

Encyclopædia Britannica;

James OR, A Fullerton

DICTIONARY

O F

A R T S and S C I E N C E S,

COMPILED UPON A NEW PLAN.

IN WHICH

The diferent SCIENCES and ARTS are digested into
distinct Treatises or Systems;

A N D

The various TECHNICAL TERMS, &c. are explained as they occur
in the order of the Alphabet.

ILLUSTRATED WITH ONE HUNDRED AND SIXTY COPPER

By a SOCIETY of GENTLEMEN in SCOTLA

IN THREE VOLUMES

VOL. I.

EDINBURGH

Printed for A. BELL and C. M

And sold by COLIN MACFARQUHAR, at

M.DCC

*frances: An AI-toolbox to
discover automatically insights
from Data Foundry collections*

Dr. Rosa Filgueira,
Assistant Professor,
Heriot-Watt University,
Email: R.Filgueira@hw.ac.uk

*frances**: New AI-toolbox to discover automatically insights

New ways to unlock the full value of NLS digital collections

- Objective 1: New facilities to run more complex text analysis queries → ML/NLP techniques.
- Objective 2: Full integration with NLS Data Foundry → Hide the large-scale text mining complexity
- Using the **Encyclopaedia Britannica** as the core dataset

(*) *Frances Wright* (September 6, 1795 – December 13, 1852)

frances: Providing Automatic ML Analysis

frances will provide **abstractions** to a variety of ML/NLP techniques
→ Extract complex knowledge without being an expert data scientist

- Train and use text embedding models
- Employ topic mining, sentiment analysis, text summarization
- Build knowledge graph(s) visualizing the results

Using the **Encyclopaedia Britannica** – *frances* will allow us automatically to

- Group similar articles
- Detect how articles have changed across editions
- Extract the relationships between articles
- Classify articles into different categories
- Summarise articles
- Analyse the sentiment expressed in an article

Overview of how the Encyclopaedia Britannica has changed over time



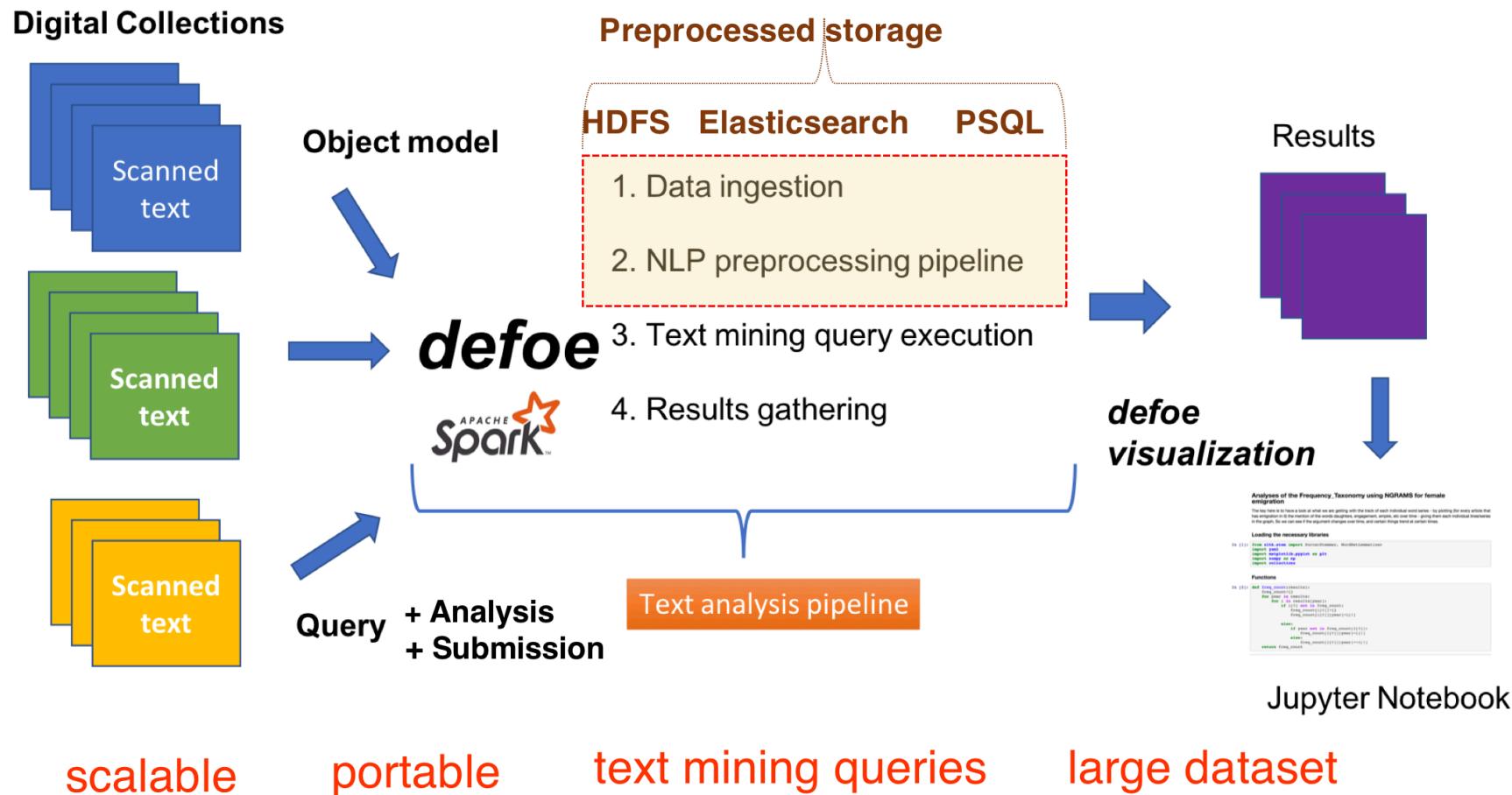
New suite of ML functionalities that can be used to analyse
any other Data Foundry collection

My Journey

- (1) From Semi-Structured EB information (ALTO & METS XML files) →**
Information Extraction → Knowledge Graph → Deep learning
Transformers → **To Augmented EB-Knowledge Graph with advance AI-methods**
- (2) Querying Augment EB Knowledge Graph:**
 - (1) Extracting information already stored in the EB-Knowledge Graph
 - (2) Processing information stored in the EB-Knowledge Graph **in parallel**
→ Create new data/results

Phase 1: Text Mining

defoe: scalable toolbox for historical research



<https://github.com/defoe-code/defoe>

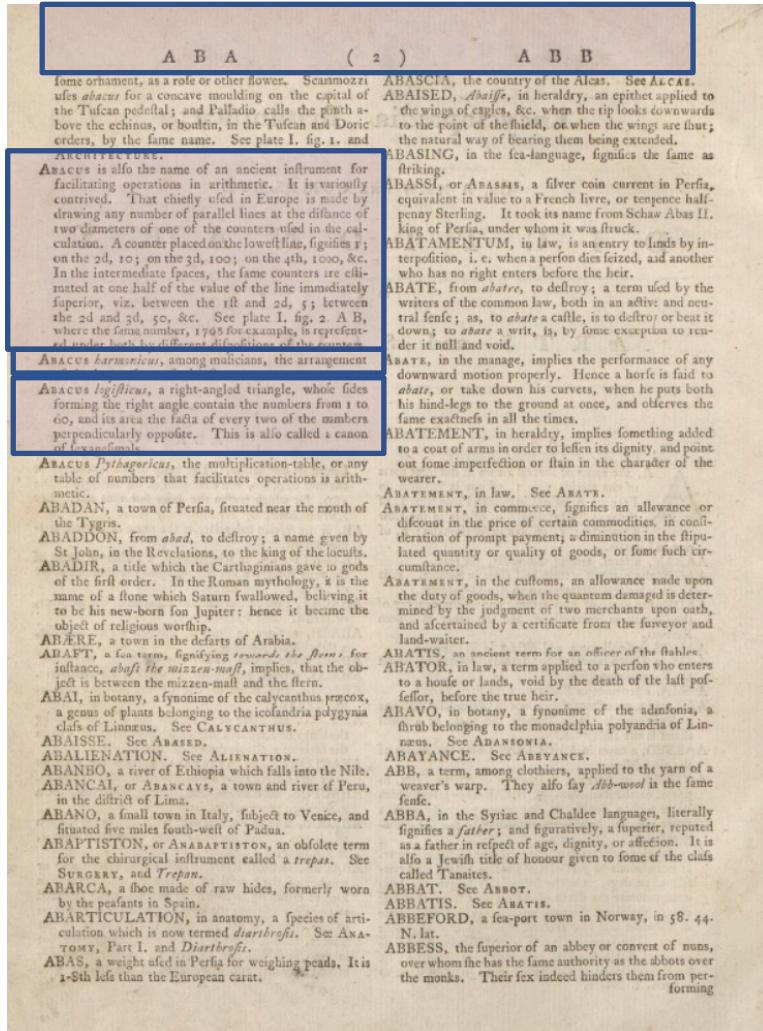
Phase 1: Text Mining

1.1. Improved **defoe Extract Terms query** → It extracts the Eb Terms by page and classify them between **articles** and **topics** (**Terms v.1**)

Extracted Terms Based on Heuristics --> Pages Layout & Text & Headers → Different heuristics for different EB editions. Using ALTO-XML information.

Phase 1: Text Mining

1. Detecting pages headers from ALTO XML
2. Using headers to classify terms into: Articles & Topics
3. Using ALTO Text for detecting the start of each article:
--> Starting a line with TERM UPPERCASE + “,” .



some ornament, as a rose or other flower. Scamozzi uses *abacus* for a concave moulding on the capital of the Tuscan pedestal; and Palladio calls the plinth above the echinus, or boutin, in the Tuscan and Doric orders, by the same name. See plate I. fig. 1. and

ARCHITECTURE. *ABACUS* is also the name of an ancient instrument for facilitating operations in arithmetic. It is variously contrived. That chiefly used in Europe is made by drawing any number of parallel lines at the distance of two diameters of one of the counters used in the calculation. A counter placed on the lowest line, figures 1; on the 2d, 10; on the 3d, 100; on the 4th, 1000, &c. In the intermediate spaces, the same counters are eliminated at one half of the value of the line immediately superior, viz. between the 1st and 2d, 5; between the 2d and 3d, 50, &c. See plate I. fig. 2. A *B*, where the same number, 1749 for example, is represented under both the different divisions of the counter.

ABACUS harminicus, among musicians, the arrangement

ABACUS logistica, a right-angled triangle, whose sides forming the right angle contain the numbers 3 to 60, and its area the ratio of every two of the numbers perpendicularly opposite. This is also called a canon of *logistica*.

ABACUS Pythagoricus, the multiplication-table, or any table of numbers that facilitates operations is arithmetic.

ABADAN, a town of Persia, situated near the mouth of the Tigris.

ABADDON, from *ab*, to destroy; a name given by St John, in the Revelations, to the king of the locusts.

ABADIR, a title which the Carthaginians gave to gods of the first order. In the Roman mythology, it is the name of a stone which Saturn swallowed, believing it to be his new-born son Jupiter: hence it became the object of religious worship.

ABARE, a town in the deserts of Arabia.

ABRAFT, a sea term, signifying towards the shores. For instance, *abaf* the mizen-mast, implies that the object is between the mizen-mast and the stern.

ABAL, in botany, a synonym of the calycinthus *procox*, a genus of plants belonging to the icocandria pygynia class of Linneus. See *CALYCANTHUS*.

ABALISE. See *ABASED*.

ABALIENATION. See *ALIENATION*.

ABANHO, a river of Ethiopia which falls into the Nile. **ABANCAL**, or *ABANCAYS*, a town and river of Peru, in the district of Lima.

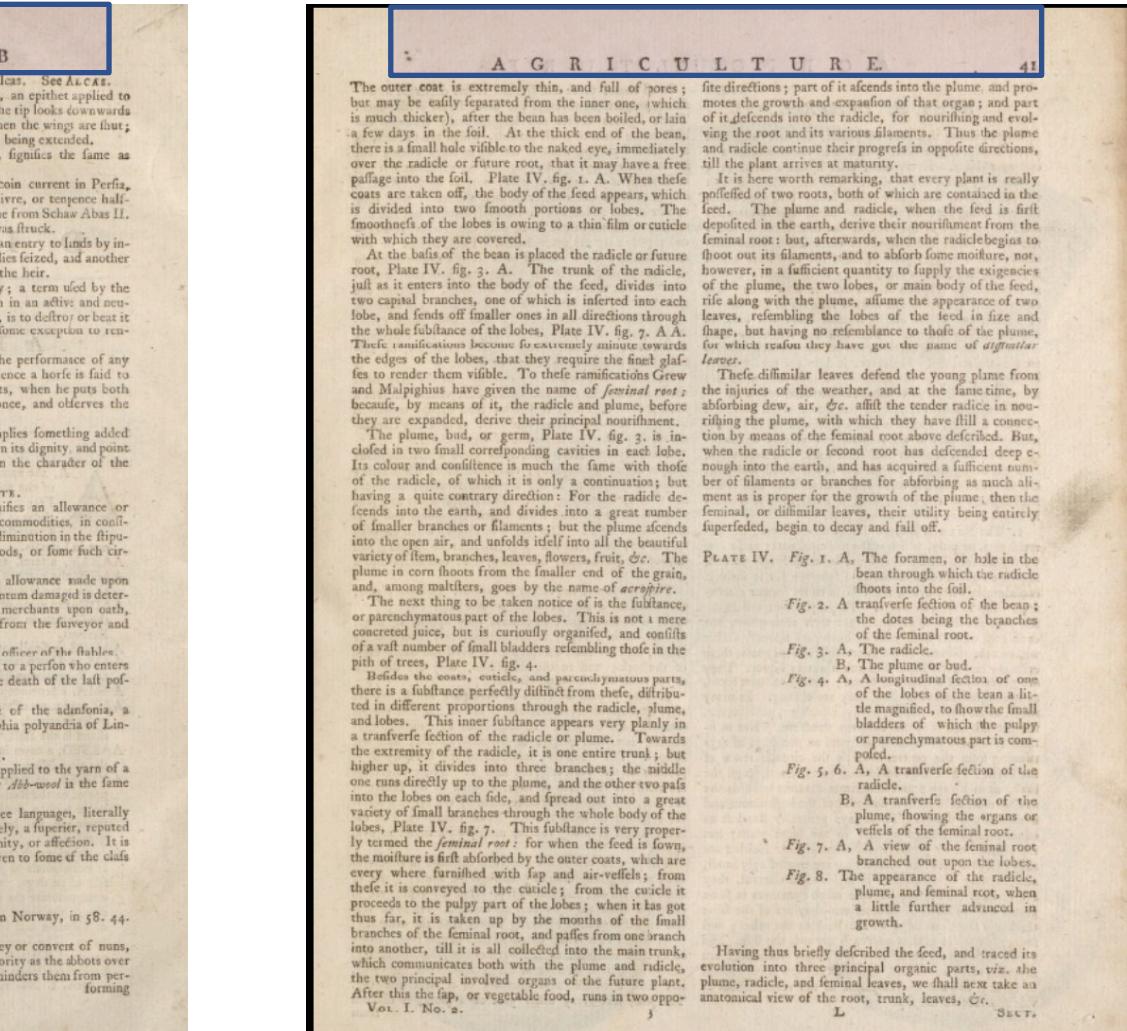
ABANO, a small town in Italy, subject to Venice, and situated five miles south-west of Padua.

ABAPTISTON, or *ANABAPTISTON*, an obsolete term for the chirurgical instrument called a *trepas*. See *SURGERY*, and *Trepas*.

ABARCA, a shoe made of raw hides, formerly worn by the peasants in Spain.

ABARTICULATION, in anatomy, a species of articulation which is now termed *diarthrosis*. See *ANATOMY*, Part I. and *Diarthrosis*.

ABAS, a weight used in Persia for weighing peads. It is 1-8th less than the European carat.



AGRICULTURE

The outer coat is extremely thin, and full of pores; but may be easily separated from the inner one, which is much thicker, after the bean has been boiled, or lain a few days in the foil. At the thick end of the bean, there is a small hole visible to the naked eye, immediately over the radicle or future root, that it may have a free passage into the soil. Plate IV. fig. 1. A. When these coats are taken off, the body of the feed appears, which is divided into two smooth portions or lobes. The smoothness of the lobes is owing to a thin film or cuticle with which they are covered.

At the basis of the bean is placed the radicle or future root, Plate IV. fig. 3. A. The trunk of the radicle, just as it enters into the body of the feed, divides into two capital branches, one of which is inserted into each lobe, and sends off smaller ones in all directions through the whole substance of the lobes. Plate IV. fig. 7. A. A. These ramifications become so extremely minute towards the edges of the lobes, that they require the finest glafes to render them visible. To these ramifications Grew and Malpighius have given the name of *feminal root*: because, by means of it, the radicle and plume, before they are expanded, derive their principal nourishment.

The plume, bud, or germs, Plate IV. fig. 3, is enclosed in two small corresponding cavities in each lobe. Its colour and consistence is much the same with those of the radicle, of which it is only a continuation; but having a quite contrary direction: For the radicle descends into the earth, and divides into a greater number of smaller branches or filaments; but the plume ascends into the open air, and unfolds itself into all the beautiful variety of stem, branches, leaves, flowers, fruit, &c. The plume in corn shoots from the smaller end of the grain, and, among multifiers, goes by the name of *ascropire*.

The next thing to be taken notice of is the fuitance, or parenchymatos part of the lobes. This is not a mere concreted juice, but is curiously organized, and consists of a vast number of small bladders resembling those in the pit of trees. Plate IV. fig. 4.

Besides the coats, cuticle, and parenchymatos parts, there is a substance perfectly distinct from these, distributed in different proportions through the radicle, plume, and lobes. This inner substance appears very plainly in a transverse section of the radicle or plume. Towards the extremity of the radicle, it is one entire trunk; but higher up, it divides into three branches; the middle one runs directly up to the plume, and the other two pass into the lobes on each side, and spread out into a great variety of small branches through the whole body of the lobes. Plate IV. fig. 7. This substance is very properly termed the *feminal root*: for when the feed is sown, the moisture is first absorbed by the outer coats, which are every where furnished with sap and air-vessels; from thence it is conveyed to the cuticle; from the cuticle it proceeds to the pulpy part of the lobes; when it has got thus far, it is taken up by the months of the small branches of the feminal root, and passes from one branch into another, till it is all collected into the main trunk, which communicates both with the plume and radicle, the two principal involved organs of the future plant.

After this the sap, or vegetable food, runs in two oppo-

site directions; part of it ascends into the plume, and promotes the growth and expansion of that organ; and part of it descends into the radicle, for nourishing and evolving the root and its various filaments. Thus the plume and radicle continue their progress in opposite directions, till the plant arrives at maturity.

It is here worth remarking, that every plant is really possessed of two roots, both of which are contained in the seed. The plume and radicle, when the seed is first deposited in the earth, derive their nourishment from the feminal root: but, afterwards, when the radicle begins to shoot out its filaments, and to absorb some moisture, not, however, in a sufficient quantity to supply the exigencies of the plume, the two lobes, or main body of the feed, rise along with the plume, affume the appearance of two leaves, resembling the lobes of the feed in size and shape, but having no resemblance to those of the plume, for which reason they have got the name of *digitular leaves*.

These dissimilar leaves defend the young plane from the injuries of the weather, and at the same time, by absorbing dew, air, &c. assist the tender radicle in nourishing the plume, with which they have still a connection by means of the feminal root above described. But, when the radicle or second root has descended deep enough into the earth, and has acquired a sufficient number of filaments or branches for absorbing as much aliment as is proper for the growth of the plume; then the feminal, or digitilar leaves, their utility being entirely superseeded, begin to decay and fall off.

Plate IV. Fig. 1. A. The foramen, or hole in the bean through which the radicle shoots into the foil.

Fig. 2. A transverse section of the bean; the dots being the branches of the feminal root.

Fig. 3. A. The radicle.

B. The plume or bud.

Fig. 4. A longitudinal section of one of the lobes of the bean a little magnified, to shew the small bladders of which the pulpy or parenchymatos part is composed.

Fig. 5, 6. A. A transverse section of the radicle.

B. A transverse section of the plume, showing the organs or vessels of the feminal root.

Fig. 7. A. A view of the feminal root branched out upon the lobes.

Fig. 8. The appearance of the radicle, plume, and feminal root, when a little further advanced in growth.

Having thus briefly described the feed, and traced its evolution into three principal organic parts, viz. the plume, radicle, and feminal leaves, we shall next take an anatomical view of the root, trunk, leaves, &c.

L SECT.

Articles

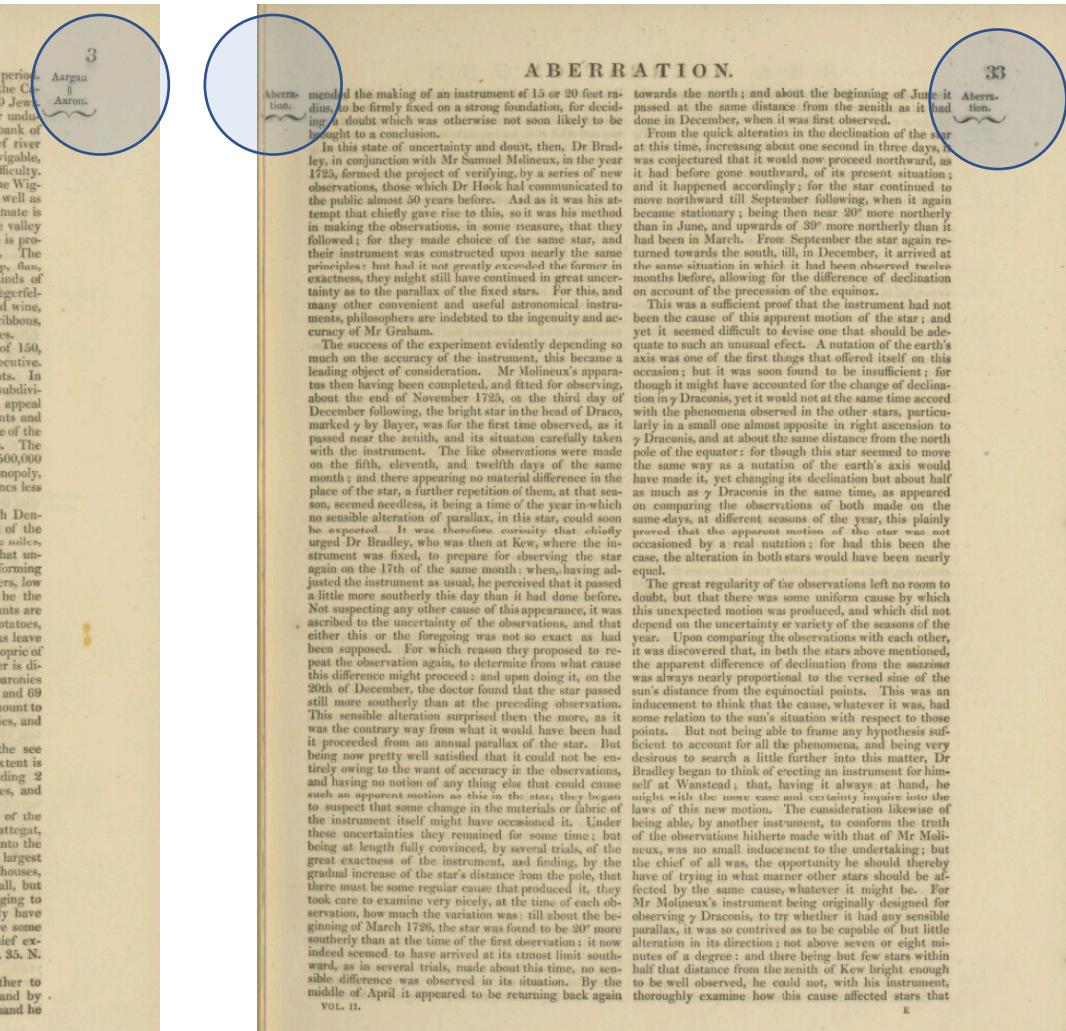
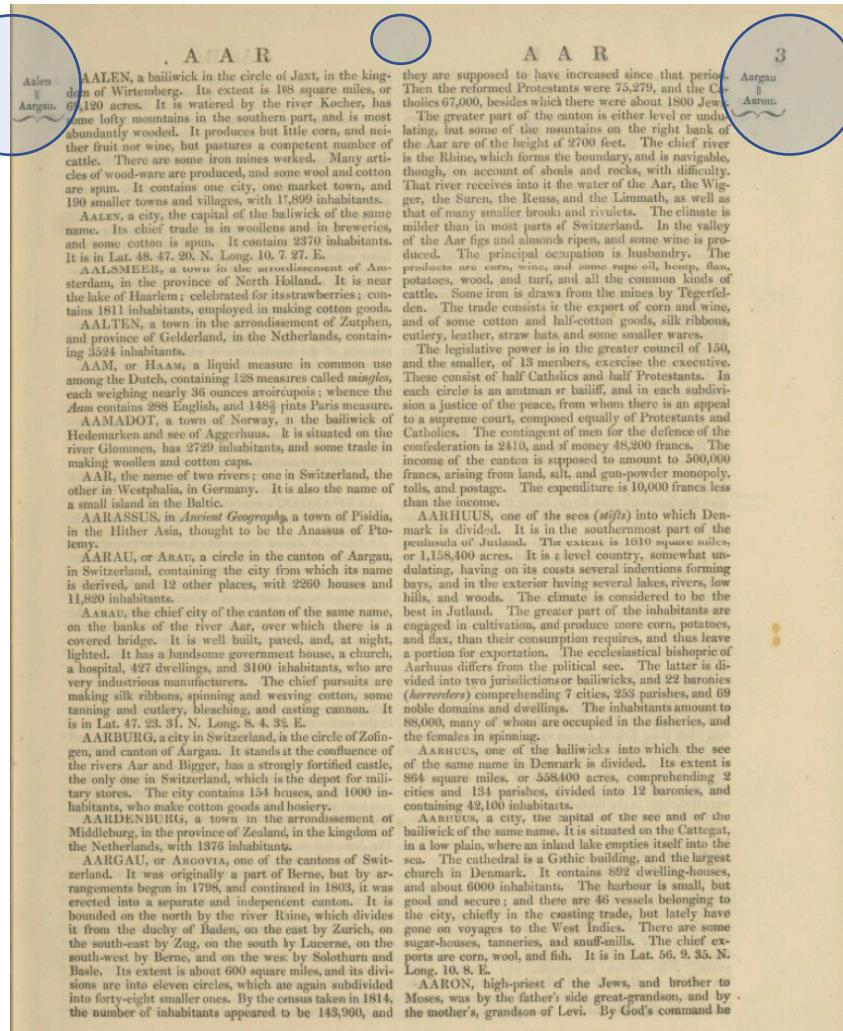
Topic

Edition 1 - 1771

Phase 1: Text Mining

1. Detecting pages headers from ALTO XML
2. Using headers to classify terms into: Articles & Topics
3. Using ALTO Text for detecting the start of each article:

--> Starting a line with TERM UPPERCASE + “,” .

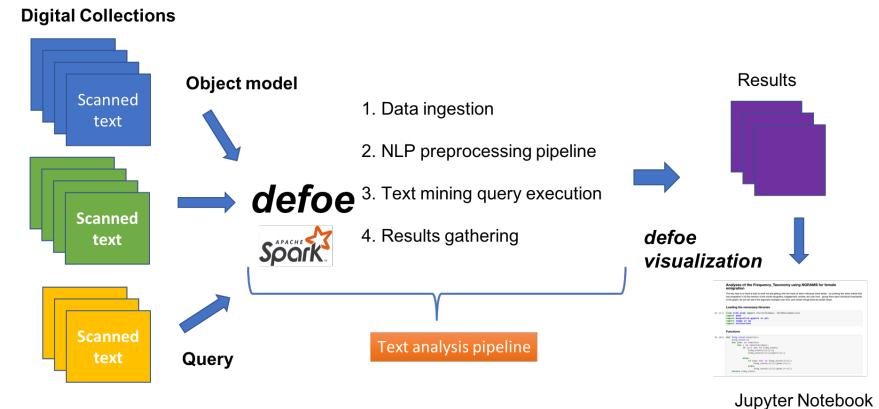


Articles

Edition 7

Topic

Extract EB Terms



term
definition
relatedTerms
header
startsAt
endsAt
numberOfTerms
numberOfWords
numberOfPages
positionPage
typeTerm
editionTitle
editionNum
supplementTitle
supplementsTo
year
place
volumeTitle
volumeNum
letters
part
altoXML
Name: 18, dtype: object

ABACTORES
or ABACTORS, a term for such as carry offer dr...
[]
EBAA
15
15
22
18
832
18
Article
First edition, 1771, Volume 1, A-B
1
[]
1771
Edinburgh
Encyclopaedia Britannica; or, A dictionary of ...
1
A-B
0
144133901/alto/188082904.34.xml

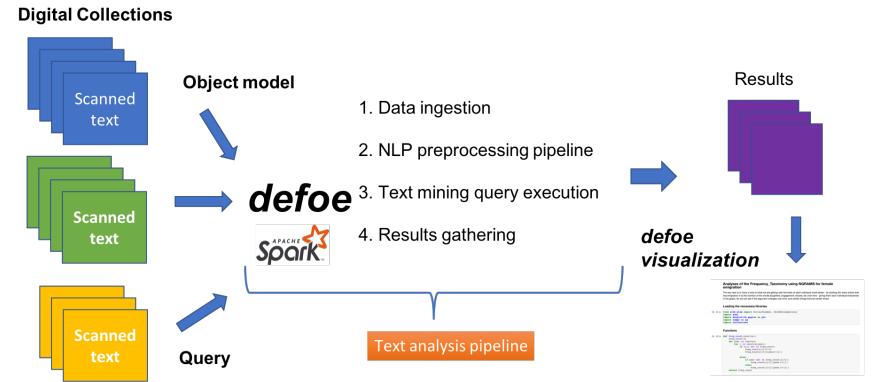
Extracted Term ABACATORES information - Edition 1, 1771, Vol A-B

Phase 1: Text Mining

1.2 Improved defoe **Metadata Extraction query** → It extracts the metadata per Edition and Volume (**Metadata v.1**)

Based on METS information

Extract collection Metadata (METS)



	MMSID	editionTitle	editor	editor_date	genre	language	termsOfAddress	numberOfPages	physicalDescription	place	...	permanentURL	Jupyter Notebook
14	997902543804341	Third edition, Volume 2, ANG-BAR	None	None	encyclopedia	eng	None	922	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149977338	
15	997902543804341	Third edition, Volume 3, BAR-BZO	None	None	encyclopedia	eng	None	856	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149977873	
16	997902543804341	Third edition, Volume 4, CAA-CIC	None	None	encyclopedia	eng	None	842	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149978642	
17	997902543804341	Third edition, Volume 5, CIC-DIA	None	None	encyclopedia	eng	None	858	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149979156	
18	997902543804341	Third edition, Volume 6, DIA-ETH	None	None	encyclopedia	eng	None	850	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149979622	
19	997902543804341	Third edition, Volume 7, ETM-GOA	None	None	encyclopedia	eng	None	882	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149981189	
20	997902543804341	Third edition, Volume 8, GOB-HYD	None	None	encyclopedia	eng	None	832	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149981670	
21	997902543804341	Third edition, Volume 9, Hydrostatics- LES	None	None	encyclopedia	eng	None	872	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149982181	
22	997902543804341	Third edition, Volume 10, LES-MEC	None	None	encyclopedia	eng	None	842	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149982692	
23	997902543804341	Third edition, Volume 11, Medals- Midwifery	None	None	encyclopedia	eng	None	862	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/149983206	
26	997902543804341	Third edition, Volume 1, A-ANG	None	None	encyclopedia	eng	None	894	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/190273291	
27	997902543804341	Third edition, Volume 12, MIE-NEG	None	None	encyclopedia	eng	None	870	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/190273372	
28	997902543804341	Third edition, Volume 13, NEH-PAS	None	None	encyclopedia	eng	None	874	18v.,plates : ill.,maps,music ; 4to	Edinburgh	...	https://digital.nls.uk/191253798	

Metadata of some of the Volumes of Edition 3

Phase 1: Text Mining

1.3. **New Post-processing python scripts (*)** to improve the previous results:

- Re-classification of **Terms v.1** (articles and topics)
- Join terms spitted across pages

We get here: **Terms & Metadata v.2**

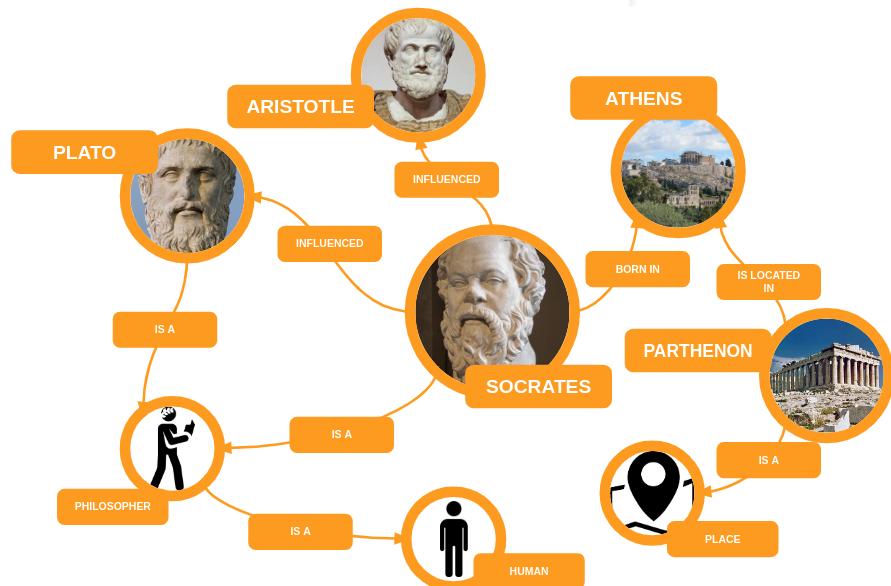
(*) **Also Based on Heuristics** --> Pages Layout & Text & Headers → different heuristics for different EB editions.

Phase 2: Knowledge Graph

Knowledge Graph: Incorporate human knowledge into intelligent systems, exploiting a semantic graph perspective

- A **knowledge graph** is a specialized graph or network of the things we want to describe and how they are related
- It is a **semantic** model since we want to capture and generate **meaning** with the model

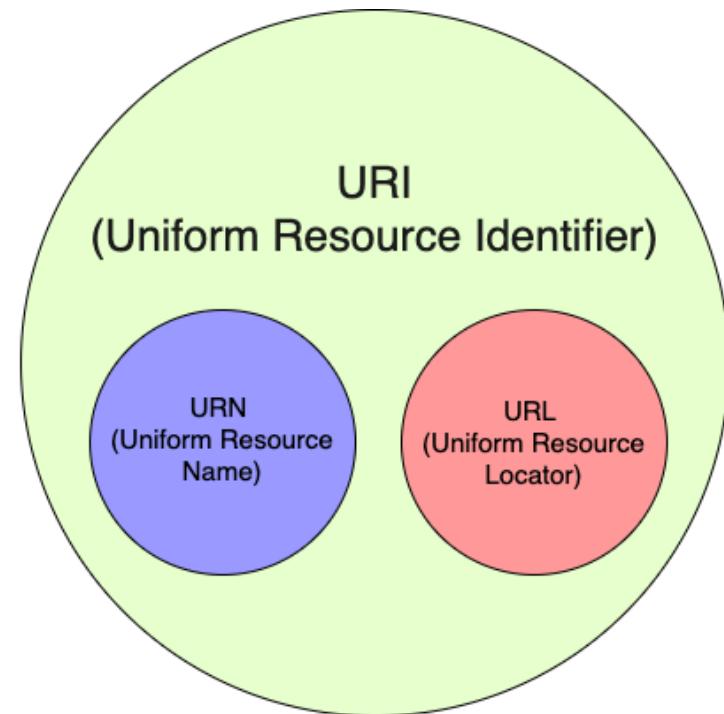
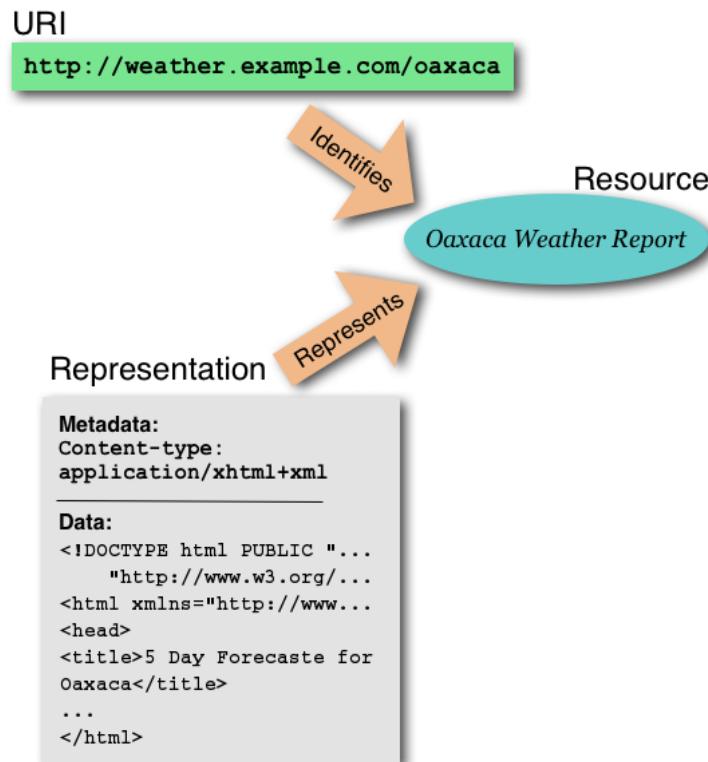
"The application of graph processing and graph DBMSs will grow at 100 percent annually through 2022 to continuously accelerate data preparation and enable more complex and adaptive data science."
– Gartner's *Top 10 Data and Analytics Technology Trends for 2019*



Phase 2: Knowledge Graph

Knowledge Graph: Incorporate human knowledge into intelligent systems, exploiting a semantic graph perspective

URI: A Universal Resource Identifier, is defined to be an ASCII string used to identify “things” on the Knowledge Graph



Phase 2: Knowledge Graph

Knowledge Graph: Incorporate human knowledge into intelligent systems, exploiting a semantic graph perspective

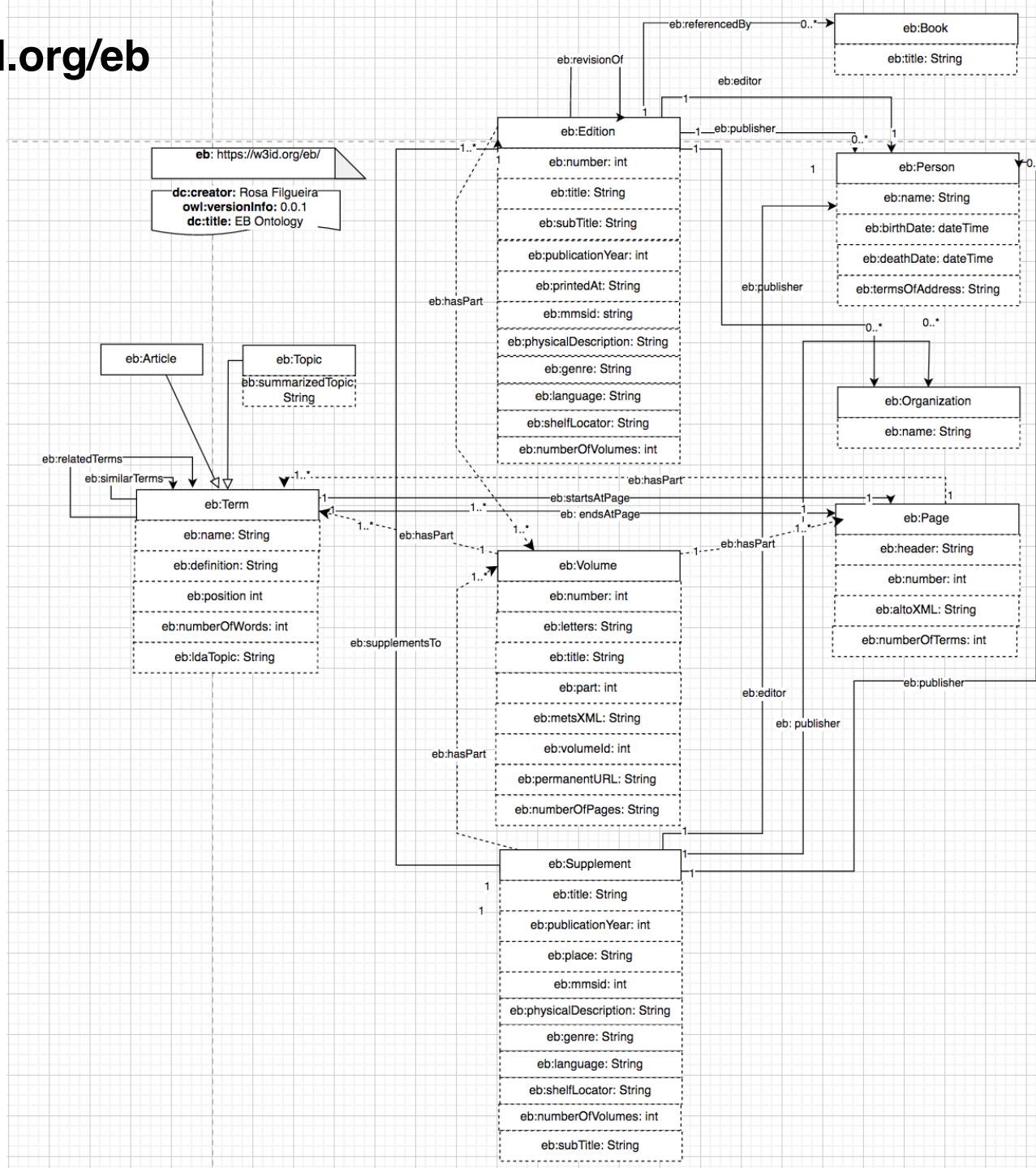
2.1. Create EB Ontology: To explain (give meaning) how our “things” are related with each other.

In order to create and publish the EB-Ontology I used:

- [diagrams.net](#) : To create an UML with the EB information (classes, properties, relationships, etc.)
- [Chowlk](#) : To convert the UML into an OWL ontology
- [Widoco](#): To publish and create an enriched and customized documentation of the ontology
- [w3id.org](#): To configure my permanent Identifier for EB ontology →
<https://w3id.org/eb/>

EB-Ontology : <https://github.com/francesNLP/EB-ontology>

<https://w3id.org/eb>



Phase 2: Knowledge Graph

Knowledge Graph: Incorporate human knowledge into intelligent systems, exploiting a semantic graph perspective

2.1. Created EB Ontology: The description of this ontology is available online at http: <https://w3id.org/eb/>

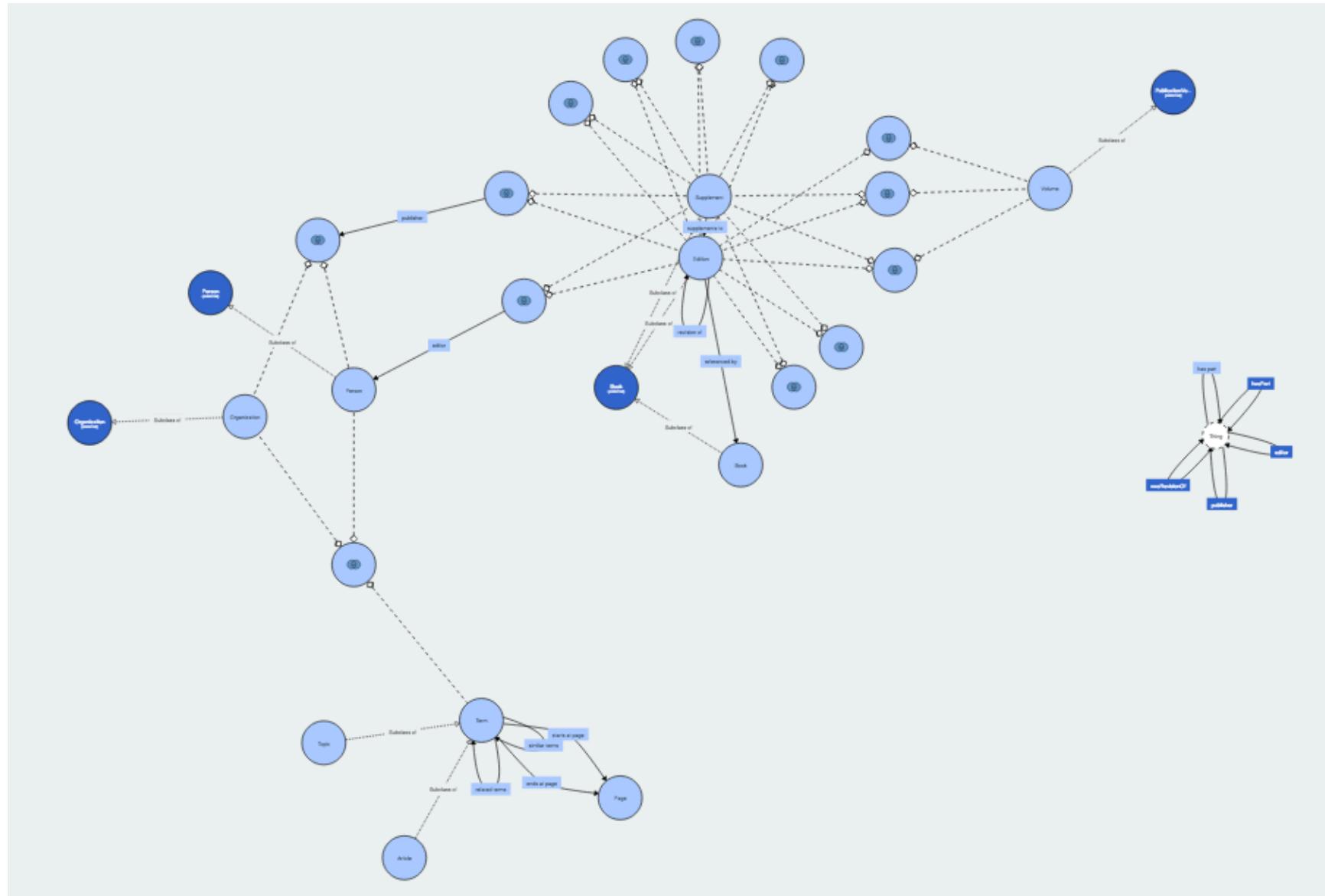
2.2 . Created **EB Knowledge Graph 1.0**: Populated the post-processed information (extracted **Terms & Metadata v.2**) into a RDF triplestore using the EB Ontology

- My RDF EB-data is stored in an Apache Jena FUSEKI SPARQL server – Used this Fuseki-docker image to set up my SPARQL server/end-point

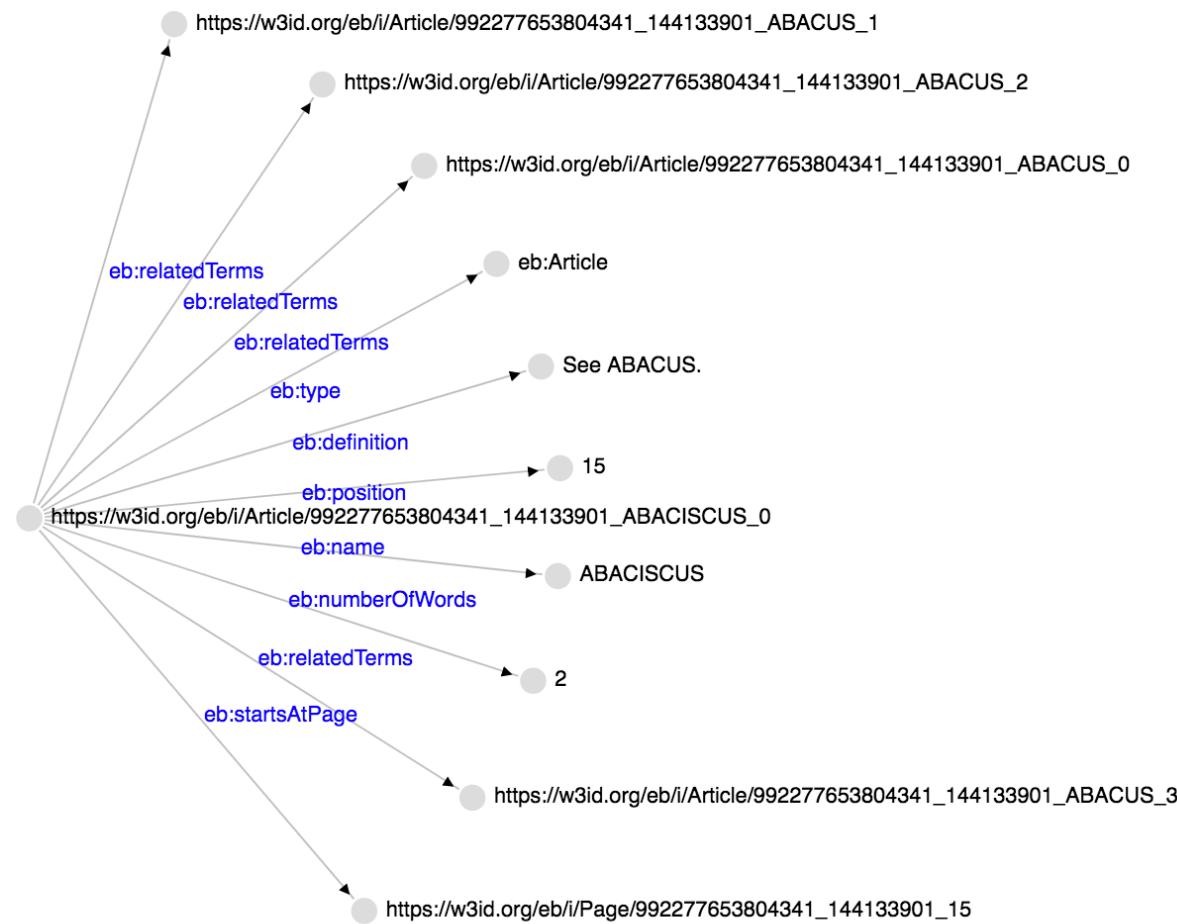
NOTE:

Each Term, Edition, Page, Volume, etc ... is a **Resource in our Knowledge Graph** and has an **URI to identify it**.

EB Knowledge Graph

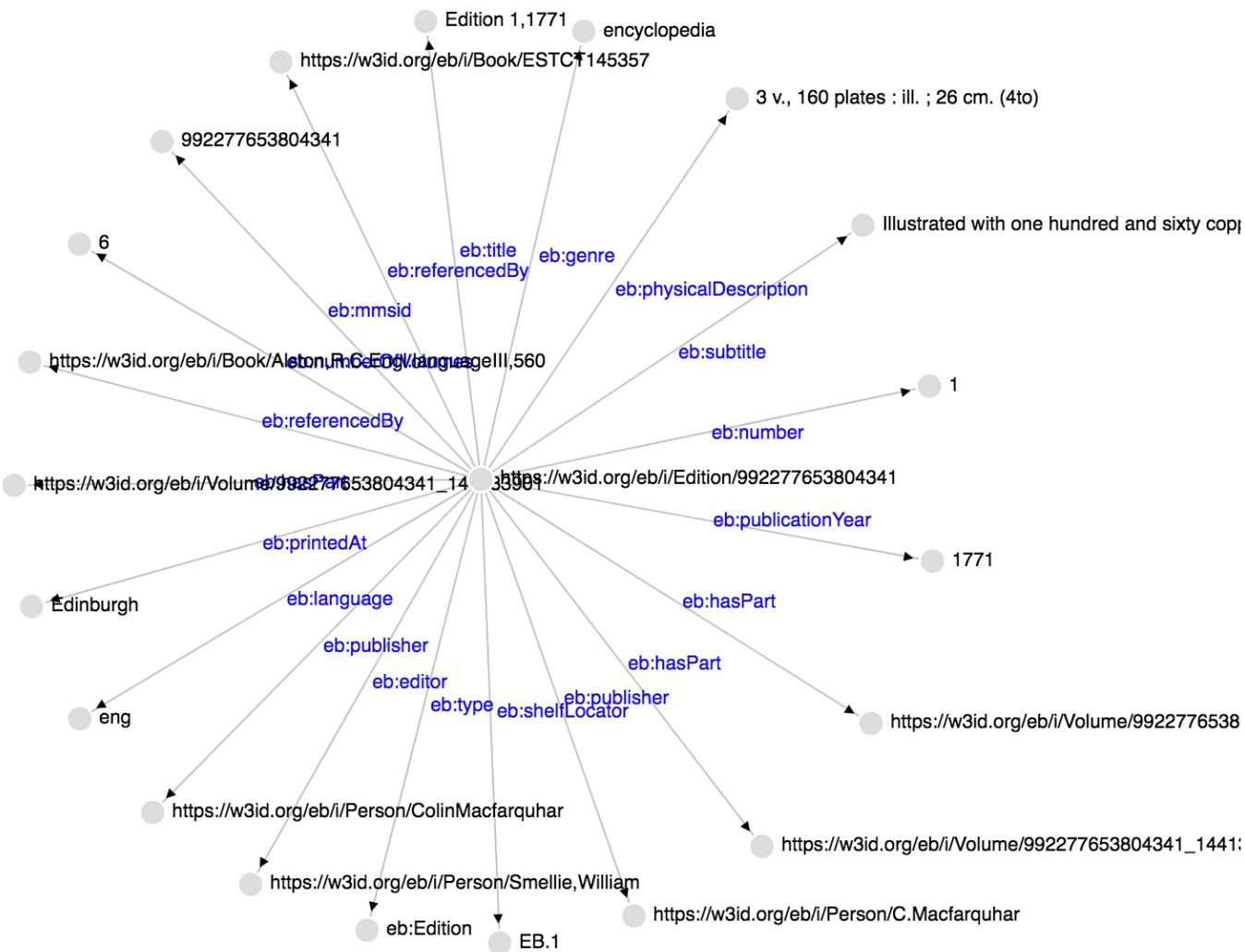


EB Knowledge Graph



Term ABACISCUS, URI: https://w3id.org/eb/i/Article/992277653804341_144133901_ABACISCUS_0

EB Knowledge Graph



Edition 1, 1771 – URI: <https://w3id.org/eb/i/Edition/992277653804341>

EB Knowledge Graph – Querying our KG

SPARQL is an **RDF query language**—that is, a semantic query language for databases—able to retrieve and manipulate data stored in Resource Description Framework (RDF) format.

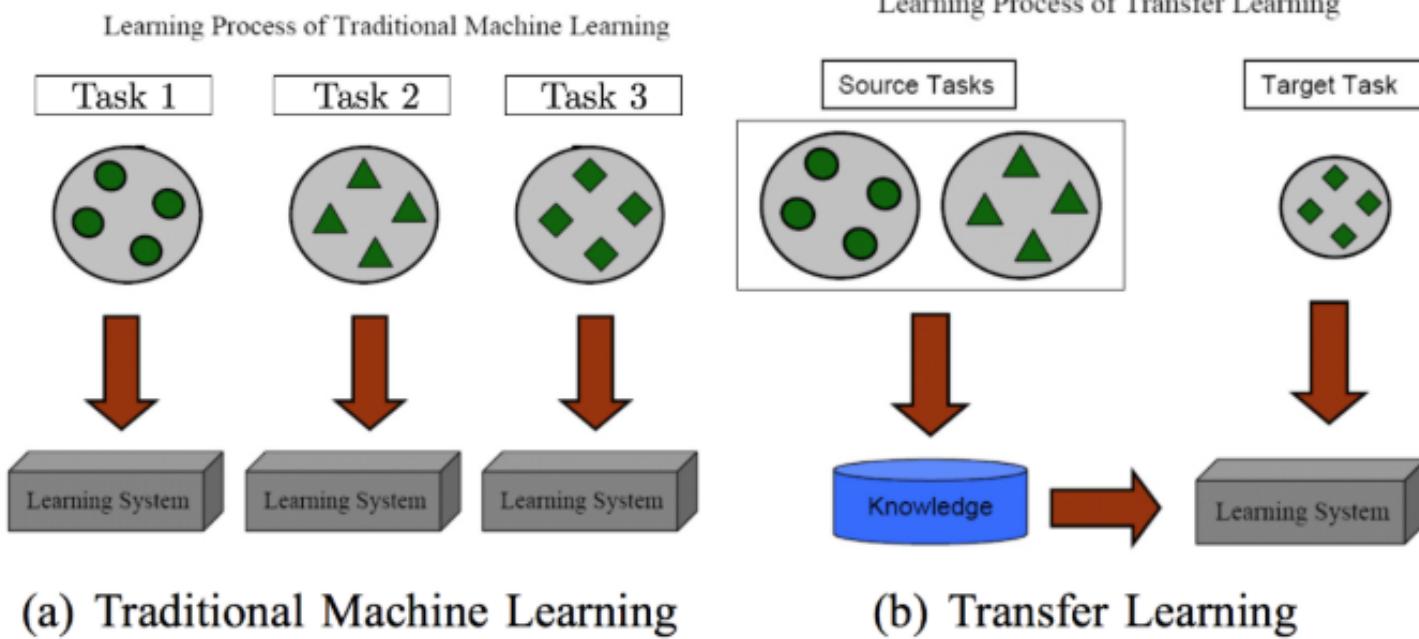
```
2 SELECT ?article ?definition
3 WHERE {
4     ?article a eb:Article .
5     ?article eb:definition ?definition
6     FILTER (CONTAINS(?definition, "Scotland"))
7     FILTER (CONTAINS(?definition, "Glasgow"))
8     OPTIONAL{FILTER CONTAINS(?definition, "Edinburgh") }
9 } LIMIT 10
10
```



Fuseki RDF triplestore

QUERY RESULTS	
Table	Raw Response
Download	
Showing 1 to 8 of 8 entries	
	<input type="text"/> Search: <input type="button"/> Show 50 entries
article	definition
1 https://w3id.org/eb/l/Article/9929192893804340_144850368_PAISLEY_0	"a town of Scotland, in the county of Renfrew, six miles west of Glasgow."
2 https://w3id.org/eb/l/Article/9929192893804340_144850366_ARGYLESHIRE_0	"a county of Scotland, lying westward of Glasgow, and comprehending the countries of Lorn, Cowal, Knapdale, Kintyre, together with the islands Mull, Jura, Iona, & Canna. It gives the title of duke to the noble family of Campbell."
3 https://w3id.org/eb/l/Article/9929192893804340_144850367_INVERARY_0	"a parliament town of Scotland, in the county of Argyle, of which it is the capital, situated in Lochay, forty five miles north-west of Glasgow : W. long. 5° 0', N. lat 36° 28'."
4 https://w3id.org/eb/l/Article/9929192893804340_144850367_HAMILTON_0	"a town of Scotland, in the county of Clydesdale, situated on the river Clyde, eleven miles south-east of Glasgow : W. long. 3° 0', N. lat. 55° 0"

Phase 3: Augmented Knowledge Graph with Deep Transfer Learning



Phase 3: Augmented Knowledge Graph with Deep Transfer Learning



Transformers

build passing license [Apache-2.0](#) website [online](#) release [v2.0.0](#)

State-of-the-art Natural Language Processing for PyTorch and TensorFlow 2.0

😊 Transformers provides thousands of pretrained models to perform tasks on texts such as classification, information extraction, question answering, summarization, translation, text generation, etc in 100+ languages. Its aim is to make cutting-edge NLP easier to use for everyone.

😊 Transformers provides APIs to quickly download and use those pretrained models on a given text, fine-tune them on your own datasets then share them with the community on our [model hub](#). At the same time, each python module defining an architecture can be used as a standalone and modified to enable quick research experiments.

😊 Transformers is backed by the two most popular deep learning libraries, [PyTorch](#) and [TensorFlow](#), with a seamless integration between them, allowing you to train your models with one then load it for inference with the other.

Phase 3: Augmented Knowledge Graph with Deep Transfer Learning



build passing license Apache-2.0 website online release v2.0.0

A High-Level Look

Let's begin by looking at the model as a single black box. In a machine translation application, it would take a sentence in one language, and output its translation in another.



Phase 3: Augmented Knowledge Graph with Deep Transferring Learning

EB Knowledge Graph 2.0: previous info + storing the result of applying different deep learning transformers analyses:

- **sentiment analyses:** Classifying text between positive and negative
 - transformer: [*siebert/sentiment-roberta-large-english*](#)
- **topic modelling:** Clustering terms into topics
 - transformer: [*all-mpnet-base-v2*](#)
- **term similarity:** Comparing text & semantic similarity
 - transformer: [*all-mpnet-base-v2*](#)
- **spelling checking:** Finding misspelling/ocr errors and fixing them
 - transformer: [*neuspell*](#) + *ElmoslstmChecker*
- **term evolution:** Checking how a term has changed over the years
 - transformer: [*all-mpnet-base-v2*](#)
- **summarization:** Summarizing the text of a topic term (XLNET)
 - transformer: [*XLNet*](#)

Phase 3: Augmented Knowledge Graph with Deep Transferring Learning

EB Knowledge Graph 2.0: previous info + storing the result of applying different deep learning transformers analyses:

- **sentiment analyses:** Classifying text between positive and negative
- **topic modelling:** Clustering terms into topics
- **term similarity:** Comparing text & semantic similarity
- **spelling checking:** Finding misspelling/ocr errors and fixing them
- **term evolution:** Checking how a term has changed over the years
- **summarization:** Summarizing the text of a topic term

Example: Spelling Checking → *Lewis* Term

Original Definition

the mort northerly of any of the w eftern islands of Scotland, lying in 8\u00b0 odd minutes W. long, and between 58\u00b0 and 59 0 odd minutes N. lat.



Cleaned Definition

the most northerly of any of the w eastern islands of Scotland , lying in 8 and odd minutes W. long , and between 58 and and 59 0 odd minutes N. land .



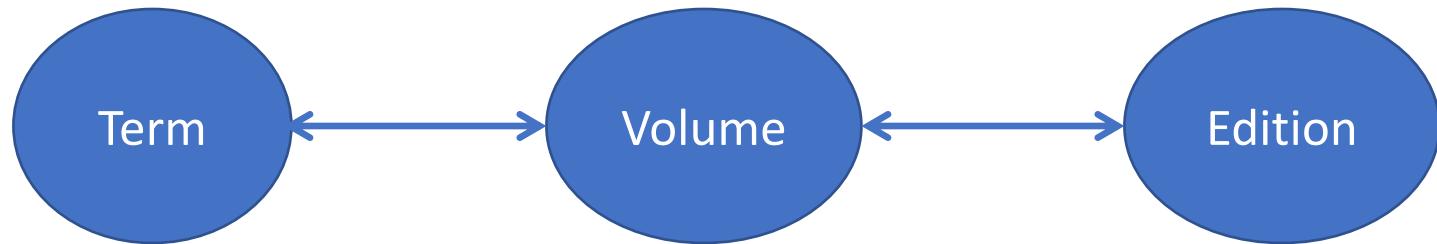
Compute Difference

the morst northerly of any of the w eftern islands of Scotland, lying in 8\u00b0 odd minutes W. long, and between 58\u00b0 and 59 0 odd minutes N. latnd .

Phase 3: Augmented Knowledge Graph with Deep Transferring Learning

Two types of “queries” against the EB Knowledge Graph 2.0 :

- **Type 1: Extracting information from the EB Knowledge → SPARQL → We “just” navigate through the KG to get the desired information.**

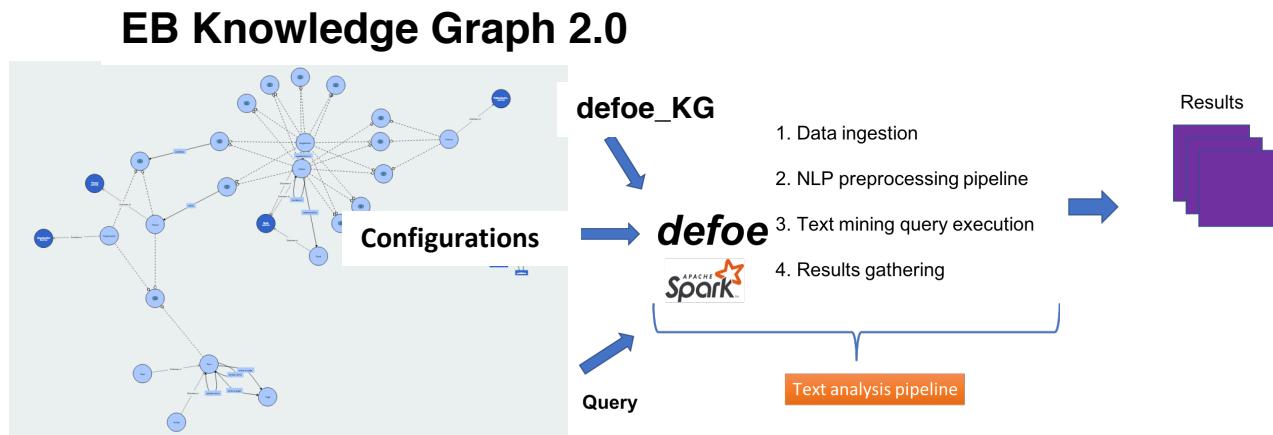


Example: Given a **Term** (e.g., *Edinburgh*), I can get the **Edition** Information (e.g., *Edition Title*)

- **Type 2: Processing information from the EB Knowledge → defoe → We are going to process further the definitions from the selected terms.**
 - But for doing this we needed to do some work on defoe first → Phase 4

Phase 4: Defoe and Knowledge Graph

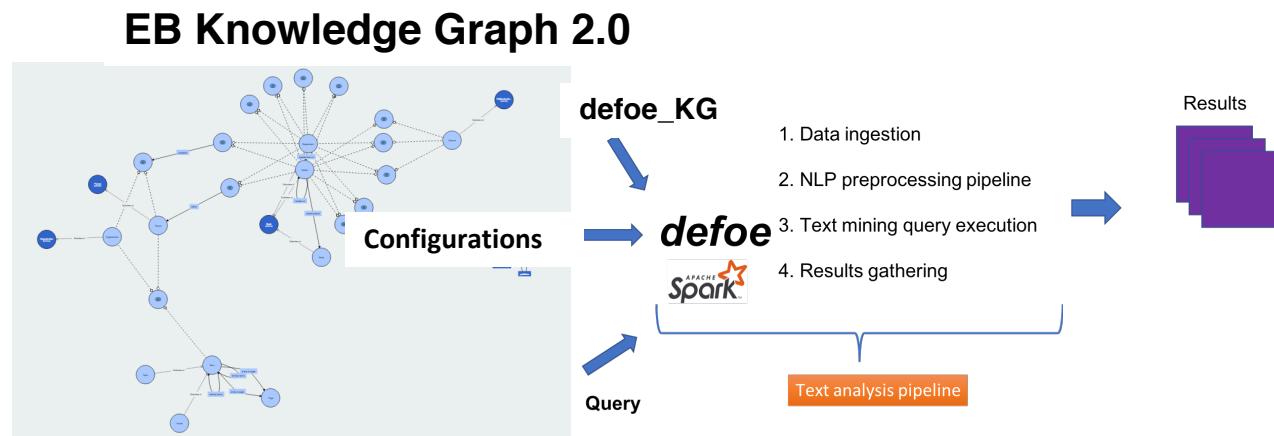
- 4.1. Created a new **KG defoe connector** (based in SPARQL) to run defoe queries using the EB Knowledge Graph as a source of data → **defoe_KG**
- 4.2. Improved **defoe queries to be fully configurable**: different filtering options, target, lexicon, etc.



Phase 4: Defoe and Knowledge Graph

4.3 defoe text mining queries:

- frequency keysearch: **Count number of terms** or times in which appear keywords or keysentences and group by years. Several filtering options.
- term fulltext keysearch: **Extract terms definitions** in which appear keywords or keysentences and group by years. Several filtering options.
- term snippet keysearch: **Extract snippets of definitions** in which appear keywords or keysentences and group by years. Several filtering options, including the snippet size.
- publication normalization: **Extract the number of documents, pages, words** per year.
- uris keysearch: **Extract uris of terms** in which appear keywords or keysentences and group by years. Several filtering options.
- geoparser terms: **Geoparsing the term definition** in which appear keywords or keysentences and group by years. Several filtering options.



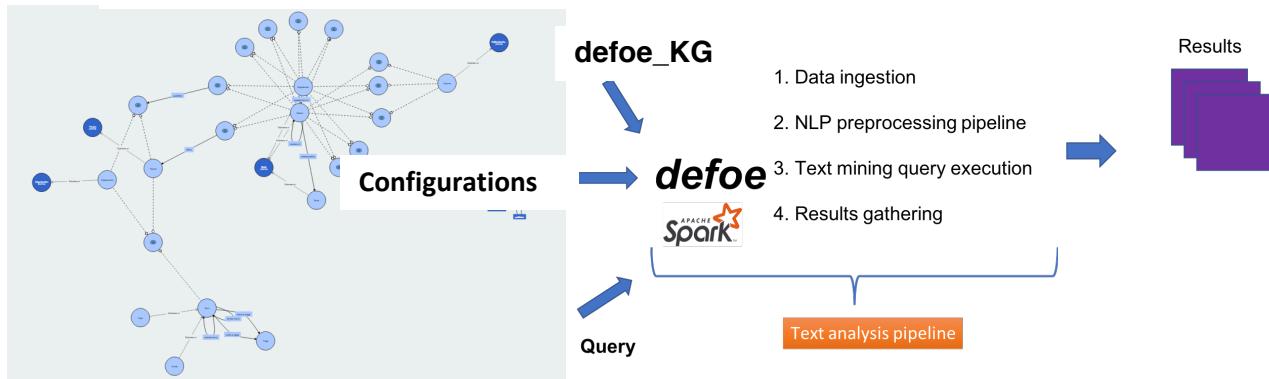
Phase 4: Defoe and Knowledge Graph

4.4 Created **new defoe ML/NLP queries** (using also deep learning transformers)

- sentiment analysis: Calculate the **sentiment** analyses of selected terms.
- topic modelling: Calculate the **topic** of selected terms. Filtering options.
- spelling checker: Check the **spelling** of selected terms. Filtering options.

(TO-DO)

EB Knowledge Graph 2.0



Phase 5: Flask Web-Application

5.1 Web-User Interface:

- So, users DO NOT have to create SPARQL or defoe queries – the web-application does it for them → Abstractions to SPARQL & DEFOE
- The web-app runs both types of queries and visualizes the results
- Functionalities:
 - Term Search
 - Term Similarity
 - Topic Modelling
 - Spelling Checker
 - Defoe queries
 - EB Details
 - Knowledge Graph Resources Visualizations

5.2 Flask + JQuery + JavaScript + HTML + CSS + Web-services + Plotly

Summary

1. EB ALTO & METS + **defoe** → Extracted **Terms & Metadata (v.1)**
2. Extracted terms and metadata (v.1) + **postprocessing scripts**:
 1. Final set of **Terms & Metadata (v.2)**
 2. **EB ontology**
3. Terms & Metadata (v.2) + EB ontology → **EB Knowledge Graph 1.0**
4. EB Knowledge Graph 1.0 + Transformers → **EB Knowledge Graph 2.0**
5. EB Knowledge Graph 2.0 + defoe → **defoe_KG**
6. defoe_KG + **configurable defoe queries**
7. EB Knowledge Graph 2.0 + defoe_KG + defoe queries (text mining/ NLP)
8. EB Knowledge Graph Queries:
 1. **Type 1: SPARQL queries:** Extracting the EB KG data
 2. **Type 2: defoe queries:** Processing the EB KG data
9. Flask-web app: **Abstractions to SPARQL** and **defoe queries**

Interacting with the EB Knowledge Graph

Term Search Term Similarity Topic Modelling Spell Checker Term Evolution Defoe Queries EB Details Visualization of Resources

Exploring the Encyclopaedia Britannica (1768-1860)

Term Search

Enter the **term** that you would like to search for. In case that the **Term Type is a Topic**, only the **summary** of the definition is displayed. If no term is introduced, it will search for the first term in the Encyclopaedia.

Results for **EDINBURGH**.

Note that if you click over an **URI** in this table, it will take you to the **Visualization of Resources page**. However, if you instead click over a **related term**, it will conduct a **term search**, showing all the searching results for that term. And if you click over a **topic model** if will take you to the **Topic Modelling page**, listing all the terms belonging to that particular topic model.

displaying 1 - 2 records in total 2

URI	Year	Edition	Volume	Start Page	End Page	Term Type	Definition/Summary	Related Terms	Topic Modelling	Sentiment_Score	Advanced Options
https://w3id.org/eb/i/Article/992277653804341_144133902_EDINBURGH_0	1771	1	2	408	409	Article	the capital city of the kingdom of Scotland, situated W. long. 3° 0', and N. lat. 56° 0' 0". We ffiall... More		15_scotland_county_edinburgh_firth	LABEL_0_0.87	Spell Checker Term Similarity Term Evolution
https://w3id.org/eb/i/Article/9929192893804340_144850367_EDINBURGH_0	1773	1	2	414	415	Article	the capital city of the kingdom of Scotland, situated W. long. 3° 0', and N. lat. 56° 0' 0". We (hall ... More		15_scotland_county_edinburgh_firth	LABEL_0_0.89	Spell Checker Term Similarity Term Evolution

FLASK Web-Application

Interacting with the EB Knowledge Graph

Type 1

Term Search

Term Similarity

Topic Modelling

Spell Checker

Term Evolution

Type 2

Defoe Queries

Type 1

EB Details

Visualization of Resources

Exploring the Encyclopaedia Britannica (1768-1860)

Term Search

Enter the **term** that you would like to search for. In case that the **Term Type is a Topic**, only the **summary** of the definition is displayed. If no term is introduced, it will search for the first term in the Encyclopaedia.

Results for **EDINBURGH**.

Note that if you click over an **URI** in this table, it will take you to the **Visualization of Resources page**. However, if you instead click over a **related term**, it will conduct a **term search**, showing all the searching results for that term. And if you click over a **topic model** it will take you to the **Topic Modelling page**, listing all the terms belonging to that particular topic model.

displaying 1 - 2 records in total 2

URI	Year	Edition	Volume	Start Page	End Page	Term Type	Definition/Summary	Related Terms	Topic Modelling	Sentiment_Score	Advanced Options
https://w3id.org/eb/i/Article/992277653804341_144133902_EDINBURGH_0	1771	1	2	408	409	Article	the capital city of the kingdom of Scotland, situated W. long. 3° 0', and N. lat. 56° 0' 0". We ffiall... More		15_scotland_county_edinburgh_firth	LABEL_0_0.87	Spell Checker Term Similarity Term Evolution
https://w3id.org/eb/i/Article/9929192893804340_144850367_EDINBURGH_0	1773	1	2	414	415	Article	the capital city of the kingdom of Scotland, situated W. long. 3° 0', and N. lat. 56° 0' 0". We (hall ... More		15_scotland_county_edinburgh_firth	LABEL_0_0.89	Spell Checker Term Similarity Term Evolution

Note: Two types of EB-KG queries

- **Type 1:** Extracting information from EB-KG – Navigating across the KG
- **Type 2:** Processing information from EB-KG -- Processing in parallel further the Term's definitions (stored in KG).

Interacting with the EB Knowledge Graph

Term Similarity

Enter an URI of a term or some text that you would like to search similar terms for. If no term is introduced, it will search for similar terms of the first term in the Encyclopaedia.

Details of this resource

URI	Edition	Year	Volume	Term	Original Definition	Topic Modelling	Sentiment_Score
https://w3id.org/eb/i/Article/992277653804341_144133902_EDINBURGH_0	1	1771	2	EDINBURGH	the capital city of the kingdom of Scotland, situated W. long. 3°, and N. lat. 56°00'00". We ffiall... More	15_scotland_county_edinburgh_firth	LABEL_0_0.87

Note that if you click over an **URI** in this table, it will search for **similar terms** of that resource in this page. However, if you instead click over a **term**, it will take you to the **Term Search page**, showing all the searching results for that term. And if you click over a **topic model** it will take you to the **Topic Modelling page**, listing all the terms belonging to that particular topic model.

displaying 1 - 10 records in total 20

« 1 2 »

URI	Edition	Year	Volume	Term	Definition	Topic Modelling	Similarity Rank	Sentiment_Score
https://w3id.org/eb/i/Article/9929192893804340_144850367_EDINBURGH_0	1	1773	2	EDINBURGH	the capital city of the kingdom of Scotland, situated W. long. 3°, and N. lat. 56°00'00". We (hall ... More	15_scotland_county_edinburgh_firth	0.99541134	LABEL_0_0.89
https://w3id.org/eb/i/Article/992277653804341_144133901_ABERDEEN_0	1	1771	1	ABERDEEN	the name of two cities in Scotland, called the Old and New Towns, situated on the German Ocean, in i... More	15_scotland_county_edinburgh_firth	0.7731124	LABEL_0_0.92
https://w3id.org/eb/i/Article/9929192893804340_144850367_DUMFERMLINE_0	1	1773	2	DUMFERMLINE	a parliament-town of Scotland, situated in the county of Fife	15_scotland_county_edinburgh_firth	0.7610539	POSITIVE_1.00

Interacting with the EB Knowledge Graph

Topic Modelling

Exploring the Encyclopaedia Britannica (1768-1860)

Topic Modelling Search

Enter a **topic modelling name** or just the **number** of a topic to see all the terms within the same topic. All topics modelling names start with a number. If no topic is introduced, it will use the first topic modelling, which name starts with '0'.

We have 127 topics modelling items.

182 terms found for the topic 15_scotland_county_edinburgh_frith

Note that if you click over an **URI** in this table, it will take you to the [Visualization of Resources page](#). However, if you instead click over a **term**, it will take you to the [Term Search page](#), showing all the searching results for that term.

displaying 1 - 10 records in total 182

« 1 2 3 4 5 ... 18 19 »

URI	Edition	Year	Volume	Term	Definition	Sentiment_Score
https://w3id.org/eb/i/Article/9929192893804340_144850368_PAISLEY_0	1	1773	3	PAISLEY	a town of Scotland, in the county of Renfrew, six miles west of Glasgow.	POSITIVE_0.99
https://w3id.org/eb/i/Article/9929192893804340_144850366_ARGYLESHERE_0	1	1773	1	ARGYLESHERE	a county of Scotland, lying westward of Glasgow, and comprehending the counties of Lorn, Cowal, Kna... More	POSITIVE_1.00
https://w3id.org/eb/i/Article/9929192893804340_144850367_LEWIS_0	1	1773	2	LEWIS	the most northerly of any of the western islands of Scotland, lying in 8°W00'00 odd minutes W. long.... More	POSITIVE_0.99

Topic Modelling Visualizations

We are visualizing the topic model 15_scotland_county_edinburgh_frith. For each topic model, we plot their most common 10 words , along with their c-TF-IDF scores .

Topic Word Scores



Interacting with the EB Knowledge Graph

Term Similarity

Spell Checker

Enter an **URI of a term** that you would like to check its spelling. If no term is introduced, it will check the spelling of the first term in the Encyclopaedia.

URI	Edition	Year	Volume	Term
https://w3id.org/eb/i/Article/9929192893804340_144850368_WIDOW_0	1	1773	3	WIDOW

Original Definition

a woman who has lost her hufbind.

Cleaned Definition

a woman who has lost her husband .



Original Definition

a woman who has lost her hufbind.

Cleaned Definition

a woman who has lost her husband.

Interacting with the EB Knowledge Graph

Term Evolution

Enter an **URI of a term** that you would like to see its evolution.

Details of this resource

URI	Edition	Year	Volume	Term	Original Definition	Topic Modelling	Sentiment_Score
https://w3id.org/eb/i/Article/992277653804341_144133902_EDINBURGH_0	1	1771	2	EDINBURGH	the capital city of the kingdom of Scotland, situated W. long. 3°, and N. lat. 56°. We ffiall... More	15_scotland_county_edinburgh_frith	LABEL_0_0.87

The evolution of EDINBURGH can be seen in this table.

URI	Edition	Year	Volume	Term	Definition	Topic Modelling	Similitud Rank	Sentiment_Score
https://w3id.org/eb/i/Article/9929192893804340_144850367_EDINBURGH_0	1	1773	2	EDINBURGH	the capital city of the kingdom of Scotland, situated W. long. 3°, and N. lat. 56°. We ffiall... More	15_scotland_county_edinburgh_frith	0.99541134	LABEL_0_0.89

Select Edition

Select Volume

Edition 1 Year 1771

1 A-B



Returned details for Edition 1 Year 1771 and Volume 1 A-B

Edition	Details
Year	1771
Edition Number	1
Edition URI	< https://w3id.org/eb/i/Edition/992277653804341 >
Edition Title	Edition 1,1771
Edition Subtitle	Illustrated with one hundred and sixty copperplates
Printed at	Edinburgh
Physical Description	3 v., 160 plates : ill. ; 26 cm. (4to)
MMSID	992277653804341
Shelf Locator	EB.1
Genre	encyclopedia
Language	English
Number of Volumes	3

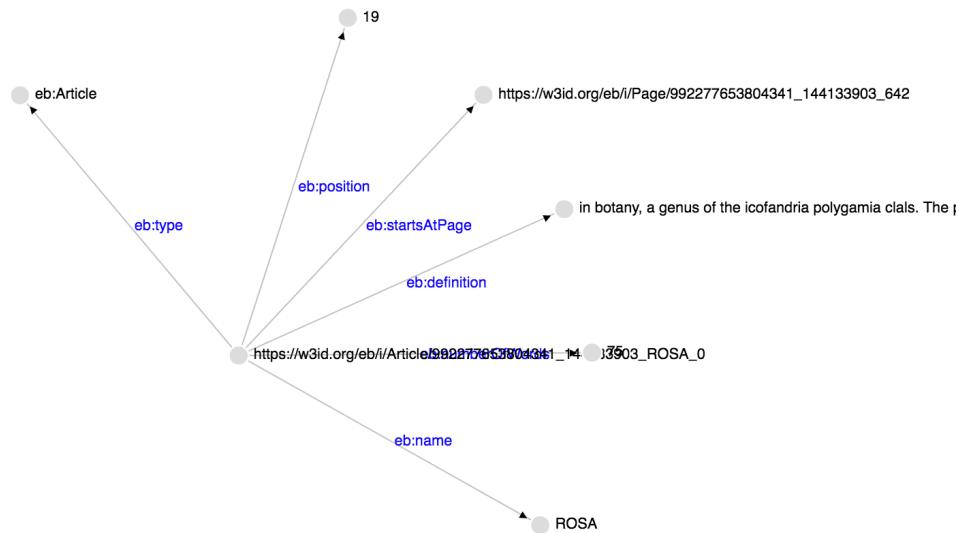
Volume	Details
Volume Number	1
Volume URI	< https://w3id.org/eb/i/Volume/992277653804341_144133901 >
Volume Title	Encyclopaedia Britannica; or, A dictionary of arts and sciences, compiled upon a new plan
Volume Letters	A-B
Volume Permanent URL	https://digital.nls.uk/144133901
Volume Number of Pages	832

Volume	Statistics
Number of Articles	2505
Number of Topics	15
Number of Distinct Articles	2440
Number of Distinct Topics	15

Interacting with the EB Knowledge Graph

Visualization of Resources

Visualization RDF graph for the resource <https://w3id.org/eb/i/Article/992277653804341_144133903_ROSA_0>



Interacting with the EB Knowledge Graph

Defoe Queries

Select the **defoe query** that you would to run, and introduce your configurations:

Defoe query frequency_keysearch_by_year ▾

Preprocess Treatment

- None
- Normalize
- Normalize & Numbers
- Normalize & Lemmatize
- Normalize & Stemming

Select the hit count (only for Freq. query) term ▾

Upload your lexicon file

peace_war.txt

Filtering Options

Introduce Target words/sentences separated by "," (Optional) peace accord,peace agr

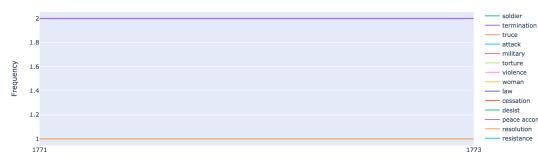
Select the Target filter

- Select terms which contains any of the target words/sentences
- Select terms which contains all the target words/sentences

Start Year (Optional) 1771

End Year (Optional) 1773

Frequency of Lexicon Terms per Years



Normalized Frequency of Lexicon Terms per Years



Results of frequency_keysearch_by_year query

Here we can see the frequency of your lexicon words/sentences taking into account the target words/sentences (if any) and other configurations, sorted by years.

Year	Keyword : Frequency
1771	• accord : 1 • soldier : 1 • termination : 1 • truce : 2 • attack : 1 • military : 1 • torture : 1 • violence : 1 • war : 1 • cessation : 1 • desert : 1 • peace accord : 2 • resolution : 1
1773	• peace accord : 2 • resolution : 1 • law : 2 • cessation : 1 • attack : 1 • military : 1 • termination : 1 • torture : 1 • violence : 1 • war : 1 • accord : 1 • agreement : 1 • soldier : 1 • termination : 1

Download Results

frances: Architecture

NSL Data Foundry Collections

