

# D6.5 Final ELEXIS interoperability report including interaction with CLARIN/DARIAH services

Author(s): Simon Krek, Gregor Leban,  
Iztok Kosem, Andraž Repar, Anže  
Sršen

Date: 31 July 2022



H2020-INFRAIA-2016-2017

Grant Agreement No. 731015

ELEXIS - European Lexicographic Infrastructure

#### NUMBER AND NAME OF THE DELIVERABLE



Deliverable Number:	D6.5
Dissemination Level:	PUBLIC
Delivery Date:	31 July 2022
Version:	1.0
Author(s):	Simon Krek, Gregor Leban, Iztok Kosem, Andraž Repar, Anže Sršen

Project Acronym: ELEXIS  
Project Full Title: European Lexicographic Infrastructure  
Grant Agreement No.: 731015

### Deliverable/Document Information

Project Acronym: ELEXIS  
Project Full Title: European Lexicographic Infrastructure  
Grant Agreement No.: 731015

### Document History

Version Date  
1.0 28/7/2022

Changes/Approval

Author(s)/Approved by  
Simon Krek, Gregor Leban, Iztok  
Kosem, Andraž Repar, Anže Sršen

## Contents

1. Executive summary .....	5
2. ELEXIS tools and services available for lexicographers .....	6
2.1. Sketch Engine .....	6
2.2. OneClick Dictionary .....	6
2.3. Lexonomy .....	6
2.4. Publex .....	7
2.5. Elexifier .....	7
2.6. EDiE: ELEXIS Dictionary Evaluator .....	7
2.7. ElexiLink: ELEXIS Dictionary Matrix .....	8
3. ELEXIS tools and services available for NLP researchers .....	8
3.1. Clusty .....	8
3.2. VerbAtlas .....	8
3.3. SyntagNet .....	9
3.4. MultiMirror .....	9
3.5. NAISC .....	9
3.6. BabelNet Linker .....	10
3.7. Cross The Word .....	10
4. ELEXIS tools and services available for everybody .....	10
4.1. Elexifinder .....	10
4.2. News feed .....	11
5. Interaction with CLARIN and DARIAH services .....	11
5.1. DARIAH-CAMPUS .....	11
5.2. CLARIN.SI repository .....	11
5.3. ELEXIS Knowledge Centre (CLARIN) .....	12
5.4. ELEXIS Association .....	14

## 1. Executive summary

This report provides an overall description of ELEXIS tools and services developed by the end of the ELEXIS project. Its goal is to provide an overview of ELEXIS infrastructure as an ecosystem of interconnected elements. Some of the elements are available as standalone pieces of software available on GitHub or other repositories, however, the part of the infrastructure exhibiting a high level of interoperability are the online dictionary tools represented in graphical form in Figure 1. In addition to dictionary tools, other resources for lexicographers and other users, e.g. VerbAtlas and SyntagNet, have been developed and made available. The third aspect crucial to ensuring interoperability is a standard for encoding dictionaries - TEI Lex-0, to which the ELEXIS project is making a considerable contribution.

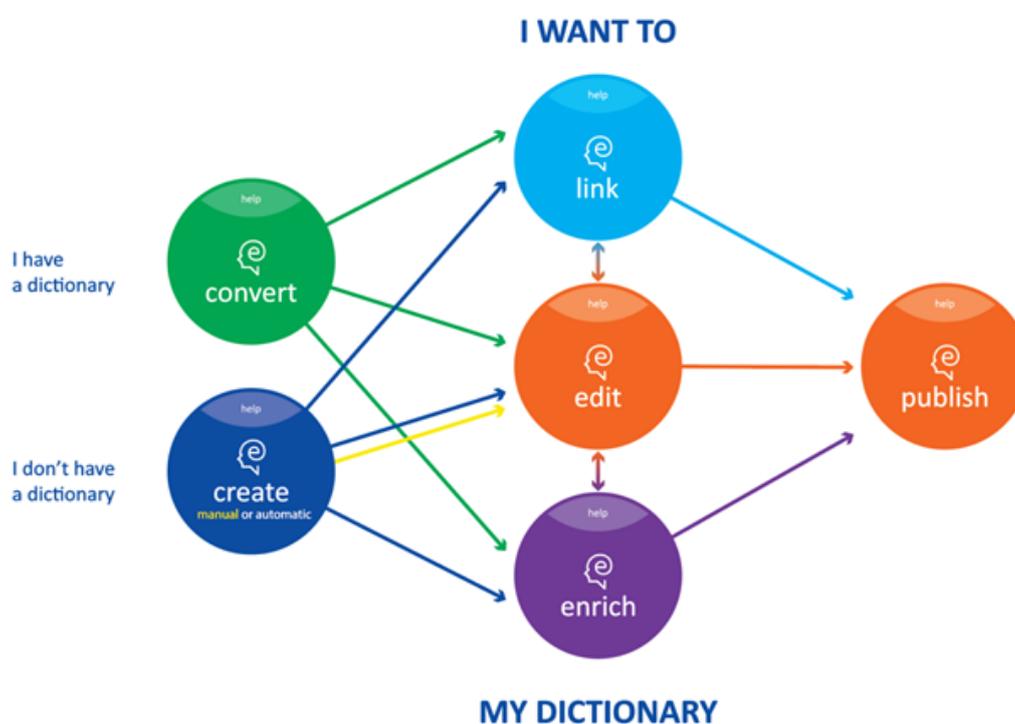


Figure 1 Graphic Guide to ELEXIS Dictionary Tools

Another important aspect of ELEXIS is providing necessary information to lexicographic community about the research literature and related developments. This is offered via Elexifinder, and relatedly LexBib bibliographic library, and ELEXIS News Feed, with interoperability of these services being improved constantly.

Since this is the final interoperability report of the ELEXIS project it is important to understand where tools and services will be maintained after the end of the project. We use three colours to indicate different repositories and infrastructures where ELEXIS tools and services will be kept alive:

- BLUE: CLARIN.SI infrastructure as part of CLARIN ELEXIS Knowledge Centre
- GREEN: DARIAH-CAMPUS for the ELEXIS Curriculum
- BROWN: GitHub repositories of ELEXIS and other project partners



## 2. ELEXIS tools and services available for lexicographers

### 2.1. Sketch Engine

The Sketch Engine corpus query, corpus building and corpus management system allows users to build and work with 550+ text corpora in over 90 languages and 30 scripts. Sketch Engine contains a number of unique tools to analyse large corpora of up to 60 billion words. Each user can benefit from fully automated dictionary-building functionality.

**The access to Sketch Engine was funded by the EU through the ELEXIS project between 1 April 2018 and 1 April 2022.** The access was provided at no cost to academic institutions and ELEXIS observers, and applied to non-commercial use. At the end of the free access period, a total of 461 institutions have been using the tool. Details about the use of Sketch Engine during the ELEXIS project are included in the deliverable D8.4 (Periodic assessment of LEX1, LEX2 and LEX3 – final report).

Additional information about the Sketch Engine tool:

- Web site (with post-ELEXIS information): <https://www.sketchengine.eu/elexis/>

### 2.2. OneClick Dictionary

OneClick Dictionary (OCD) is a dictionary drafting module. It interconnects a corpus management system (e.g. SketchEngine, noSketch Engine) or even excel sheets with ELEXIS dictionary writing and online dictionary publishing system Lexonomy and provides an automatically created dictionary draft (e.g. headwords, wordforms, collocations, examples), to be post-edited in Lexonomy by the lexicographer. OneClick Dictionary enables lexicographers to shift lexicographers work and intellectual input into the post-editing phase instead of manually analyzing the input data before creating a dictionary draft. Hence, the tool is not limited to professionals but also designed for spontaneous lexicography – small projects of lexicographic nature such as glossaries and domain-specific wordlists and dictionaries often prepared by teachers or other professionals without formal training in lexicography.

Source code:

- ELEXIS GitHub repository: <https://github.com/elexis-eu/ocd>



Additional information about the OneClick Dictionary module:

- Deliverable 4.2: [Dictionary Drafting Module](#)

### 2.3. Lexonomy

Lexonomy is a cloud-based dictionary-writing and also online-dictionary-publishing system which is highly scalable to adapt to large dictionary projects as well as small lexicographic works such as editing and online publishing of domain-specific glossaries or terminology resources. Lexonomy already interacts with Sketch Engine and other online concordancers, e.g. Kontext, NoSketchEngine as a part of CLARIN.SI infrastructure. Sketch Engine can push lexicographic data into Lexonomy to create automatically generated dictionary drafts and Lexonomy can pull data from Sketch Engine's or other corpora during the entry editing process.

Online installation:

- Web site: <https://lexonomy.elex.is/>
- AVAILABLE AFTER 31 July 2022 AS PART OF CLARIN ELEXIS Knowledge Centre



Source code:

- ELEXIS GitHub repository: <https://github.com/elexis-eu/lexonomy>



Additional information about Lexonomy:

- Deliverable 4.1: [Online Dictionary Post-Editing and Presentation Module](#)
- Deliverable 4.4: [Dictionary Enhancement Module](#)

## 2.4. Publex

Publex is a software that allows you to publish your XML dictionary data on the web. No special technical skills are needed for publication of dictionaries apart from knowing your data, especially how it is annotated. The data upload works by importing the data from a Git repository. When importing the data, Publex captures all the different tags, attributes and associated attribute values the dictionary is annotated with. Publex offers three different search options for the published dictionaries: a lemma list search, a fulltext search and an advanced search. Each published dictionary can be accessed directly via its own URL.

- Online installation: <http://publex.uni-trier.de>
- AVAILABLE AFTER 31 July 2022 AS PART OF CLARIN ELEXIS Knowledge Centre



Additional information about Publex:

- ELEXIS Curriculum: [Publishing Legacy Dictionaries with Publex](#)



## 2.5. Elexifier

Elexifier is a cloud-based dictionary conversion service. It uses advanced XML parsing and machine learning techniques to help you convert your PDF and XML dictionaries in a standardized machine-readable format. Users can upload their PDF and custom XML dictionaries to Elexifier, define mapping rules for XML transformation or create a machine learning training set for PDF conversion and download the transformed XML or PDF dictionary in a TEI-compliant file format.

Online installation:

- Web site: <https://elexifier.elex.is/>
- AVAILABLE AFTER 31 July 2022 AS PART OF CLARIN ELEXIS Knowledge Centre



Source code (ELEXIS GitHub repository):

- Front end: <https://github.com/elexis-eu/elexifier>
- Back end: <https://github.com/elexis-eu/elexifier-api>
- Landing page: <https://github.com/elexis-eu/elexifier-landing-page>
- PDF conversion: <https://github.com/elexis-eu/elexifier-pdf>



Additional information about Elexifier:

- Deliverable 1.4: [ELEXIS Conversion Tools](#)

## 2.6. EDiE: ELEXIS Dictionary Evaluator

EDiE is a tool evaluating the availability and usability of linked lexical resources and dictionaries published, using the ELEXIS dictionary API, which are accessible when using the ELEXIS infrastructure. It allows users to assess different aspects of dictionaries based on their metadata and entries. Furthermore, aggregated metrics over dictionaries of interests/contexts let users compare different dictionaries for their specific use cases.

Source code (ELEXIS GitHub repository):

- Software: <https://github.com/elexis-eu/edie>
- Dictionary API: <https://elexis-eu.github.io/elexis-rest/elexis.html>



Additional information about EDiE:

- Publication: <https://elex.is/wp-content/uploads/Edie-4.pdf>

### 2.7. ElexiLink: ELEXIS Dictionary Matrix

ElexiLink is a single-page web application built with Angular (a fast and scalable development platform). It utilizes Lexonomy API to acquire language and dictionary definitions, and search for computed mappings between dictionaries. The application uses a modular service that seamlessly loads all the relevant data on each route/path change. To speed up loading, data acquired from Lexonomy API is temporarily cached and appropriately invalidated when changes occur.

Online installation:

- Web site: <https://matrix.elex.is/>
- AVAILABLE AFTER 31 July 2022 AS PART OF CLARIN ELEXIS Knowledge Centre



Source code:

- ELEXIS GitHub repository: <https://github.com/elexis-eu/elexilink>



## 3. ELEXIS tools and services available for NLP researchers

### 3.1. Clusty

Clusty is an innovative algorithm designed to perform lexical-semantic analytics for NLP: sense clustering. The team at the Linguistic Computing Laboratory of the Sapienza University of Rome investigated clustering approaches which allow to effectively and easily scale across languages whilst dropping the requirement of large amounts of data which is typically needed when employing neural networks. Clusty's results can be used for improving word sense disambiguation systems.

The demonstration of the efficacy of Clusty for performing one of the most challenging tasks in natural language processing, sense clustering, is presented in deliverable D3.1.

Source code:

- ELEXIS GitHub repository: <https://github.com/elexis-eu/D3.1>



Additional information about Clusty:

- Deliverable D3.1: [Semantic Analytics for NLP: Sense Clustering](#)

### 3.2. VerbAtlas

VerbAtlas is a novel large-scale manually-crafted semantic resource for wide-coverage, intelligible & scalable Semantic Role Labeling. The goal of VerbAtlas is to manually cluster WordNet synsets that share similar semantics into sets of semantically-coherent frames. The main features are:

- 466 semantically-coherent frames using 26 cross-frame VerbNet-inspired semantic roles for their argument structure.
- Available both for download and via RESTful API.
- Full coverage of WordNet 3.0 verb synsets (13,000+).
- Complete linkage to BabelNet 4.0, which supports 280+ languages (new version to come later this year!).

- Manual mapping to PropBank of all CoNLL-2009 and CoNLL-2012 dataset occurrences (5000+ mappings).
- Selectional preferences: the superconcept most probably associated with a semantic role in a frame (e.g. food for the patient role of the EAT frame).
- Default/shadow arguments: arguments logically implied or already incorporated into a verb.
- Implicit arguments: arguments that are implicit in the argument structure of a verb.

Additional information about VerbAtlas:

- API documentation: <https://verbatlas.org/api-documentation>
- Download: <https://verbatlas.org/download>
- Publication (EMNLP-IJCNLP): <https://aclanthology.org/D19-1058/>

### 3.3. SyntagNet

SyntagNet is a manually-curated large-scale lexical-semantic combination database which associates pairs of concepts with pairs of co-occurring words. The goal of SyntagNet is to capture sense distinctions evoked by syntagmatic relations, hence providing information which complements the essentially paradigmatic knowledge shared by currently available Lexical Knowledge Bases such as WordNet. Its main features are: Wide coverage, with 78,000 noun-verb and noun-noun lexical combinations extracted from the English Wikipedia and the British National Corpus.

High-quality, fully manual disambiguation for all of the lexical combinations, according to the WordNet 3.0 sense inventory. A resulting Lexical Knowledge Base made up of 88,019 semantic combinations linking 20,626 WordNet 3.0 unique synsets with a relation edge. A user-friendly web interface for looking up terms and their lexical-semantic combinations, with complete linkage to BabelNet 4.0.

Additional information about SyntagNet:

- API documentation: <http://syntagnet.org/api-documentation>
- Download: <http://syntagnet.org/data/SyntagNet1.0.zip>
- Publication (EMNLP-IJCNLP): <https://aclanthology.org/D19-1359/>

### 3.4. MultiMirror

Neural Cross-lingual Word Alignment for Multilingual Word Sense Disambiguation: MultiMirror is a cross-lingual sense projection approach for multilingual WSD based on a novel discriminative word alignment model, capable of jointly aligning all source and target tokens with each other, surpassing its competitors across several language combinations. The sense-tagged datasets it produces lead a standard WSD classifier to achieve state-of-the-art performances on established benchmarks in French, German, Italian, Spanish and Japanese.

Source code:

- SapienzaNLP GitHub repository: <https://github.com/SapienzaNLP/multimirror>



Additional information about MultiMirror:

- Publication (IJCAI): <https://www.ijcai.org/proceedings/2021/539>

### 3.5. NAISC

NAISC 1.0 ('NAISC' means 'links' in Irish and is pronounced 'nashk') is a tool for linking datasets and was created by the SFI Insight Centre for Data Analytics and the ELEXIS project. NAISC serves as a

system for aligning RDF datasets: It takes as input 2 RDF documents (referred to as ‘left’ and ‘right’) and outputs an alignment (set of RDF triples) between these two documents. NAISC typically relies on a configuration, which is a JSON document.

Source code:

- Insight Centre GitHub repository: <https://github.com/insight-centre/naisc>



Additional information about NAISC:

- Presentation [slides](#)
- Video introduction: [https://www.youtube.com/watch?v=maYEv8rG0\\_k](https://www.youtube.com/watch?v=maYEv8rG0_k)

### 3.6. BabelNet Linker

The BabelNet Linker is a linking web service which produces a mapping between two dictionary definitions in a cross-lingual scenario. The BabelNet-linker API allows a dictionary to be linked to BabelNet at definition level. Specifically, this API allows a definition in any language to be mapped to a semantically-equivalent English definition in BabelNet by relying on state-of-the-art Transformer-based architectures. Importantly, this API will make it possible to map the dictionaries made available within the ELEXIS Consortium at definition level by pivoting through BabelNet.

Source code:

- GitHub repository: <https://github.com/elexis-eu/BabelNet-linker>



Additional information about BabelNet Linker:

- Deliverable D2.4: [Cross-lingual Lexical Resource Linking Web Service](#)

### 3.7. Cross The Word

CrossTheWord is a crossword puzzle game for Android with small and big crossword puzzles, available for free download via the GooglePlay Store.

Features:

- Hundreds of automatically generated crosswords (in constant growth!)
- Power-ups to boost your game experience and help you solve the unsolvable!
- A dynamic tap and swipe interface to surf through crosswords
- A subgame of lexical substitution to earn extra points!

Crosswords are currently available in English only.

Source code:

- GitHub repository: <https://github.com/elexis-eu/CrossTheWord>



Additional information about lexicographic Cross the Word:

- Google Play: <https://play.google.com/store/apps/details?id=it.uniroma1.lcl.crucy>

## 4. ELEXIS tools and services available for everybody

### 4.1. Elexifinder

The search tool ELEXIFINDER is dedicated to helping lexicographers and other researchers find scientific output in lexicography and related fields. It enables users to search through papers and

videos, using concepts, i.e. words or set of words with a Wikipedia page, and various other conditions, e.g. source (conference etc.), author, language etc. Each paper/video is linked to its page where the users can download or view it.

Online installation:

- Web site: <https://finder.elex.is/>
- AVAILABLE AFTER 31 July 2022 AS PART OF CLARIN ELEXIS Knowledge Centre 

Additional information about Elexifinder:

- On the ELEXIS web page: <https://elex.is/tools-and-services/elexifinder/>
- Publication: [https://elex.link/elex2019/wp-content/uploads/2019/09/eLex\\_2019\\_29.pdf](https://elex.link/elex2019/wp-content/uploads/2019/09/eLex_2019_29.pdf)
- LexBib Zotero Group: <https://www.zotero.org/groups/lexbib>
- LexBib Wikibase : <https://lexbib.elex.is>

## 4.2. News feed

Lexicographic news feed is an ELEXIS service that uses the Event Registry API to extract latest news articles identified to be related to lexicography. News articles are extracted from 30,000 news sources, and over 35 languages are currently supported.

Additional information about lexicographic News feed:

- Web site: <https://elex.is/tools-and-services/lexicographic-news/>

## 5. Interaction with CLARIN and DARIAH services

### 5.1. DARIAH-CAMPUS

DARIAH-Campus is a discovery framework and hosting platform for learning resources. It is maintained by DARIAH, which is a European Research Infrastructure Consortium (ERIC) and a European Strategy Forum on Research Infrastructures (ESFRI) Landmark. The legal status of DARIAH offers a stable framework for sustaining the ELEXIS learning resources beyond the end of ELEXIS as a funded H2020 project, while the DARIAH-Campus' alignment with the European Open Science Cloud (EOSC) and the Social Sciences and Humanities (SSHOC) Marketplace will guarantee future discoverability and interoperability of the ELEXIS Curriculum.

The staging url of the ELEXIS Curriculum: <https://elexis.humanistika.org/curriculum/the-elexis-curriculum>

The published url of the ELEXIS Curriculum: <https://campus.dariah.eu/curriculum/the-elexis-curriculum>

### 5.2. CLARIN.SI repository

After the end of the ELEXIS project, the tools and services that have been developed within ELEXIS will be hosted by CLARIN.SI infrastructure, and (some of the) activities started in the project will continue within the proposed ELEXIS Knowledge Centre for Lexicography as a part of CLARIN ERIC knowledge infrastructure (for more on ELEXIS-KC see deliverable D6.5 - Final ELEXIS interoperability report including interaction with CLARIN/DARIAH services).

Within the CLARIN.SI infrastructure, a special metadata collection was created to ensure consistency in the metadata that is associated with data that has been contributed to ELEXIS. See: <https://www.clarin.si/repository/xmlui/handle/11356/1479>. All the tools in the ELEXIS infrastructure pull the metadata from this collection.

### 5.3. ELEXIS Knowledge Centre (CLARIN)

At the end of ELEXIS project, the following partner and observer institutions decided to propose the formation of a CLARIN Knowledge Centre (ELEXIS-KC) for Lexicography within CLARIN infrastructure:

	Country	CLARIN	ELEXIS KC partner
1	Slovenia	<a href="#">CLARIN.SI</a>	Jožef Stefan Institute
2	Austria	<a href="#">CLARIAH-AT</a>	Austrian Academy of Sciences (ACDH-CH)
3	Latvia	<a href="#">CLARIN-LV</a>	Institute of Mathematics and Computer Science, University of Latvia
4	Denmark	<a href="#">CLARIN-DK</a>	University of Copenhagen
5	Lithuania	<a href="#">CLARIN-LT</a>	Vytautas Magnus University
6	Denmark	<a href="#">CLARIN-DK</a>	Danish Society for Language and Literature
7	The Netherlands	<a href="#">CLARIAH-NL</a>	Dutch Language Institute
8	Croatia	<a href="#">HR-CLARIN</a>	Institut za hrvatski jezik i jezikoslovlje
9	Estonia	<a href="#">CLARIN Estonia</a>	Institute for the Estonian language
10	Portugal	<a href="#">PORTULAN CLARIN</a>	NOVA CLUNL
11	Norway	<a href="#">CLARINO</a>	The National Library of Norway
12	Italia	<a href="#">CLARIN-IT</a>	Università Cattolica del Sacro Cuore
13	Germany	<a href="#">CLARIN-D</a>	University of Trier
14	Ukraine	--	Institute of Philology of Borys Grinchenko Kyiv University
15	Italia	<a href="#">CLARIN-IT</a>	Institute for Computational Linguistics A. Zampolli, Italian National Research Council
16	Hungary	<a href="#">HunCLARIN</a>	Research Institute for Linguistics, Hungarian Academy of Sciences
17	Bulgaria	<a href="#">CLaDA-BG</a>	Bulgarian Academy of Sciences

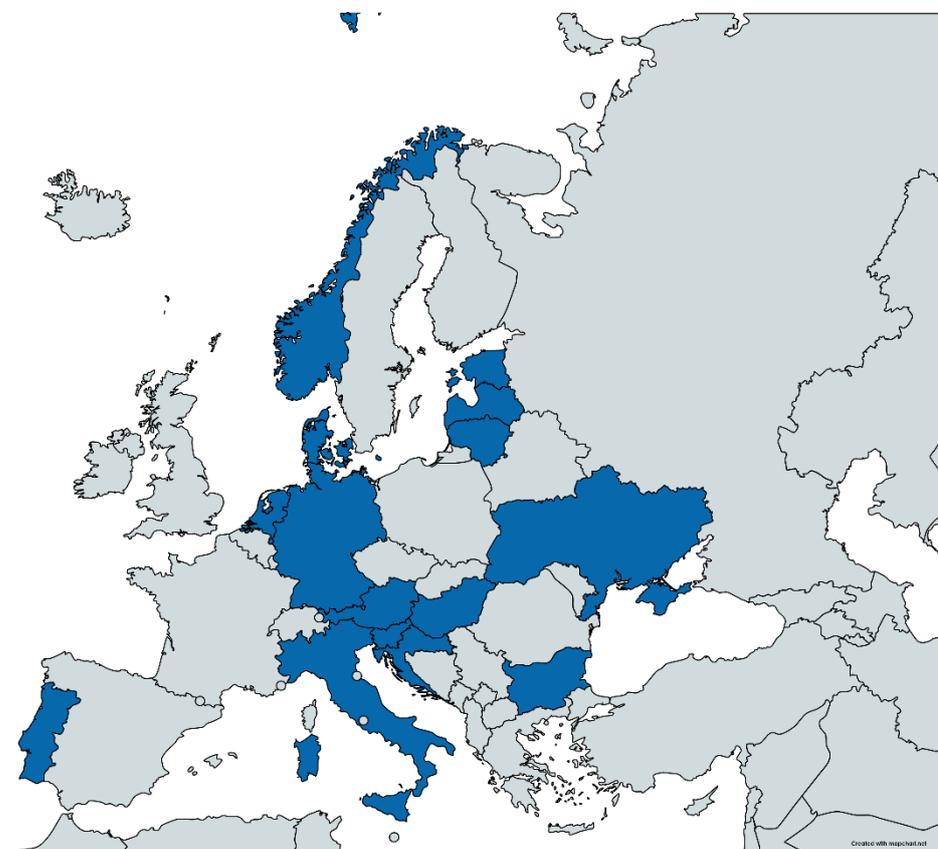
The ELEXIS Knowledge Centre for Lexicography will operate as a distributed virtual centre supported by five leading institutions in existing CLARIN National Consortia, and eight lexicographic institutions, seven of them from countries with established CLARIN consortia. The Knowledge Centre will be coordinated by the Jožef Stefan Institute (Ljubljana, Slovenia), and the main point of contact to CLARIN ERIC is Simon Krek ([simon.krek@ijs.si](mailto:simon.krek@ijs.si)).

ELEXIS-KC is supported by the following institutions:

- Jožef Stefan Institute (SLOVENIA, CLARIN.SI)
  - URL(s): <https://www.ijs.si/>, <http://www.clarin.si/info/about/>
  - expertise: Existing ELEXIS infrastructure, Helpdesk
- Austrian Centre for Digital Humanities and Cultural Heritage (ACDH-CH), Austrian Academy of Sciences (AUSTRIA, CLARIAH-AT)
  - URL(s): <https://www.oeaw.ac.at/acdh/acdh-ch-home>
  - expertise: Existing ELEXIS infrastructure, lexicography, corpus linguistics, language resources and text technologies
- Institute of Mathematics and Computer Science, University of Latvia (LATVIA, CLARIN-LV)

- URL(s): clarin.lv, ailab.lv, tezaurs.lv
- expertise: language resources and technology for Latvian, creation of WordNet-like and FrameNet-like lexical resources, computational lexicography, semantic parsing
- University of Copenhagen (DENMARK, CLARIN-DK)
  - URL(s): Centre for Language Technology: <https://cst.ku.dk/english/> Danish CLARIN: <https://clarin.dk/clarindk/forside.jsp>
  - expertise: Computational lexicography, wordnets, lexical-semantic resources for NLP, practical lexical projects for NLP
- Vytautas Magnus University (LITHUANIA, CLARIN-LT)
  - URL(s): <http://clarin-lt.lt/?lang=en>
  - expertise: Lithuanian language technology, corpus linguistics, lexicography, terminology
- Danish Society for Language and Literature (DENMARK)
  - URL(s): <https://dsl.dk/>, <https://ordnet.dk/>
  - expertise: lexicography (historical and contemporary) for the Danish language, computational lexicography and the creation of formal semantic lexicons for Danish (WordNet, FrameNet etc.)
- Dutch Language Institute (NETHERLANDS, CLARIAH-NL/CLARIAH-VL)
  - URL(s): <https://ivdnt.org/>; <https://portal.clarin.ivdnt.org/>; <https://kdutch.ivdnt.org>
  - expertise: lexicography (historical and contemporary) for the Dutch language, terminology, corpus linguistics, language resources
- Institut za hrvatski jezik i jezikoslovlje (CROATIA)
  - URL(s): <http://ihji.hr/>
  - expertise: lexicography, historical lexicography, etymology, semantics, terminology, general linguistics
- Institute for the Estonian language (ESTONIA)
  - URL(s): <https://portaal.eki.ee/>
  - expertise: lexicography, corpus lexicography, (multilingual) lexicographic resources for Estonian, corpus linguistics
- NOVA CLUNL (PORTUGAL)
  - URL(s): [www.clunl.fcs.unl.pt](http://www.clunl.fcs.unl.pt)
  - expertise: lexicography, terminology, standards, computational lexical semantics
- The National Library of Norway (NORWAY)
  - URL(s): <https://www.nb.no/sprakbanken/en/resource-catalogue/>, <https://www.nb.no/dh-lab/>
  - expertise: computational lexical semantics, Natural Language Processing, digital resources for written and spoken modern Norwegian and for medieval Norwegian Latin
- Università Cattolica del Sacro Cuore (ITALY)
  - URL(s): <https://centridiricerca.unicatt.it/circse-home?rdeLocaleAttr=en>
  - expertise: Digital Lexical and Textual Resources for Latin; Linguistic Linked Open Data; Natural Language Processing; valency and subcategorization
- Institute of Philology of Borys Grinchenko Kyiv University (UKRAINE)
  - URL(s): <https://partner.kubg.edu.ua/2012-02-09-16-42-36/institute-of-philology/65-institute-of-philology/557-institute-chairs.html#department-of-romance-philology-and-comparative-typological-linguistics>
  - expertise: digital neography, digital terminography, linguistic innovation modelling, corpus linguistics, digital humanities, digital learning
- University of Trier/Trier Center for Digital Humanities
  - URL(s): <https://tcdh.uni-trier.de/en>, [www.woerterbuchnetz.de](http://www.woerterbuchnetz.de)
  - expertise: (retro)digitization, XML-markup, interlinking and online publication of dictionaries, federated content search, digital humanities

- Hungarian Research Centre for Linguistics
  - URL(s): [www.nytud.hu](http://www.nytud.hu)
  - expertise: language resources and technology for Hungarian, lexicography, computational lexicography, terminology, corpus linguistics, general linguistics
- Institute of Information and Communication Technologies, Bulgarian Academy of Sciences
  - URL(s): <https://www.iict.bas.bg/EN/>
  - expertise: language technology and resources for Bulgarian; language modelling; lexical semantics; ontologies; electronic dictionaries; lexicon-corpus interface
- Institute for Computational Linguistics A. Zampolli, Italian National Research Council
  - URL(s): CLARIN-IT: national CLARIN B data center -- <https://ilc4clarin.ilc.cnr.it/> (hosting lexical data in different formats)
  - expertise: computational lexicography, standardization, (meta)models for lexical data, training courses (language resources for lexicography) legacy resources: ItalWordNet, SIMPLE Lexicon for Italian, lexical resources for ancient languages



#### 5.4. ELEXIS Association

Before the end of the project, legal department at Jožef Stefan Institute prepared a „partnership agreement“ as one of the possible post-project sustainability measures. The agreement defines basic parameters: the aims of the association, governance and administration. This option is meant for those institution that would like to be actively involved in further development of (lexicographic) data, tools and services, search for funding options and general collaboration. The draft of the agreement is included in Appendix 1 in this document.

In the initial survey conducted in June and July 2022, the following institutions expressed interest in becoming a member of the Association:

<b>University</b>	
University of Copenhagen	Denmark
Sapienza University of Rome	Italy
Università Cattolica del Sacro Cuore, Milan	Italy
Institute of the Czech National Corpus, Charles University, Prague	Czechia
Mykolas Romeris University	Lithuania
NOVA FCSH, Lisbon	Portugal
Institute of Journalism of Borys Grinchenko Kyiv University	Ukraine
Faculty of History and Philology, University of Tirana	Albania
National University of Ireland, Galway (NUI Galway)	Ireland
<b>Academy/Institute</b>	
Jožef Stefan Institute	Slovenia
Institute of the Estonian Language	Estonia
Society for Danish Language and Literature (DSL)	Denmark
Hungarian Research Centre for Linguistics	Hungary
Istituto di Linguistica Computazionale - CNR	Italy
Academia das Ciências de Lisboa	Portugal
Croatian Academy of Sciences and Arts	Croatia
"A. Philippide" Institute of Philology, Romanian Academy – Iasi	Romania
Institute for Language & Speech Processing, ATHENA Research Centre	Greece
Eurac Research	Italy
National Library of Norway	Norway
<b>Companies</b>	
K Dictionaries - Lexicala	Israel
Babelscape	Italy
Lexical Computing	Czechia

It is expected that the detailed scope of services and activities will be agreed between the members of the association in 6-12 months after the end of the project.

Appendix 1:

**PARTNERSHIP AGREEMENT (DRAFT)**

**BY AND BETWEEN THE FOLLOWING PARTIES:**

Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia

(hereinafter referred to as "Coordinator")

and

Party/Member

jointly and individually hereinafter referred to as "Party or Member" or "Parties or Members"

**Article 1**

**Preliminary observations**

The Parties to this Agreement initially note that by concluding this Agreement, the Parties enter into a contractual partnership called EUROPEAN LEXICOGRAPHIC INFRASTRUCTURE Association or ELEXIS Association.

**Article 2**

**Objectives and activities**

The objective of the ELEXIS Association is organisation and coordination of activities related to lexicography, and activities related to natural language processing tasks on the topic of semantics, insofar they are of interest to lexicography.

To this end, the ELEXIS Association seeks to engage in:

- the development of lexicographic tools and (web) services,
- the exchange and/or linking of lexicographic data,
- the development of lexicographic standards facilitating the interoperability of lexicographic data,
- scientific research on lexicography and semantics, including the organisation of conferences, the engagement in research projects and similar,
- the exchange of expertise including the organisation of training, research visits and similar,

- all other activities which contribute to its aims.

### **Article 3**

#### **Services and activities**

The detailed scope of services and/or activities, in line with the objectives in Article 2, including possible financial arrangements and allocation of ownership and terms of exercising, protection and exploitation of results, must be agreed between the Parties of this Agreement and specified in a separate written document.

### **Article 4**

#### **Board of Representatives**

The Parties shall establish the Board of Representatives, within thirty (30) days after the conclusion of this Agreement. The Board of Representatives shall be composed of one (1) authorised representative of each Party. Parties appoint their representatives in accordance with and in the manner determined by their internal affairs acts.

The Board of Representatives:

- The Board of Representatives advises and decides on important matters relating to this Agreement, ELEXIS Association and its activities.
- It decides on activities, fees and other contributions, establishes the text of the Agreement, accepts or rejects new candidate Members, establishes the action lines and performs other activities.
- The Board of Representatives shall be chaired by the President of the Board.
- The Board of Representatives selects the President among its members by a majority of votes, for a two-year period.
- The President selects a Deputy among the members of the Board of Representatives who takes the duties of the President in her/his absence.
- It is called by the President to convene the Board of Representatives at least once a year.
- All Members must receive at least four (4) weeks' advance written notice of Board of Representatives session.
- Board of Representatives decisions are based on a simple majority of the votes of the representatives. The session is quorate if the majority of representatives of the Board of Representatives are present.

### **Article 5**

#### **Membership**

An institution wishing to become a Party to this Agreement and a Member of the ELEXIS Association must apply for membership by sending an application letter to the Coordinator.

The admission of the candidate institution is decided by the Board of Representatives by a majority vote, not later than three (3) months after the receipt of application.

## **Article 6 Coordinator**

The Coordinator shall have the following functions:

- keeping the address list of Members and other contact persons updated and available;
- administration and provision of the President of the Board of Representatives and follow-up of Board's decisions;
- transmission of any relevant documents and information connected with this Agreement;
- any other duties that are agreed on between the Board of Representatives and the Coordinator.

## **Article 7 Administrative fee**

The Coordinator shall be entitled to receive an annual administrative fee from each Party of this Agreement for performing the activities as set in Article 6.

The annual administrative fee is x EUR.

The Coordinator will issue a separate invoice to each Party of this Agreement.

## **Article 8**

Nothing in this Agreement shall constitute or be deemed to constitute any formal business organisation or legal entity between the Parties. Each Party shall act as independent contractor and not as the agent of any of the other Parties.

## **Article 9 Language**

This Agreement is drawn up in English which language shall govern all documents, notices and meetings for its application and/or extension or in any other way relative thereto.

## **Article 10 Applicable law**

This Contract shall be construed in accordance with and governed by the laws of Belgium.

## **Article 11 Dispute settlement**

The Parties shall endeavour to settle their disputes amicably.

All disputes, controversies or differences which may arise between the Parties, out of or in relation to or in connection with this Agreement and cannot be settled amicably shall be finally settled by the courts of Brussels, Belgium

**Article 12**  
**Amendments**

Amendments or changes to this Agreement shall be valid only if made in writing and signed by an authorized signatory of each of the Parties.

**Article 13**  
**Effectiveness, Term and Termination**

This Agreement shall become effective upon signature by each of the Parties hereto.

This Agreement shall remain in force for a term of x (x) years.

A Party to this Agreement may, by written notice to all other Parties, withdraw from this Agreement and such notice will take effect x (x) months from the date of that notice.

Date

Signatures