Linking and the Dictionary Matrix

John P. McCrae
Data Science Institute, NUI Galway
ELEXIS Architecture

Members’ Dictionaries

Retro-digitisation

TEI-Lex0 Dictionaries

Continuous Quality Evaluation

Linking Tool (Naisc)

REST Interface

Dictionary Matrix

NLP Tools

Sketch Engine

Dictionary Authoring (Lexonomy)

Crowdsourcing

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731015.
What is it?

- Uniform input standard for the dictionary matrix
- Collects basic dictionary and entry information
- Supports many formats and use cases
- Called by the dictionary matrix
Routes into the Dictionary Matrix

- Retrodigitization Tool (D1.3)
- Conversion Tool (D1.4)
- TEI Lex-0
- REST Interface (D2.2)
- Dictionary Matrix
- TEI Lex-0
- Ontolex-Lemon

Non-ELEXIS Implementations

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015.
ELEXIS Protocol for accessing dictionaries (Draft)

Download OpenAPI specification: Download

URL: https://elexis.eu/ | License: Public Domain

This protocol allows data to be shared with the ELEXIS platform and should be implemented by all providers of data to the ELEXIS platform.

Get dictionaries

List all of the dictionaries that are available at this endpoint.

Responses

- **200 Success**
- **500 Internal Server Error**

About the dictionary

Get the metadata about the dictionary, including the conditions.

PATH PARAMETERS

- **dictionary**
  
  String Identifier of the dictionary to describe.

[Documentation](https://elexis-eu.github.io/elexis-rest/elexis.html)
The REST API workflow

Get Dictionaries ➔ About the Dictionary ➔ List All Lemmas ➔ Entry As JSON
OR
Get Dictionaries ➔ About the Dictionary ➔ Headword Lookup ➔ Entry As RDF
OR
Get Dictionaries ➔ About the Dictionary ➔ Headword Lookup ➔ Entry As TEI
About a dictionary

**Release**: Public, Non-Commercial, Research, Private

**Source Language**: The language of the headwords

**Target Languages**: The languages of the content

**Genre**: General, Learner’s, Etymological, Specialist, Historical, Spelling or Terminological Dictionary (or any combination)

**License**

**Title, Creator, Publisher, …** (plus 48 other properties)
Dictionary Entries

Minimal requirements

- ID
- A lemma
- A language
- One or more part-of-speech from Universal Dependencies:

<table>
<thead>
<tr>
<th>ADJ: adjective</th>
<th>INTJ: interjection</th>
<th>PUNCT: punctuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADP: adposition</td>
<td>NOUN: noun</td>
<td>SCONJ: subordinating conjunction</td>
</tr>
<tr>
<td>ADV: adverb</td>
<td>NUM: numeral</td>
<td>SYM: symbol</td>
</tr>
<tr>
<td>AUX: auxiliary</td>
<td>PART: particle</td>
<td>VERB: verb</td>
</tr>
<tr>
<td>CCONJ: coordinating conjunction</td>
<td>PRON: pronoun</td>
<td>X: other</td>
</tr>
<tr>
<td>DET: determiner</td>
<td>PROPN: proper noun</td>
<td></td>
</tr>
</tbody>
</table>
ELEXIS REST Server

Available by end of 2019

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731015.
ELEXIS Linking
Activity involving mental or physical effort done in order to achieve a purpose or result.

'he was tired after a day's work in the fields' + More example sentences + Synonyms

1.1 Mental or physical activity as a means of earning income; employment.

'I'm still looking for work' + More example sentences + Synonyms

1.2 The place where one is employed.

'I was returning home from work on a packed subway' + More example sentences

1.3 The period of time one spends in paid employment.

'he was going to the theatre after work' + More example sentences

1.4 West Indian [count noun] A job.

'I decided to get a work' + More example sentences + Synonyms

2 A task or tasks to be undertaken.

'they made sure the work was progressing smoothly' + More example sentences + Synonyms

2.1 The materials for a task.

'she frequently took work home with her' + More example sentences

2.2 (in combination or with modifier works) British Activity involving construction or repair.

'extra costs caused by additional building works' + More example sentences

1. (heading, uncountable) Employment.

1.1 Labour, occupation, job. [quotations ▼]

My work involves a lot of travel.

2. The place where one is employed.

He hasn't come home yet, he's still at work.

3. One's employer

'I want to go to the R.E.M. reunion concert but I'm not sure if my work will let me off.'

2. (heading, uncountable) Effort.

1. Effort expended on a particular task.

Holding a brick over your head is hard work. It takes a lot of work to write a dictionary.

2. Sustained human effort to overcome obstacles and achieve a result.

We know what we must do. Let's go to work.

3. Something on which effort is expended.

There's lots of work waiting for me at the office.

4. (physics) A measure of energy expended in moving an object; most commonly, force times distance.

Work is done against friction to drag a bag along the ground.

5. (physics, more generally) A measure of energy that is usefully extracted from a process.

We don't have much time. Let's get to work piling up those sandbags.

3. Sustained effort to achieve a goal or result, especially overcoming obstacles. [quotations ▼]

We've got some paperwork to do before we can get started. The piece was decorated.

4. (countable) A literary, artistic, or intellectual production.

it is a work of art.

the poetic works of Alexander Pope

4. (countable) A fortification.

William the Conqueror fortified many castles, throwing up new ramparts, bastions and...
Monolingual Linking

Dictionary 1

Links

Dictionary 2

Available in 2020

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731015.
The Dictionary Matrix

Search Results
Bank (English Noun)
1. “The land alongside or sloping down to a river or lake.” - From Oxford Dictionaries
   “sloping land (especially the slope beside a body of water);” - From Princeton WordNet
2. “A bank is a financial institution that accepts deposits from the public and creates credit” - From Wikipedia
Linking Workflow

**Blocking**: Identify Entries to Link

**Lens**: Extract Relevant Facts

**Features**: Analyze Textual Content

**Score** similarity with advanced AI

**Match** dictionaries holistically

**Constraints** reject unlikely or impossible links

---

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731015.
Cross-lingual Linking

Available Summer 2021

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731015.
ELEXIS Quality Assurance
Validation and Quality Assurance

Technical
- XML Validity: 75%
- Normalized tags

Operational
- Uptime: 60%
- Ave. Response: 2000 ms

Scientific
- Sense per lemma: 3.2
- Definition Similarity: Low

Available Late 2021
Questions?

Or come see me at the booth.