

WORKSHOP

SSHOC Train-the-Trainer Bootcamp for Librarians

June 23, 2020 | 10:00 CEST



www.liberconference.eu
@LIBERconference | [#LIBER2020](https://twitter.com/LIBER2020)



SSH Open Cloud

Train-the-Trainer Bootcamp for Librarians

LIBER Online 23.06.2020

Tanya Yankelevich, Darja Fiser, Jakob Lenardic, Elisa Gorgaini & Ricarda Braukmann



Housekeeping...

- ❑ Keep your **microphone muted** and **camera off** during presentations
- ❑ **Questions?** Put them in the **chat** during presentations. Raise your hand or use your microphone during the discussion
- ❑ **Slides** and other useful links – in the **”materials”** section. They will also be made available after the bootcamp.
- ❑ Joined from the **web** and **having a problem?** Refresh your page! Not solved? Contact Tatsiana in the chat.



Programme of today

10.00 – 10.05 Introduction to SSHOC and CLARIN | Ricarda Braukmann

10.05 – 10.20 The SSHOC Training Toolkit - A Discovery Tool | Ricarda Braukmann

Short break with music

10.30 – 11.00 Hands-on session CLARIN | Darja Fiser

Short break with music

11:10 – 11.40 Bringing events online | Tanya Yankelevich

Programme of today

10.00 – 10.05 Introduction to SSHOC and CLARIN | Ricarda Braukmann

10.05 – 10.20 The SSHOC Training Toolkit - A Discovery Tool | Ricarda Braukmann

Short break with music

10.30 – 11.00 Hands-on session CLARIN | Darja Fiser

Short break with music

11:10 – 11.40 Bringing events online | Tanya Yankelevich



SSHOC
social sciences & humanities open cloud



Horizon 2020
European Union Funding
for Research & Innovation

Type of action & funding:
Research and Innovation action
(INFRAEOSC-04-2018)

Partners: 45

(20 beneficiaries + 25 LTPs)

SSH ESFRI Landmarks and Projects
& international SSH data infrastructures

Project budget:
€ 14,455,594.08

Duration: 40 months
(January 2019 – 30 April 2022)

Project website:
www.SSHopencloud.eu



Objectives:

- creating the social sciences and humanities (**SSH**) part of European Open Science Cloud (**EOSC**)
- maximising **re-use** through **Open Science** and **FAIR** principles (standards, common catalogue, access control, semantic techniques, training)
- interconnecting existing and new infrastructures (clustered cloud infrastructure)
- establishing appropriate **governance model** for SSH-EOSC

EXPECTED IMPACT



The Social Sciences and Humanities are seamlessly integrated in the European Open Science Cloud



Availability of an EU-wide, easy-to-use SSH Open Marketplace, where tools and data are openly accessible



EU-wide availability of high quality "cloud ready" SSH tools and high quality SSH data



EU-wide availability of trusted and secure access mechanisms for SSH data, conforming to EU legal requirements



State of the art Research Infrastructure in several pilot domains advanced through dedicated SSH data pilots cluster projects



Maximising reuse through Open Science and FAIR principles (standards, common catalogue, access control, semantic techniques, training)



Building the SSH Open Cloud

- ❑ **SSHOC Marketplace** of tools and services for the Social Sciences and Humanities

→ sshopencloud.eu/ssh-open-marketplace

- ❑ **SSHOC Training Toolkit** is a **discovery tool** for new and existing **training resources**

→ sshopencloud.eu/join-ssh-training-community

- ❑ **SSHOC Training Community** for knowledge exchange on best practices

→ sshopencloud.eu/join-ssh-training-community

- ❑ **SSHOC Workshops, Webinars and Bootcamps**

→ sshopencloud.eu/events



What is it?

- Common Language Resources and Technology Infrastructure
- **ESFRI** ERIC status since 2012, Landmark since 2016

Aim

Provide 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)
- to **advanced tools** to discover, explore, exploit, annotate, analyse or combine data, wherever they are located

How?

Through a **single sign-on** environment that serves as an ecosystem for **knowledge sharing**



CLARIN & Open Science

- promotion for the **sharing and re-use of data through sustainable data registries**
- all integrated data is available in **open access** for research purposes
- support for **linguistic diversity**
 - data covering more than 1500 languages
 - tools for many languages
 - language resources in all modalities
- adherence to **FAIR data principles**
 - Findable, Accessible, Interoperable, Re-usable
 - promotion of responsible data science

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: clarin.eu/resource-families

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!

Who are you?

Go to menti.com

& enter the code

40 94 93



Programme of today

10.00 – 10.05 Introduction to SSHOC and CLARIN | Ricarda Braukmann

10.05 – 10.20 The SSHOC Training Toolkit - A Discovery Tool | Ricarda Braukmann

Short break with music

10.30 – 11.00 Hands-on session CLARIN | Darja Fiser

Short break with music

11:10 – 11.40 Bringing events online | Tanya Yankelevich

SSHOC Launches
Toolkit
for Trainers in SSH



training-toolkit.sshopencloud.eu

SSHOC Training Toolkit

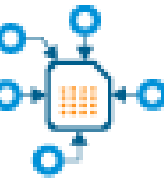
training-toolkit.sshopencloud.eu

- Collection of available materials for SSH trainers
- Materials to reuse, develop and improve your SSH trainings
 - Searchable inventory
 - Materials, modules, slides, didactics, videos, games, etc.
 - Various topics including Open Science, RDM etc.
- Launched on April 20th 2020
- training-toolkit.sshopencloud.eu

SSHOC Training Toolkit

training-toolkit.sshopencloud.eu

- The Toolkit currently lists 41 sources and 76 items for trainers
 - Source = *organisation, group or project providing training materials.*
 - Item = *the individual material that is provided by a given source.*



Overview

Search for sources and items



Search

Content - Topics

training-toolkit.sshopencloud.eu

Didactics

FAIR data

Research Data Management

Text encoding

Open Science

Survey data

Programming

...

Content - Types of materials

training-toolkit.sshopencloud.eu

Courses

Videos

Workshops

Games

Items

Slides

e-learning
modules

Reports

Webinars

...

Demo

training-toolkit.sshopencloud.eu

Use case

“I am looking for materials to aid researchers interested in language data.”

Demo

training-toolkit.sshopencloud.eu



SSHOC
social sciences & humanities open cloud

Training Toolkit

Overview

Search for sources and items



Search

The SSHOC Training toolkit provides an inventory of SSH training materials. More information can be found on the [SSHOC Training website](#).

Search entities

Displaying 1 - 117 of 117

Source

ID	Title	Description	Curated topics
----	-------	-------------	----------------

Entity type

-  Item (76)
-  Source (41)

Demo

training-toolkit.sshopencloud.eu



Overview

language research



Search

The SSHOC Training toolkit provides an inventory of SSH training materials. More information can be found on the [SSHOC Training website](#).

Search entities

Displaying 1 - 14 of 14

Source

ID **Title**

Description

Curated topics

Entity type

-  Item (76)
-  Source (41)

Search entities

Displaying 1 - 14 of 14

Source			
ID	Title	Description	Curated topics
230	CLARIN Depositing Services	One of the fundamental services of the CLARIN infrastructure is making sure that language resources can be archived and made available to the community in a reliable manner.	Open Science , Research data management/FAIR data , Text encoding and TEI
229	CLARIN Resource Families	The aim of the CLARIN Resource Families initiative is to provide a user-friendly overview of the available language resources in the CLARIN infrastructure for researchers from digital humanities, s	Research data management/FAIR data , Text encoding and TEI

Item				
ID	Title	Description	Curated topics	Source of item
263	CLARIN-DK-UCPH	The mission of CLARIN-DK is to provide easy and sustainable access for scholars in the humanities and social sciences to digital language data (in	Research data management/FAIR data , Text encoding and TEI	CLARIN Depositing Services

Demo

Search entities

Displaying 1 - 14 of 14

Source

ID	Title
230	CLARIN Depositing Services
229	CLARIN Resource Families

Item

ID	Title
263	CLARIN-DK-UCPH

Entity type

- Item (12)
- Source (2)

Collections

- Training Toolkit (13)

Intended audience

- Data creator (13)
- Data steward (12)
- Data provider (11)
- Researcher (7)
- Service provider (7)
- Student (2)
- Trainer (1)

Language

- English (11)
- German (3)
- Danish (1)
- Dutch (1)
- Greek (1)
- Spanish (1)

Source of item

- CLARIN Depositing Services (6)
- CLARIN Resource Families (4)
- FD Mentor (1)
- The Programming Historian | Programming Historian (1)

Curated topic

- Research data management/FAIR data (13)
- Text encoding and TEI (12)
- Data visualization (1)
- Open Science (1)
- Programming with R (1)

Topic

- Long time archiving of language resources and tools (6)
- Manually curated language resource overviews, curated by discipline (5)
- Dynamic Network Analysis (1)
- Temporal Network Analysis (1)

Discipline

- All disciplines (5)
- corpus linguistics (5)
- Digital humanities (3)
- computational linguistics (2)
- Digital history (2)
- Political Science (2)
- Social Sciences (2)
- Computational Linguistics (1)
- Phonetics (1)
- Phonology (1)
- Teaching (1)

Format

- Items (6)
- Course (1)
- E-learning module (1)

Search entities

Displaying 1 - 14 of 14

Source

ID	Title	Description	Curated topics
230	CLARIN Depositing Services	One of the fundamental services of the CLARIN infrastructure is making sure that language resources can be archived and made available to the community in a reliable manner.	Open Science , Research data management/FAIR data , Text encoding and TEI
229	CLARIN Resource Families	The aim of the CLARIN Resource Families initiative is to provide a user-friendly overview of the available language resources in the CLARIN infrastructure for researchers from digital humanities, s	Research data management/FAIR data , Text encoding and TEI

Item

ID	Title	Description	Curated topics	Source of item
263	CLARIN-DK-UCPH	The mission of CLARIN-DK is to provide easy and sustainable access for scholars in the humanities and social sciences to digital language data (in	Research data management/FAIR data , Text encoding and TEI	CLARIN Depositing Services

CLARIN Resource Families



The aim of the CLARIN Resource Families initiative is to provide a user-friendly overview of the available language resources in the CLARIN infrastructure for researchers from digital humanities, social sciences and human language technologies. The overviews are organized according to the types of data in the resources and include listings sorted by language.

The listings 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, CLARIN Resource Families provides an overview of other existing valuable language resources which have not yet been integrated in the infrastructure.

CLARIN Resource Families also provides 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.

Responsible organisation

[CLARIN ERIC](#)

Disciplines

[Digital humanities](#)
[corpus linguistics](#)
[computational linguistics](#)
[Social Sciences](#)
[Political Science](#)
[Digital history](#)

Intended audience

[Data creator](#)
[Data provider](#)
[Data steward](#)
[Researcher](#)

Access points

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

Collections

Training Toolkit



Resource families

Introduction

The aim of the CLARIN Resource Families initiative is to provide a user-friendly overview of the available language resources in the CLARIN infrastructure for researchers from digital humanities, social sciences and human language technologies. The overviews are organized according to the types of data in the resources and include listings sorted by language.

The listings 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 we provide an overview of other existing valuable language resources which have not yet been integrated in the infrastructure.

We 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.

We currently offer overviews of 10 corpora families, 5 families of lexical resources, and 3 tool families:

Corpora

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

Lexical Resources

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

Tools

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

Questions?

training-toolkit.sshopencloud.eu



→ Contact us at training@sshopencloud.eu

→ Join our **Training Community** at sshopencloud.eu/join-ssh-training-community

Programme of today

10.00 – 10.05 Introduction to SSHOC and CLARIN | Ricarda Braukmann

10.05 – 10.20 The SSHOC Training Toolkit - A Discovery Tool | Ricarda Braukmann

Short break with music

10.30 – 11.00 Hands-on session CLARIN | Darja Fiser

Short break with music

11:10 – 11.40 Bringing events online | Tanya Yankelevich

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-hr 1.0

Size: 14.1 million tokens
Annotation: tokenised, MSD-tagged, lemmatised, named entities
Licence: CC-BY

Croatian

The corpus contains minutes of the National Assembly of the Republic of Croatia and currently covers its VIth mandate from 15 November 2016 to 21 November 2018. 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 [KonText](#) and [noSketchEngine](#), as well as through a [dedicated webpage](#).

🔍 Concordancer

📄 Download

Slovenian parliamentary corpus siParl 1.0

Size: 227.8 million tokens
Annotation: tokenised, PoS-tagged, lemmatised
Licence: CC BY

Slovenian

The corpus contains Slovenian parliamentary debates from 1990 to 2018. It differs from the SlovParl 2.0 corpus (listed below) in that it contains only basic meta-data about the speakers, a typology of sessions and structural and editorian annotations.

The corpus is available for download from the CLARIN.SI repository and through the concordancers [KonText](#) and [noSketchEngine](#).

🔍 Concordancer

📄 Download



- 2 For both corpora, check their descriptions to see that:
 - In terms of linguistic annotation, both corpora are **annotated for syntactic and morphological features** (“MSD-tagged”), **lemmatized** and **marked for named entities**.
 - In terms of extra-linguistic annotation, both corpora are marked for **speaker metadata (gender, age, education, party affiliation)**.
 - The CC-BY licence shows that the corpus is publicly available, either for **download** or **on-line querying**.

- 3 Let's start by analysing the Slovenian corpus. First click on [Slovenian parliamentary corpus ParlaMeter-sl 1.0](#) in the CLARIN Resource Families. This takes you to the record for this corpus in the CLARIN.SI repository:

- 4 The CLARIN.SI repository shows how the corpus has to be cited to ensure proper authorship attribution, and offers a **persistent identifier** for the resource – <http://hdl.handle.net/11356/1208>.

- 5 The corpus can be queried via two concordancers – **KonText** and **noSketch Engine**. Both offer very versatile search environments in which complex queries can be narrowed down on the basis of the **speaker metadata (age, party affiliation, etc)**.

- 6 Let's query the corpus by using the **noSketch Engine**. In the repository, click on the downward arrow next to “noSketch” and then select search.

The screenshot shows the noSketch Engine interface for the ParlaMeter-si (parliament) corpus. The left sidebar contains navigation links: Home, Search, Word list, Corpus info, My jobs, and User guide. The main content area displays the following information:

ParlaMeter-si (parliament)
 Korpus parlamentarnih razprav Republike Slovenije: Mandat VII (2014-08 - 2018-06) // Corpus of parliamentary debates of the Republic of Slovenia: Mandate VII (2014-08 - 2018-06)

Counts	General info	Lexicon sizes	Tags legend	Lempos suffixes
Tokens: 40,987,516	Corpus description: Document	word: 263,007	samostalnik: S.*	samostalnik: -s
Words: 34,882,499	Language: Slovenian	tempos: 109,066	glagol: G.*	glagol: -g
Sentences: 1,833,147	Encoding: UTF-8	tag_en: 1,080	pridevnik: P.*	pridevnik: -p
Paragraphs: 133,287	Compiled: 12/31/2018 22:16:01	tag: 1,080	pristov: R.*	pristov: -r
Documents: 1,338	Tagset: Description	lc: 228,682	zaimek: Z.*	zaimek: -z
		norm: 228,682	predlog: D.*	predlog: -d
		lemma: 104,247	veznik: V.*	veznik: -v
		lemma_lc: 100,467	členek: L.*	členek: -l
			medmet: M.*	medmet: -m

- 7 Let's recall our task: we're interested how parliamentary speakers are represented in the corpus in terms of gender. In other words, how many words of the total 34,882,499 are spoken by female parliament speakers and how many by male speakers?

- 8 We can figure this out by creating a **word list** and narrowing it down to the “Female” subcorpus.

The screenshot shows the noSketch Engine interface for creating a word list. The left sidebar contains navigation links: Home, Search, Word list, Corpus info, My jobs, and User guide. The main content area displays the following options:

Word list options

Corpus: ParlaMeter-si (parliament)

Subcorpus: Female (circled in red) [info](#) [create new](#)

Search attribute: word

use n-grams. Value of n: from 2 to 2

hide/nest sub-n-grams

Filter options:

Filter word list by: Regular expression:

Minimum frequency: 5

Maximum frequency: 0 (0 = no maximum frequency)

Whitelist: no file selected

Blacklist: no file selected [format](#)

Include non-words

- 9 After clicking on **Make word list**, we get the [result](#) for female speakers. We repeat the procedure for the “Male” subcorpus and [see](#) that the male speakers say 2.5 times more words than their female counterparts.

Repeat the procedure for the Croatian ParlaMeter corpus.

- What is the gender division in terms of words between male and female speakers in this corpus?
- Is the difference greater or smaller than that in the Slovenian corpus?

Additional Task

We can also construct word lists for individual word classes, such as nouns, verbs, adjectives, etc.

Which are the most frequent nouns used by the speakers in the Slovenian corpus?

- a. Under search attribute, change from “word” to “tag_en”. This specifies that you’re searching for parts of speech rather than individual words.
- b. In the filter option “Regular expression”, write N.*. This specifies that you’re searching for all nouns.
- c. Under output options, select “lemma” under “Change output attributes”. This ensures that all inflectional variants are all accounted for under a single base form of the word.
- d. Click [here](#) to see the result for such a query.

Repeat the procedure for the Croatian ParlaMeter corpus. Are the results similar to the Slovenian corpus?

Research bite

In the Slovenian ParlaMeter corpus, the most frequent topics among the female speakers are *health* and *labour*, *family* and *social affairs*, which are followed by *public administration and education*, *science and sport*.

Most of the 100 top-ranking keywords uttered by female speakers, on the other hand, could not be classified into a single topic because they were used either to achieve a *stylistic effect*, were general words that were used in *multiple topics*, such as descriptive adjectives or legal terms, or *ideological expressions*, all of which indicate a more discursive, debating *style* of the male speakers, but could also stem from the fact that the leading roles in that term were predominantly held by male members of parliament (Source: [Parlamenteer – a Corpus of Contemporary Slovene Parliamentary Proceedings](#) by Darja Fišer, Nikola Ljubešić and Tomaž Erjavec).

Use Case 2

Creating a linguistically annotated corpus of 19th century English novels

Background

The digital humanities provide a new conception of the world of literature. Not only is this world larger – the sheer volume of the material we can access is unprecedented – but it is open to levels of analysis that could never be achieved by human brainpower alone. Hierarchies and themes fade into the background as patterns and networks emerge. These methods simultaneously divide texts into new categories and connect them to each other to form new wholes. (Source: [When computers read: Literary analysis and digital technology](#) by Sarah Jones)

CLARIN Resources and services used

- [Virtual Language Observatory \(VLO\)](#)
- [Language Resource Switchboard](#)
- [WebLicht](#)

Teacher

How can my students create an annotated corpus of 19th century English novels from scratch in an easy-to-use online environment?



Researcher

Can you help me find resources and tools to research the stylistic differences between 19th century female and male novelists?

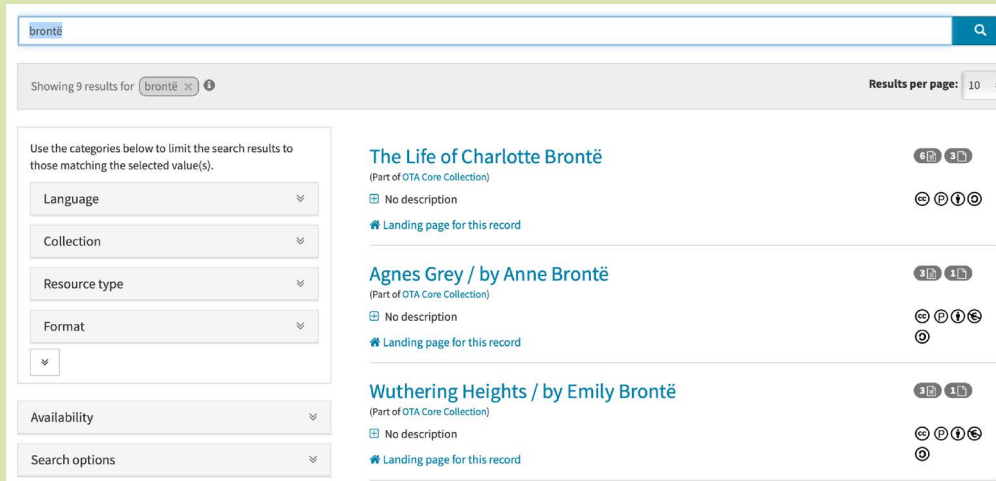


Step-by-step guide

The **Language Resource Switchboard (LRS)** aims at bridging the gap between resources (as identified in the [VLO](#), [Federal Content Search](#), and the [CLARIN Virtual Collection](#)) and tools that can process these resources in one way or another. For a given resource in question, it identifies all tools that can process the resource. It then sorts the tools in terms of the tasks they perform, and presents a task-oriented list to the user. Users can then select and invoke the tool of their choosing. (Source: Adapted from [The Language Resource Switchboard](#) by Claus Zinn)

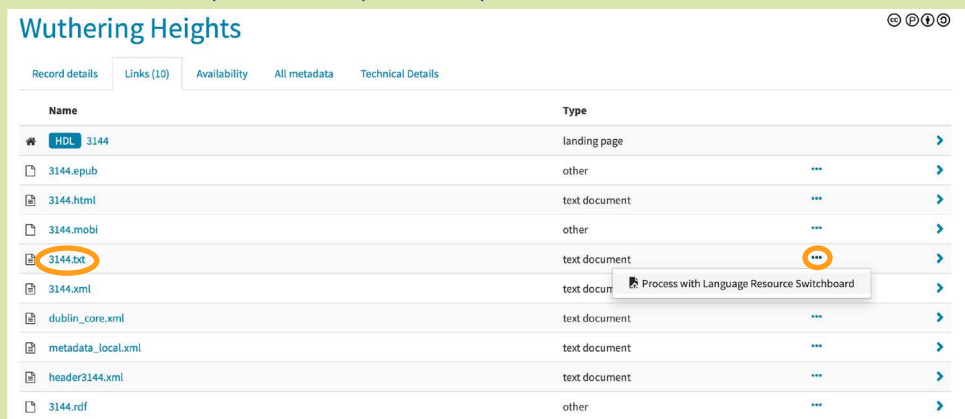


1 Search the VLO with the simple query Brontë.

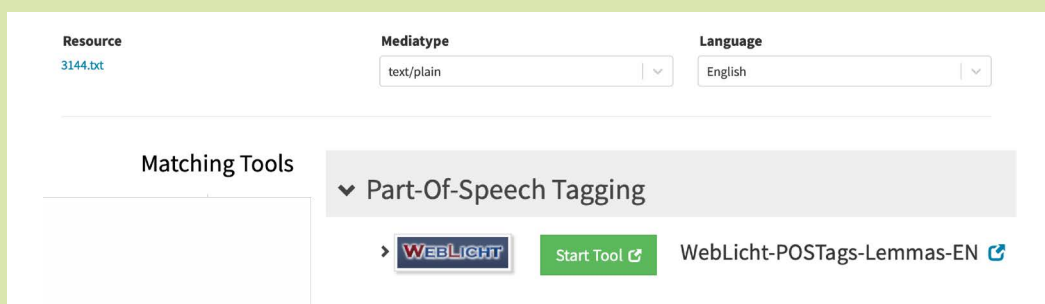


This query gives you VLO records for 19th century English novels by the Brontë sisters, such as Wuthering Heights by Emily Brontë, The Tenant of Wildfell Hall by Anne Brontë, and Jane Eyre by Charlotte Brontë.

2 In each VLO record under the “Links” tab, we can find the complete novels in the form of .txt files. Each file can be processed through the Language Resource Switchboard by clicking on “...” next to the .txt file (in our case, 3144.txt).



3 The first step in linguistic annotation typically involves **part-of-speech tagging**, with which each word in a corpus is assigned a part of speech, like *noun*, *verb*, and *adjective*. In the LRS, we see that part-of-speech tagging is performed by WebLicht.



4 In the WebLich application, we select *PoS tags/lemmas* under “Available Annotations for English Plain text”

5 After clicking on *Run Tools*, the entire *Wuthering Heights* novel becomes tagged for parts of speech.

The annotated novel can now be queried like a regular corpus either for simple words or by using the [TIGERSearch corpus query language](#). To find all the adjectives in the newly tagged corpus, type `[pos = /J.*/]` in the Query field. Make sure to enclose the query in square brackets. Try visualizing the results. Which are the most and least common adjectives in the novel? Hint: In the Statistics visualisation under “Add/remove columns”, try adding the values PoS and lemma.

WebLich also allows you to create additional annotation chains (“New Chain”). Try to repeat the task above by also tagging [The Tenant of Wildfell Hall](#) for **parts of speech** as well as for **named entities**. By creating several annotation chains in this way, you are able to create a full-fledged linguistically annotated corpus, consisting of several novels which were originally in simple plain text.



Research bite

By analysing a corpus of novels by Charles Dickens, Mahlberg et al. (2019) have studied how fictional dialogue is used by the author to create a sense of realism and authenticity. The authors have shown that Dickens consistently writes dialogue characterised by linguistic features that fictional and real people share (e.g., question fragments, set expressions conveying politeness and vagueness), which contributes to a sense of naturalness to speech in fiction. By contrast, the range of frequent word combinations in fictional dialogue is more limited than that in spoken fiction, so it is possible that literature adds an iconic or heightened meaningful effect onto these forms. Particularly, shorter lexical combinations (e.g. *I mean, you know* and *I don't know*) are less frequent in fiction. (Source: [Speech-bundles in the 19th-century English novel](#) by Michaela Malmberg, Viola Wiegand, Peter Stockwell, and Anthony Hennessey)

Programme of today

10.00 – 10.05 Introduction to SSHOC and CLARIN | Ricarda Braukmann

10.05 – 10.20 The SSHOC Training Toolkit - A Discovery Tool | Ricarda Braukmann

Short break with music

10.30 – 11.00 Hands-on session CLARIN | Darja Fiser

Short break with music

11:10 – 11.40 Bringing events online | Tanya Yankelevich



BRINGING EVENTS ONLINE

From rethinking your event
strategy to choosing the right
platform

Tanya Yankelevich, Training Coordinator | LIBER



www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE

SESSION DESIGN AND CONTENT

- RE-think and RE-structure
to maintain value
- Design event brief with learning goals and the audience.



www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE

SESSION DESIGN AND CONTENT: example

Original event GOALS:

- Connect with peers
- Learn from experts
- Get feedback on their work
- Practice new skills



www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE

SESSION DESIGN AND CONTENT: example

From GOALS to ONLINE FORMAT DESIGN:

Connect with peers → 1 on 1 & small group discussions

Learn from experts → short lectures with Q&A

Get feedback on their work, practice new skills → assignments
& small groups



www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE

SESSION DESIGN AND CONTENT: polling

Pre-recorded content: YES or NO?



www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE

SESSION DESIGN AND CONTENT cont.

Introduction

- Housekeeping rules are more important online than offline

Ice-Breakers

- Ask a question (favorite food / drink / pet / villain / ice cream flavor; if you were able to travel through time, where would you go etc.)
- Ask participants to share a picture that has a story (e.g. shoes they are wearing – they wore them on their favorite vacation, etc.) or to hold up their tea/coffee mug and tell the group about it.



www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE

ENGAGEMENT

- Polling (built-in or an additional tool, like mentimeter or slido) and voting
- Quizzes
- Breakouts, discussion and collective problem-solving
- Structured Q&A (e.g. in Google Slides)



www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE

ENGAGEMENT

- Games
- Chat and social media (incl. photos, etc)
- Contests



www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE

LOGISTICS: What not to forget

- Choose the right online platform
- Remote access map
- Soundchecks and Guidelines for speakers/trainers
- Moderators vs Hosts
- When something goes wrong
- Timing and breaks



www.liberconference.eu
@LIBERconference |  #LIBER2020



Thank You for Participating!

EBSCO



ExLibris
a ProQuest Company

DIGITAL
science

OpenEdition

RETRONews
Le site de presse de la BnF

www.liberconference.eu
@LIBERconference |  #LIBER2020

LIBER
★ 2020 ★
ONLINE