC non-commercial use only
+ID login is required



+PLAN research plan is required
+PRIV contains personal data



+NORED redistribution is not allowed
+DEP modified versions can be redistributed via
CLARIN

and other resource-specific conditions, if required (e.g., data protection terms and conditions)

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



DIGITAL HUMANITIES COURSE REGISTRY



STORIES

INFO

LOGIN

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 Hu ... | 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 ↻ | |



Imprint

©2014-2021 CC-BY 4.0

FIN-CLARIN Online courses



Corpus Linguistics and Statistical Methods (5 cr)

Word attributes

- baseform: kieli
- baseform (compound boundaries): kieli
- part-of-speech: noun
- msd: NUM_Sg|CASE_Nom
- dependency relation: nominal subject
- word position in a name: outside (O)
- original thread
- original message
- Show Dependency Tree
- Näytä dependenssipuu



Introduction to Speech Analysis (5 cr)

Waveform and spectrogram analysis of a speech segment. The spectrogram shows frequency components over time, with a blue line indicating a specific frequency track. The waveform shows the amplitude of the speech signal.



Data Clinic (5 cr)

Dependency tree and bar chart analysis. The dependency tree shows the syntactic structure of a sentence, and the bar chart displays statistical data related to the analysis.



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



CLARIN as ecosystem for knowledge exchange



Project: [Academic Organisations](#) » CLARIN - Common Language Resources and Technology Infrastructure

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



- Watching Now
- This page - CLARIN - Common Language
1
 - Category: Computers - VideoLectures.NET
1
 - International Center for Advanced
1
 - Course on Information Theory, Pattern
1
 - PhD Thesis Defense: Dynamics of large
1

CLARIN - COMMON LANGUAGE RESOURCES AND TECHNOLOGY INFRASTRUCTURE

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

RECENT EVENTS [MORE...](#)



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 ...



Baltic Summer School of Digital Humanities Essentials of Coding and Encoding

CLARIN as ecosystem for knowledge exchange

Funding opportunities



Getting involved in CLARIN

- Subscribe to the 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**

Thank you!

CLARIN

www.clarin.eu

↓
Mietta Lennes
fin-clarin@helsinki.fi

