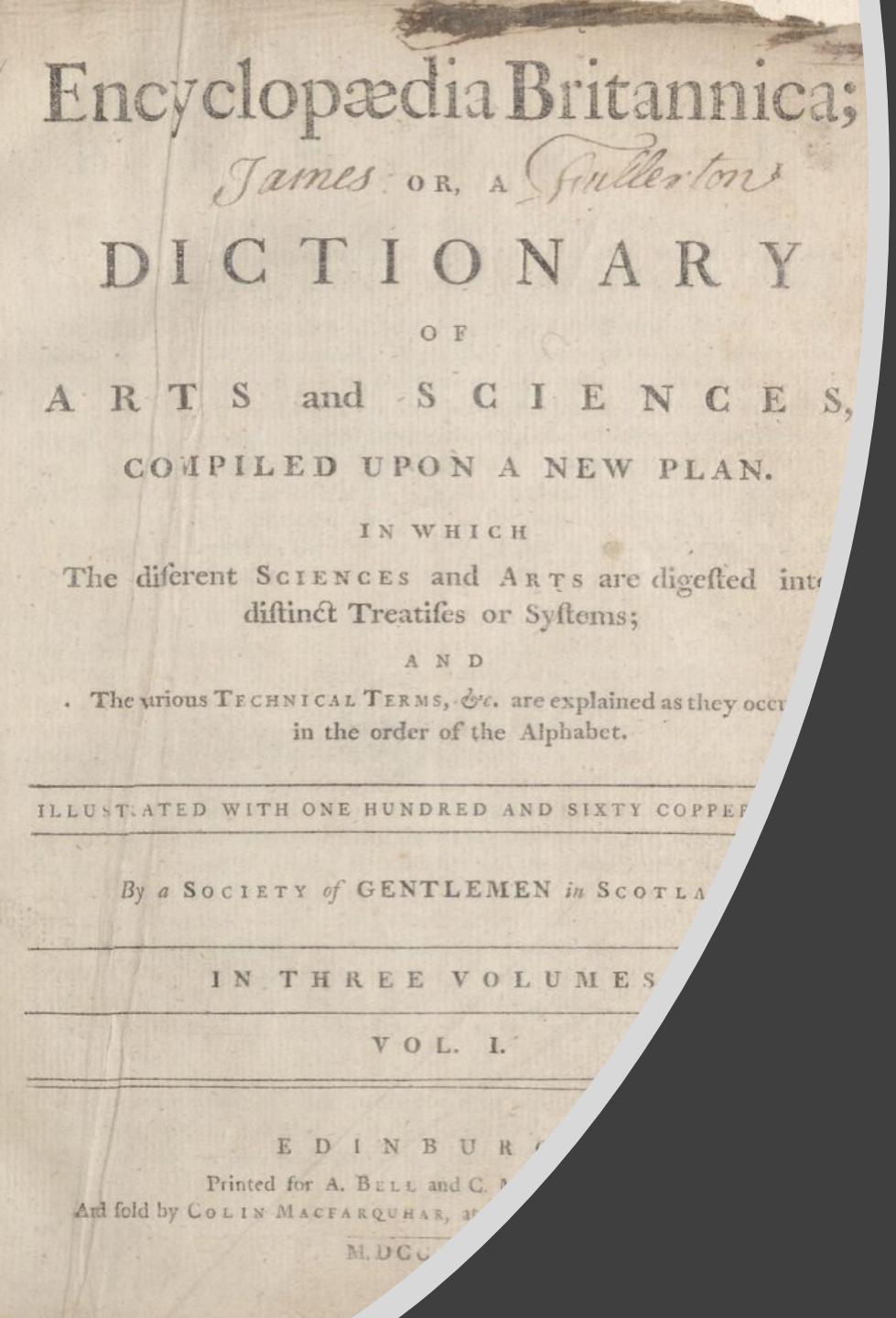

Encyclopædia Britannica;

OR, A James Fullerton's

DICTIONARY

OF

ARTS and SCIENCES,

COMPILED UPON A NEW PLAN.

IN WHICH

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

AND

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.

EDINBUR

Printed for A. BELL and C.

And sold by COLIN MACFARQUHAR, a

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 : N*ew 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**

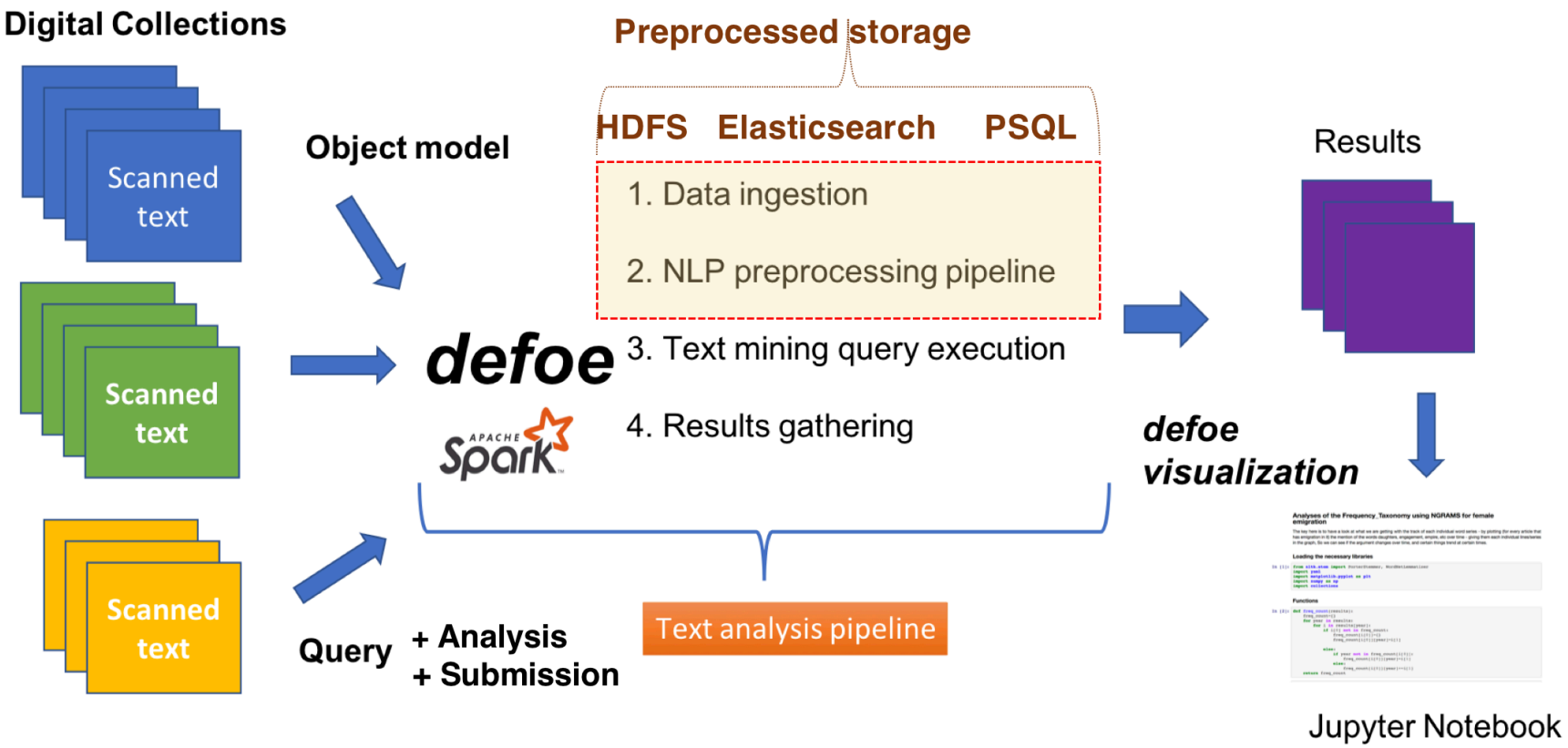
(2) Querying Augment EB Knowledge Graph:

(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



scalable portable text mining queries large dataset

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

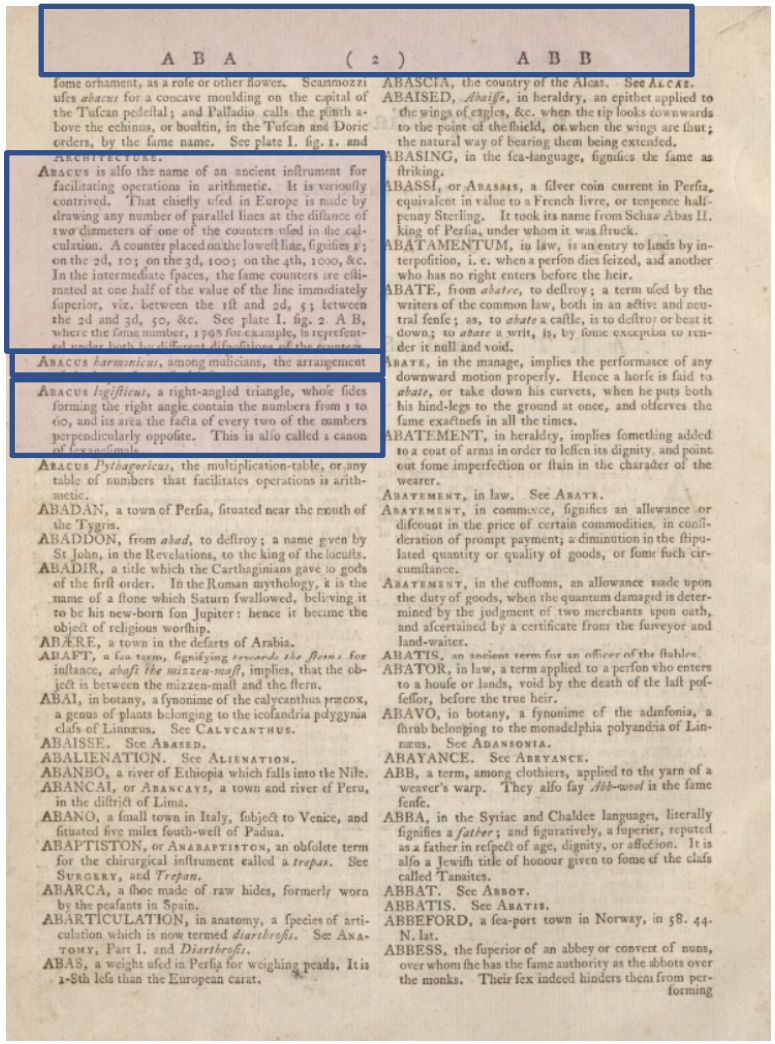
Phase 1: Text Mining

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 + "," .



A B A (2) A B B

some ornament, as a rose or other flower. Scammozzi uses *abacus* for a concave moulding on the capital of the Tuscan pedestal; and Palladio calls the plinth above the echinus, or boultin, 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, signifies 1; on the 2d, 10; on the 3d, 100; on the 4th, 1000, &c. In the intermediate spaces, the same counters are estimated 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, 1765 for example, is represented under both the different dispositions of the counters.

ABACUS *harmonicus*, among musicians, the arrangement

ABACUS *logisticus*, a right-angled triangle, whose sides forming the right angle contain the numbers from 1 to 60, and its area the facta of every two of the numbers perpendicularly opposite. This is also called a canon of sexagesimals.

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

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

ABADDON, from *abad*, 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.

ABÆRE, a town in the deserts of Arabia.

ABAFT, a sea term, signifying *towards the stern*: for instance, *abaft the mizzen-mast*, implies, that the object is between the mizzen-mast and the stern.

ABAI, in botany, a synonime of the calycanthus præcox, a genus of plants belonging to the icosandria polygynia class of Linnæus. See CALYCANTHUS.

ABAISSE. See ABASED.

ABALIENATION. See ALIENATION.

ABANBO, a river of Ethiopia which falls into the Nile.

ABANCAI, 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 *trepan*. See SURGERY, and *Trepan*.

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 pearls. It is 1-8th less than the European carat.

ABASCIA, the country of the Alans. See ALANS.

ABAISED, *Abaisse*, in heraldry, an epithet applied to the wings of eagles, &c. when the tip looks downwards to the point of the shield, or when the wings are shut; the natural way of bearing them being extended.

ABASING, in the sea-language, signifies the same as striking.

ABASSI, or ABASSIS, a silver coin current in Persia, equivalent in value to a French livre, or tenpence halfpenny Sterling. It took its name from Schaw Abas II. king of Persia, under whom it was struck.

ABATAMENTUM, in law, is an entry to lands by interposition, i. e. when a person dies seized, and another who has no right enters before the heir.

ABATE, from *abatre*, to destroy; a term used by the writers of the common law, both in an active and neutral sense; as, to *abate* a castle, is to destroy or beat it down; to *abate* a writ, is, by some exception to render it null and void.

ABATE, in the manage, implies the performance of any downward motion properly. Hence a horse is said to *abate*, or take down his curvets, when he puts both his hind-legs to the ground at once, and observes the same exactness in all the times.

ABATEMENT, in heraldry, implies something added to a coat of arms in order to lessen its dignity, and point out some imperfection or stain in the character of the wearer.

ABATEMENT, in law. See ABATE.

ABATEMENT, in commerce, signifies an allowance or discount in the price of certain commodities, in consideration of prompt payment; a diminution in the stipulated quantity or quality of goods, or some such circumstance.

ABATEMENT, in the customs, an allowance made upon the duty of goods, when the quantum damaged is determined by the judgment of two merchants upon oath, and ascertained by a certificate from the surveyor and land-waiter.

ABATIS, an ancient term for an officer of the stables.

ABATOR, in law, a term applied to a person who enters to a house or lands, void by the death of the last possessor, before the true heir.

ABAVO, in botany, a synonime of the adansonia, a shrub belonging to the monadelphia polyandria of Linnæus. See ADANSONIA.

ABAYANCE. See ABEYANCE.

ABB, a term, among clothiers, applied to the yarn of a weaver's warp. They also say *Abb-wool* in the same sense.

ABBA, in the Syriac and Chaldee languages, literally signifies a *father*; and figuratively, a superior, reputed as a father in respect of age, dignity, or affection. It is also a Jewish title of honour given to some of the class called Tanaites.

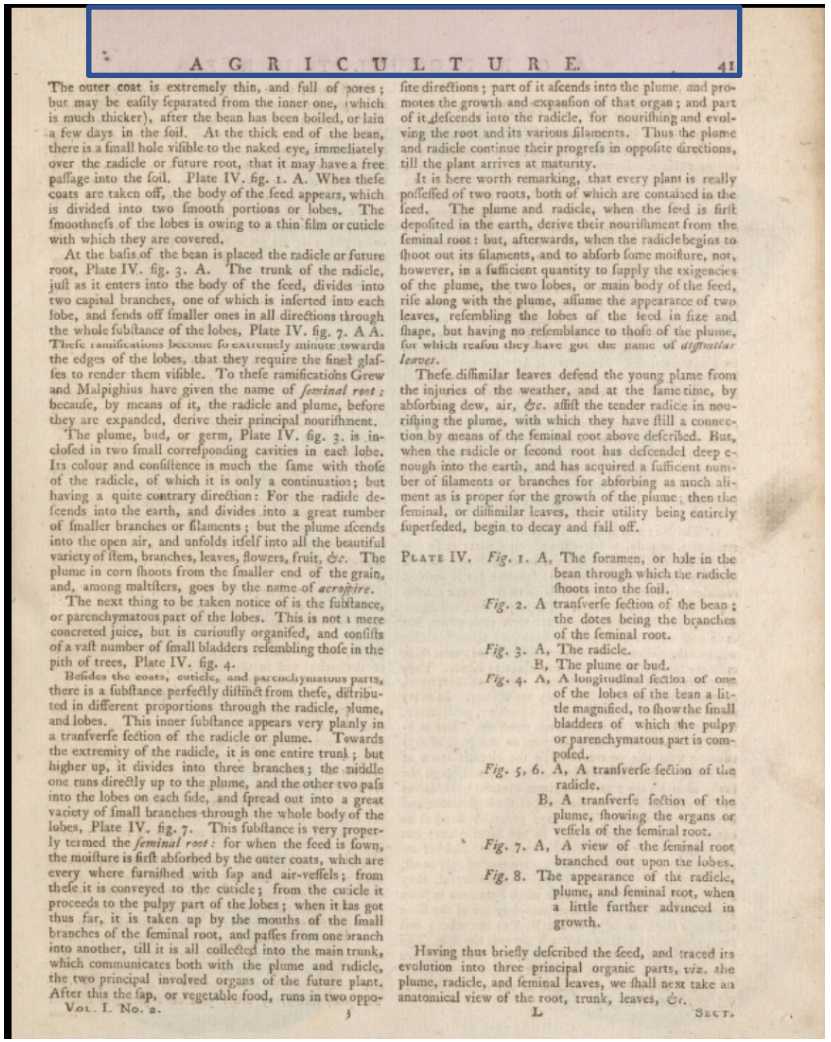
ABBAT. See ABBOT.

ABBATIS. See ABATIS.

ABBEFORD, a sea-port town in Norway, in 58. 44. N. lat.

ABBESS, the superior of an abbey or convent of nuns, over whom she has the same authority as the abbots over the monks. Their sex indeed hinders them from performing

Articles



A G R I C U L T U R E. 41

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 soil. 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 seed 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 seed, divides into two capital branches, one of which is inferted 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 glasses to render them visible. To these ramifications Grew and Malpighius have given the name of *seminal root*; because, by means of it, the radicle and plume, before they are expanded, derive their principal nourishment.

The plume, bud, or germ, Plate IV. fig. 3. is inclosed 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 great 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 maltsters, goes by the name of *acrospire*.

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

Besides the coats, cuticle, and parenchymatous 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 *seminal root*: for when the seed is sown, the moisture is first absorbed by the outer coats, which are every where furnished with sap and air-vessels; from these 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 mouths of the small branches of the seminal 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 seminal 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 seed, rise along with the plume, assume the appearance of two leaves, resembling the lobes of the seed in size and shape, but having no resemblance to those of the plume, for which reason they have got the name of *dissimilar leaves*.

These dissimilar leaves defend the young plume 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 seminal 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 seminal, or dissimilar leaves, their utility being entirely superseded, 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 soil.

Fig. 2. A transverse section of the bean; the dotes being the branches of the seminal root.

Fig. 3. A, The radicle.
B, The plume or bud.

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

Fig. 5, 6. A, A transverse section of the radicle.
B, A transverse section of the plume, shewing the organs or vessels of the seminal root.

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

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

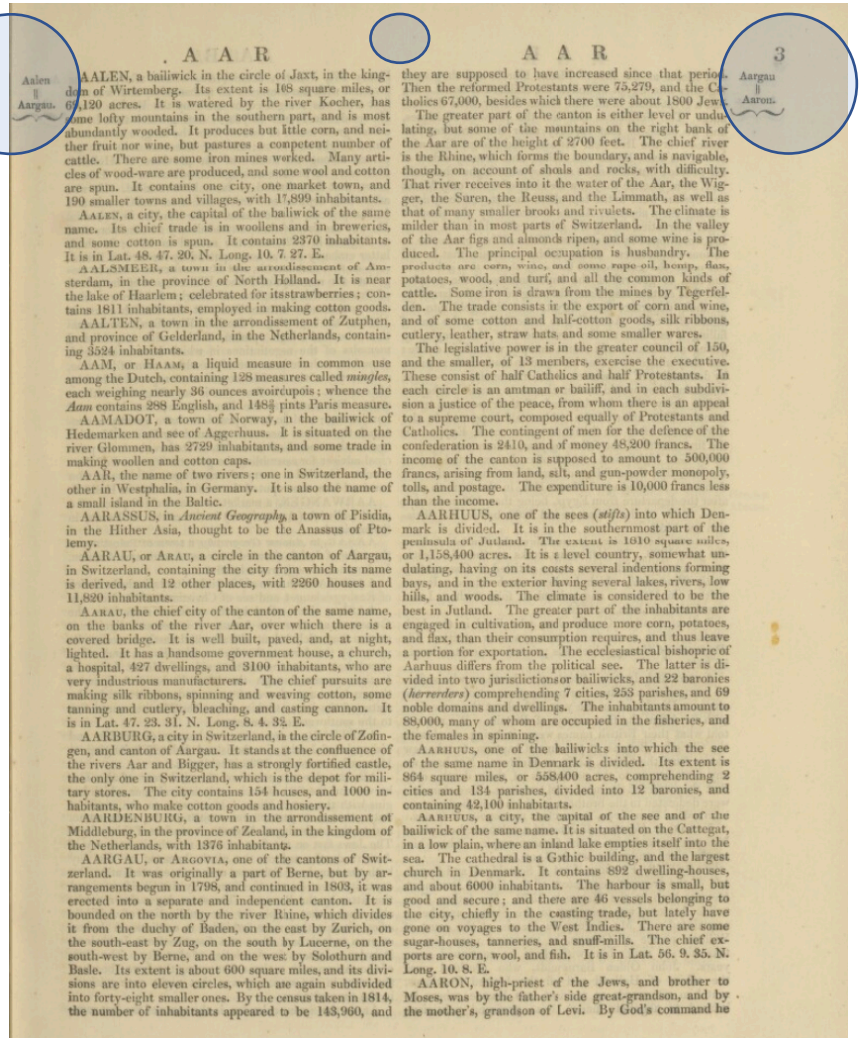
Having thus briefly described the seed, and traced its evolution into three principal organic parts, viz. the plume, radicle, and seminal leaves, we shall next take an anatomical view of the root, trunk, leaves, &c.

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 + “,” .



A A R A A R 3

AALEN, a bailiwick in the circle of Jaxt, in the kingdom of Wirtemberg. Its extent is 168 square miles, or 67,120 acres. It is watered by the river Kocher, has some lofty mountains in the southern part, and is most abundantly wooded. It produces but little corn, and neither fruit nor wine, but pastures a competent number of cattle. There are some iron mines worked. Many articles of wood-ware are produced, and some wool and cotton are spun. It contains one city, one market town, and 190 smaller towns and villages, with 17,899 inhabitants.

AALEN, a city, the capital of the baliwick of the same name. Its chief trade is in woollens and in breweries, and some cotton is spun. It contains 2370 inhabitants. It is in Lat. 48. 47. 20. N. Long. 10. 7. 27. E.

AALSMEER, a town in the arrondissement of Amsterdam, in the province of North Holland. It is near the lake of Haarlem; celebrated for itsstrawberries; contains 1811 inhabitants, employed in making cotton goods.

AALTEN, a town in the arrondissement of Zutphen, and province of Gelderland, in the Netherlands, containing 3524 inhabitants.

AAM, or HAAM, a liquid measure in common use among the Dutch, containing 128 measures called *mingles*, each weighing nearly 36 ounces avoirdupois; whence the *Aam* contains 288 English, and 148⅘ pints Paris measure.

AAMADOT, a town of Norway, in the bailiwick of Hedemarken and see of Aggerhuus. It is situated on the river Glommen, has 2729 inhabitants, and some trade in making woollen and cotton caps.

AAR, the name of two rivers; one in Switzerland, the other in Westphalia, in Germany. It is also the name of a small island in the Baltic.

AARASSUS, in *Ancient Geography*, a town of Pisidia, in the Hither Asia, thought to be the Anassus of Ptolemy.

AARAU, or ARAU, a circle in the canton of Aargau, in Switzerland, containing the city from which its name is derived, and 12 other places, with 2260 houses and 11,820 inhabitants.

AARAU, the chief city of the canton of the same name, on the banks of the river Aar, over which there is a covered bridge. It is well built, paved, and, at night, lighted. It has a handsome government house, a church, a hospital, 427 dwellings, and 3100 inhabitants, who are very industrious manufacturers. The chief pursuits are making silk ribbons, spinning and weaving cotton, some tanning and cutlery, bleaching, and casting cannon. It is in Lat. 47. 23. 31. N. Long. 8. 4. 32. E.

AARBURG, a city in Switzerland, in the circle of Zofingen, and canton of Aargau. It stands at the confluence of the rivers Aar and Bigger, has a strongly fortified castle, the only one in Switzerland, which is the depot for military stores. The city contains 154 houses, and 1000 inhabitants, who make cotton goods and hosiery.

AARDENBURG, a town in the arrondissement of Middleburg, in the province of Zealand, in the kingdom of the Netherlands, with 1376 inhabitants.

AARGAU, or ARGOVIA, one of the cantons of Switzerland. It was originally a part of Berne, but by arrangements begun in 1798, and continued in 1803, it was erected into a separate and independent canton. It is bounded on the north by the river Rhine, which divides it from the duchy of Baden, on the east by Zurich, on the south-east by Zug, on the south by Lucerne, on the south-west by Berne, and on the west by Solothurn and Basle. Its extent is about 600 square miles, and its divisions are into eleven circles, which are again subdivided into forty-eight smaller ones. By the census taken in 1814, the number of inhabitants appeared to be 143,960, and they are supposed to have increased since that period. Then the reformed Protestants were 75,279, and the Catholics 67,000, besides which there were about 1800 Jews.

The greater part of the canton is either level or undulating, but some of the mountains on the right bank of the Aar are of the height of 2700 feet. The chief river is the Rhine, which forms the boundary, and is navigable, though, on account of shoals and rocks, with difficulty. That river receives into it the water of the Aar, the Wigger, the Suren, the Reuss, and the Limmath, as well as that of many smaller brooks and rivulets. The climate is milder than in most parts of Switzerland. In the valley of the Aar figs and almonds ripen, and some wine is produced. The principal occupation is husbandry. The products are corn, wine, and some rape oil, hemp, flax, potatoes, wood, and turf, and all the common kinds of cattle. Some iron is drawn from the mines by Tegerfelden. The trade consists in the export of corn and wine, and of some cotton and half-cotton goods, silk ribbons, cutlery, leather, straw hats and some smaller wares.

The legislative power is in the greater council of 150, and the smaller, of 13 members, exercise the executive. These consist of half Cathelics and half Protestants. In each circle is an amtman or bailiff, and in each subdivision a justice of the peace, from whom there is an appeal to a supreme court, composed equally of Protestants and Catholics. The contingent of men for the defence of the confederation is 2410, and of money 48,200 francs. The income of the canton is supposed to amount to 500,000 francs, arising from land, silt, and gun-powder monopoly, tolls, and postage. The expenditure is 10,000 francs less than the income.

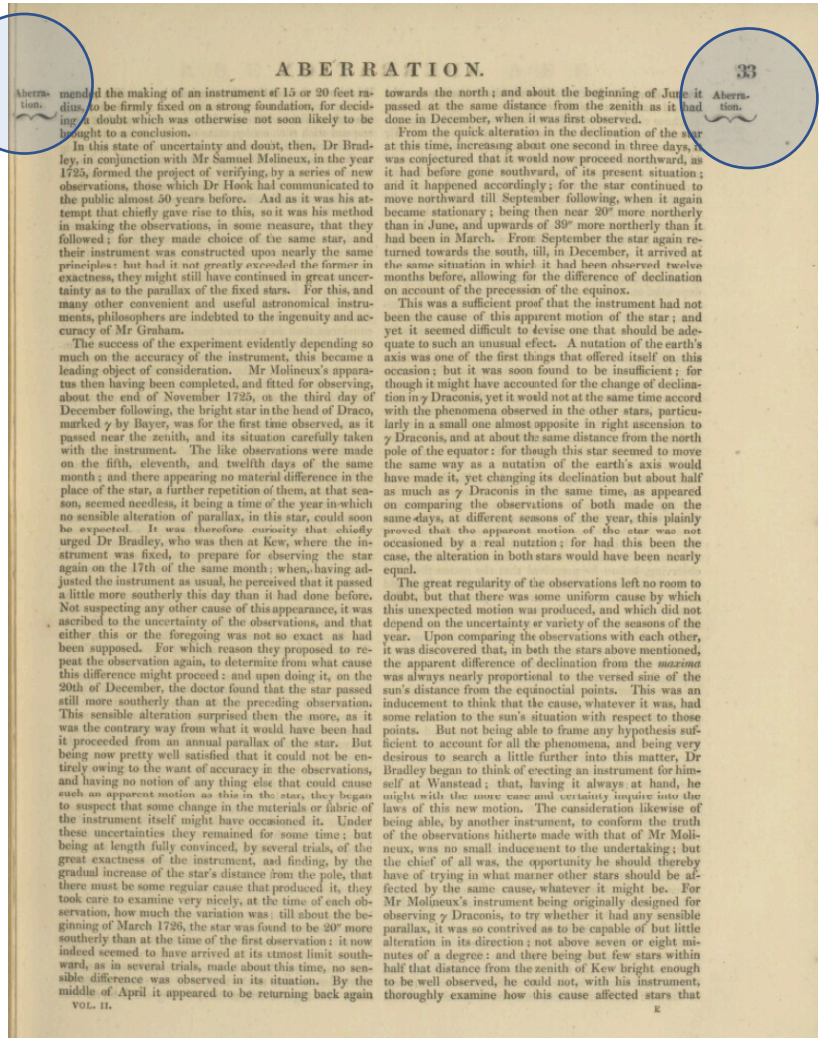
AARHUUS, one of the sees (*stifts*) into which Denmark is divided. It is in the southernmost part of the peninsula of Jutland. The extent is 1610 square miles, or 1,158,400 acres. It is a level country, somewhat undulating, having on its coasts several indentions forming bays, and in the exterior having several lakes, rivers, low hills, and woods. The climate is considered to be the best in Jutland. The greater part of the inhabitants are engaged in cultivation, and produce more corn, potatoes, and flax, than their consumption requires, and thus leave a portion for exportation. The ecclesiastical bishopric of Aarhuus differs from the political see. The latter is divided into two jurisdictions or bailiwicks, and 22 baronies (*herrerders*) comprehending 7 cities, 253 parishes, and 69 noble domains and dwellings. The inhabitants amount to 88,000, many of whom are occupied in the fisheries, and the females in spinning.

AARHUUS, one of the bailiwicks into which the see of the same name in Denmark is divided. Its extent is 864 square miles, or 558,400 acres, comprehending 2 cities and 134 parishes, divided into 12 baronies, and containing 42,100 inhabitants.

AARHUUS, a city, the capital of the see and of the bailiwick of the same name. It is situated on the Cattegat, in a low plain, where an inland lake empties itself into the sea. The cathedral is a Gothic building, and the largest church in Denmark. It contains 892 dwelling-houses, and about 6000 inhabitants. The harbour is small, but good and secure; and there are 46 vessels belonging to the city, chiefly in the coasting trade, but lately have gone on voyages to the West Indies. There are some sugar-houses, tanneries, and snuff-mills. The chief exports are corn, wool, and fish. It is in Lat. 56. 9. 35. N. Long. 10. 8. E.

AARON, high-priest of the Jews, and brother to Moses, was by the father's side great-grandson, and by the mother's, grandson of Levi. By God's command he

Articles



ABERRATION. 33

mended the making of an instrument of 15 or 20 feet radius, to be firmly fixed on a strong foundation, for deciding a doubt which was otherwise not soon likely to be brought to a conclusion.

In this state of uncertainty and doubt, then, Dr Bradley, in conjunction with Mr Samuel Molineux, in the year 1725, formed the project of verifying, by a series of new observations, those which Dr Hook had communicated to the public almost 50 years before. And as it was his attempt that chiefly gave rise to this, so it was his method in making the observations, in some measure, that they followed; for they made choice of the same star, and their instrument was constructed upon nearly the same principles: but had it not greatly exceeded the former in exactness, they might still have continued in great uncertainty as to the parallax of the fixed stars. For this, and many other convenient and useful astronomical instruments, philosophers are indebted to the ingenuity and accuracy of Mr Graham.

The success of the experiment evidently depending so much on the accuracy of the instrument, this became a leading object of consideration. Mr Molineux's apparatus then having been completed, and fitted for observing, about the end of November 1725, on the third day of December following, the bright star in the head of Draco, marked γ by Bayer, was for the first time observed, as it passed near the zenith, and its situation carefully taken with the instrument. The like observations were made on the fifth, eleventh, and twelfth days of the same month; and there appearing no material difference in the place of the star, a further repetition of them, at that season, seemed needless, it being a time of the year in which no sensible alteration of parallax, in this star, could soon be expected. It was therefore curiosity that chiefly urged Dr Bradley, who was then at Kew, where the instrument was fixed, to prepare for observing the star again on the 17th of the same month; when, having adjusted the instrument as usual, he perceived that it passed a little more southerly this day than it had done before. Not suspecting any other cause of this appearance, it was ascribed to the uncertainty of the observations, and that either this or the foregoing was not so exact as had been supposed. For which reason they proposed to repeat the observation again, to determine from what cause this difference might proceed: and upon doing it, on the 20th of December, the doctor found that the star passed still more southerly than at the preceding observation. This sensible alteration surprised them the more, as it was the contrary way from what it would have been had it proceeded from an annual parallax of the star. But being now pretty well satisfied that it could not be entirely owing to the want of accuracy in the observations, and having no notion of any thing else that could cause such an apparent motion as this in the star, they began to suspect that some change in the materials or fabric of the instrument itself might have occasioned it. Under these uncertainties they remained for some time; but being at length fully convinced, by several trials, of the great exactness of the instrument, and finding, by the gradual increase of the star's distance from the pole, that there must be some regular cause that produced it, they took care to examine very nicely, at the time of each observation, how much the variation was: till about the beginning of March 1726, the star was found to be 20″ more southerly than at the time of the first observation: it now indeed seemed to have arrived at its utmost limit southward, as in several trials, made about this time, no sensible difference was observed in its situation. By the middle of April it appeared to be returning back again towards the north; and about the beginning of June it passed at the same distance from the zenith as it had done in December, when it was first observed.

From the quick alteration in the declination of the star at this time, increasing about one second in three days, it was conjectured that it would now proceed northward, as it had before gone southward, of its present situation; and it happened accordingly; for the star continued to move northward till September following, when it again became stationary; being then near 20″ more northerly than in June, and upwards of 39″ more northerly than it had been in March. From September the star again returned towards the south, till, in December, it arrived at the same situation in which it had been observed twelve months before, allowing for the difference of declination on account of the precession of the equinox.

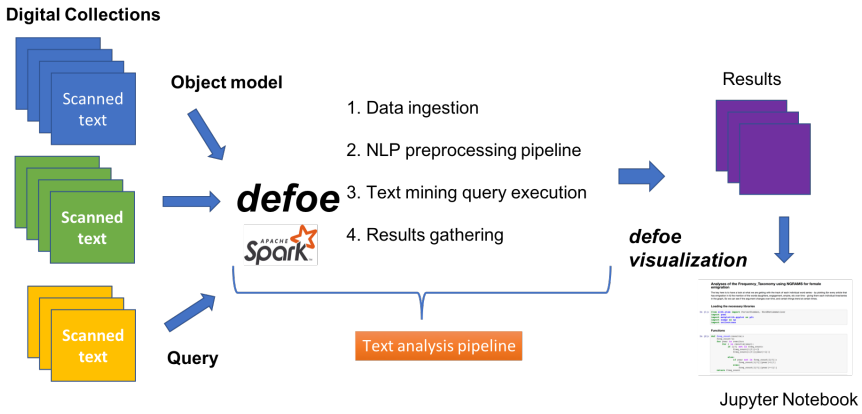
This was a sufficient proof that the instrument had not been the cause of this apparent motion of the star; and yet it seemed difficult to devise one that should be adequate to such an unusual effect. A nutation of the earth's axis was one of the first things that offered itself on this occasion; but it was soon found to be insufficient; for though it might have accounted for the change of declination in γ Draconis, yet it would not at the same time accord with the phenomena observed in the other stars, particularly in a small one almost opposite in right ascension to γ Draconis, and at about the same distance from the north pole of the equator: for though this star seemed to move the same way as a nutation of the earth's axis would have made it, yet changing its declination but about half as much as γ Draconis in the same time, as appeared on comparing the observations of both made on the same days, at different seasons of the year, this plainly proved that the apparent motion of the star was not occasioned by a real nutation; for had this been the case, the alteration in both stars would have been nearly equal.

The great regularity of the observations left no room to doubt, but that there was some uniform cause by which this unexpected motion was produced, and which did not depend on the uncertainty or variety of the seasons of the year. Upon comparing the observations with each other, it was discovered that, in both the stars above mentioned, the apparent difference of declination from the *maxima* was always nearly proportional to the versed sine of the sun's distance from the equinoctial points. This was an inducement to think that the cause, whatever it was, had some relation to the sun's situation with respect to those points. But not being able to frame any hypothesis sufficient to account for all the phenomena, and being very desirous to search a little further into this matter, Dr Bradley began to think of erecting an instrument for himself at Wanstead; that, having it always at hand, he might with the more ease and certainty inquire into the laws of this new motion. The consideration likewise of being able, by another instrument, to conform the truth of the observations hitherto made with that of Mr Molineux, was no small inducement to the undertaking; but the chief of all was, the opportunity he should thereby have of trying in what manner other stars should be affected by the same cause, whatever it might be. For Mr Molineux's instrument being originally designed for observing γ Draconis, to try whether it had any sensible parallax, it was so contrived as to be capable of but little alteration in its direction; not above seven or eight minutes of a degree: and there being but few stars within half that distance from the zenith of Kew bright enough to be well observed, he could not, with his instrument, thoroughly examine how this cause affected stars that

Topic

Edition 7

Extract EB Terms



term	ABACTORES
definition	or ABACTORS, a term for such as carry offer dr...
relatedTerms	[]
header	EBAA
startsAt	15
endsAt	15
numberOfTerms	22
numberOfWords	18
numberOfPages	832
positionPage	18
typeTerm	Article
editionTitle	First edition, 1771, Volume 1, A–B
editionNum	1
supplementTitle	
supplementsTo	[]
year	1771
place	Edinburgh
volumeTitle	Encyclopaedia Britannica; or, A dictionary of ...
volumeNum	1
letters	A–B
part	0
altoXML	144133901/alto/188082904.34.xml

Name: 18, dtype: object

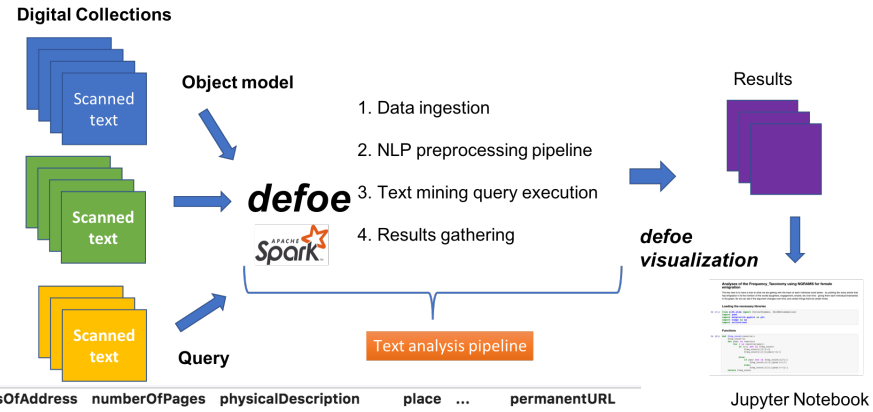
Extracted Term ABACATORES information - Edition 1, 1771, Vol A-B

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
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	encyc	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
		Third edition,										

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**

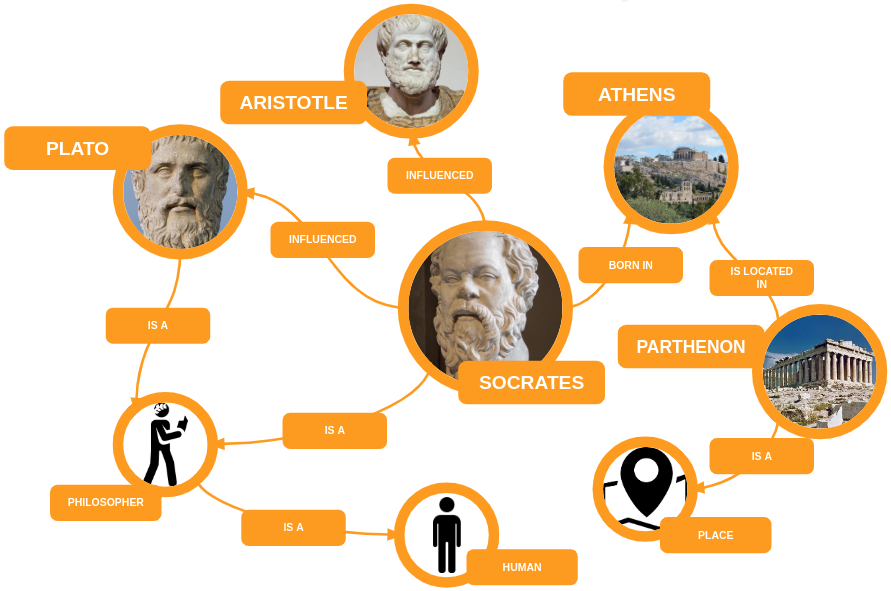
(*) **Also Based on Heuristics -->** Pages Layout & Text & Headers → different heuristics for different EB editions.

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

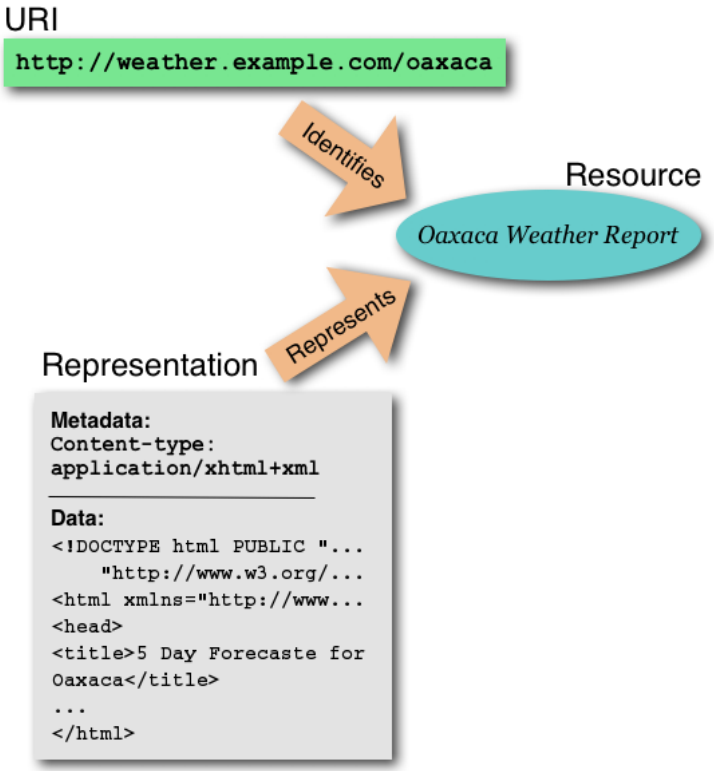
"The application of graph processing and **graph DBMSs will grow at 100 percent annually through 2022** to continuously accelerate data preparation and enable more complex and adaptive data science."
– Gartner's *Top 10 Data and Analytics Technology Trends for 2019*

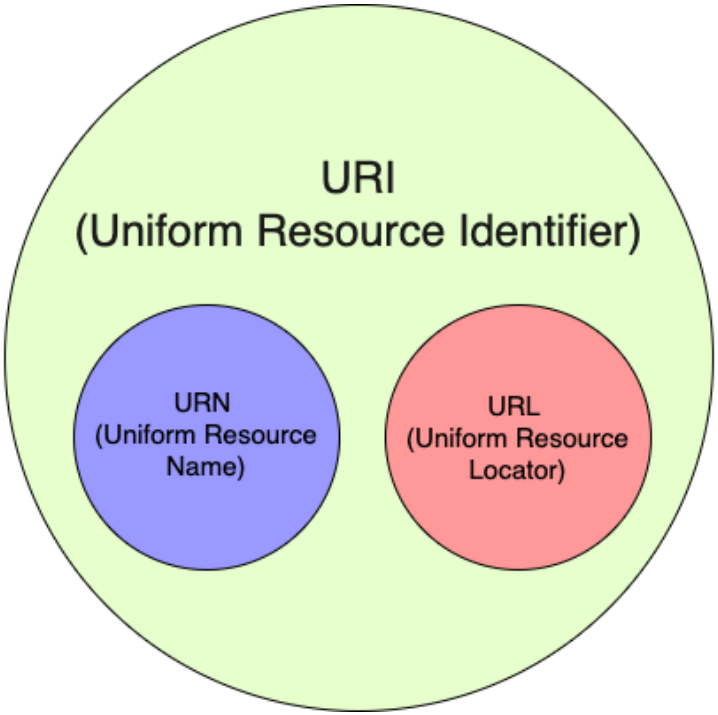


Phase 2: Knowledge Graph

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

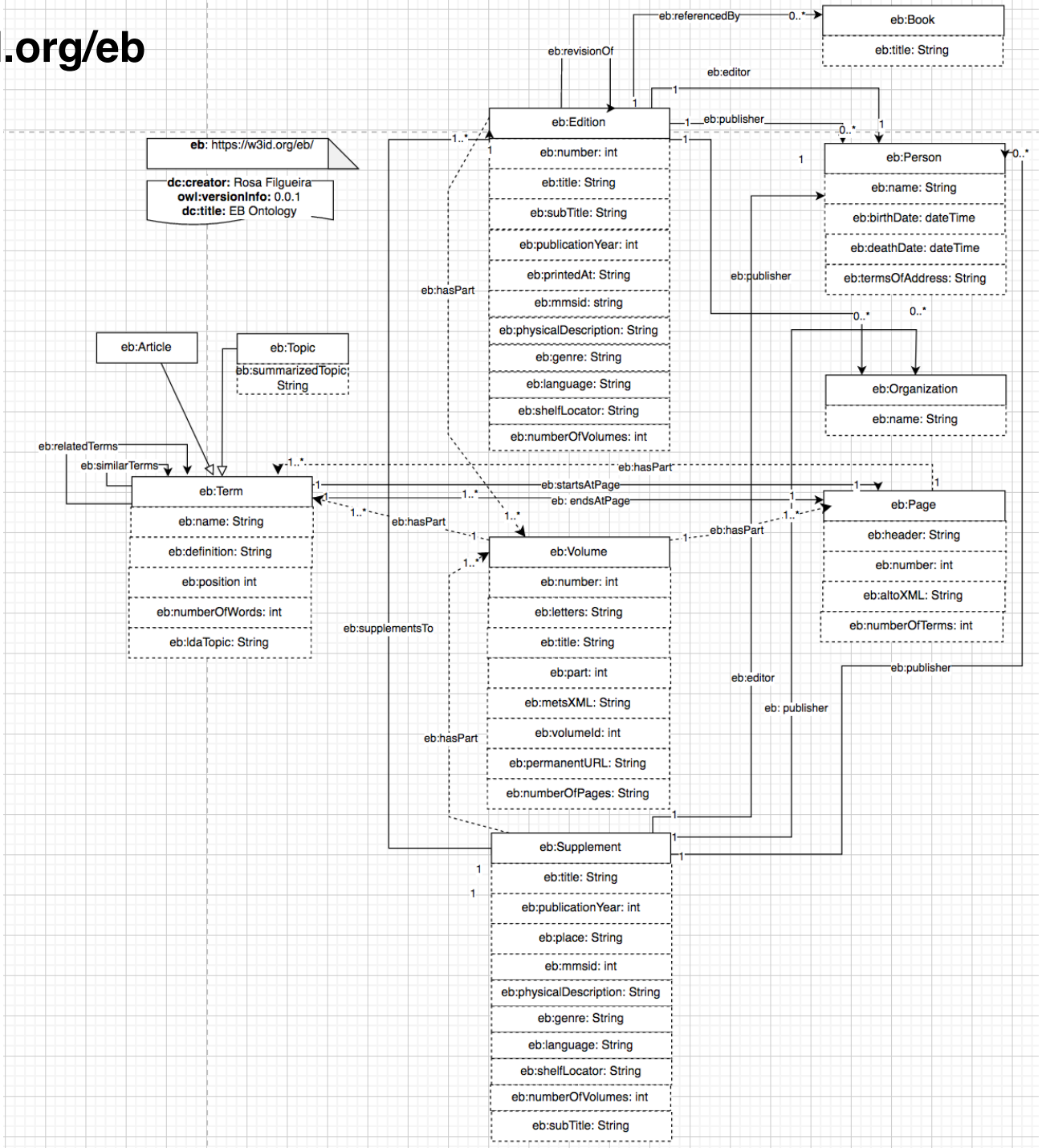
2.1. Create EB Ontology: To explain (give meaning) how our "things" are related with each other.

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/

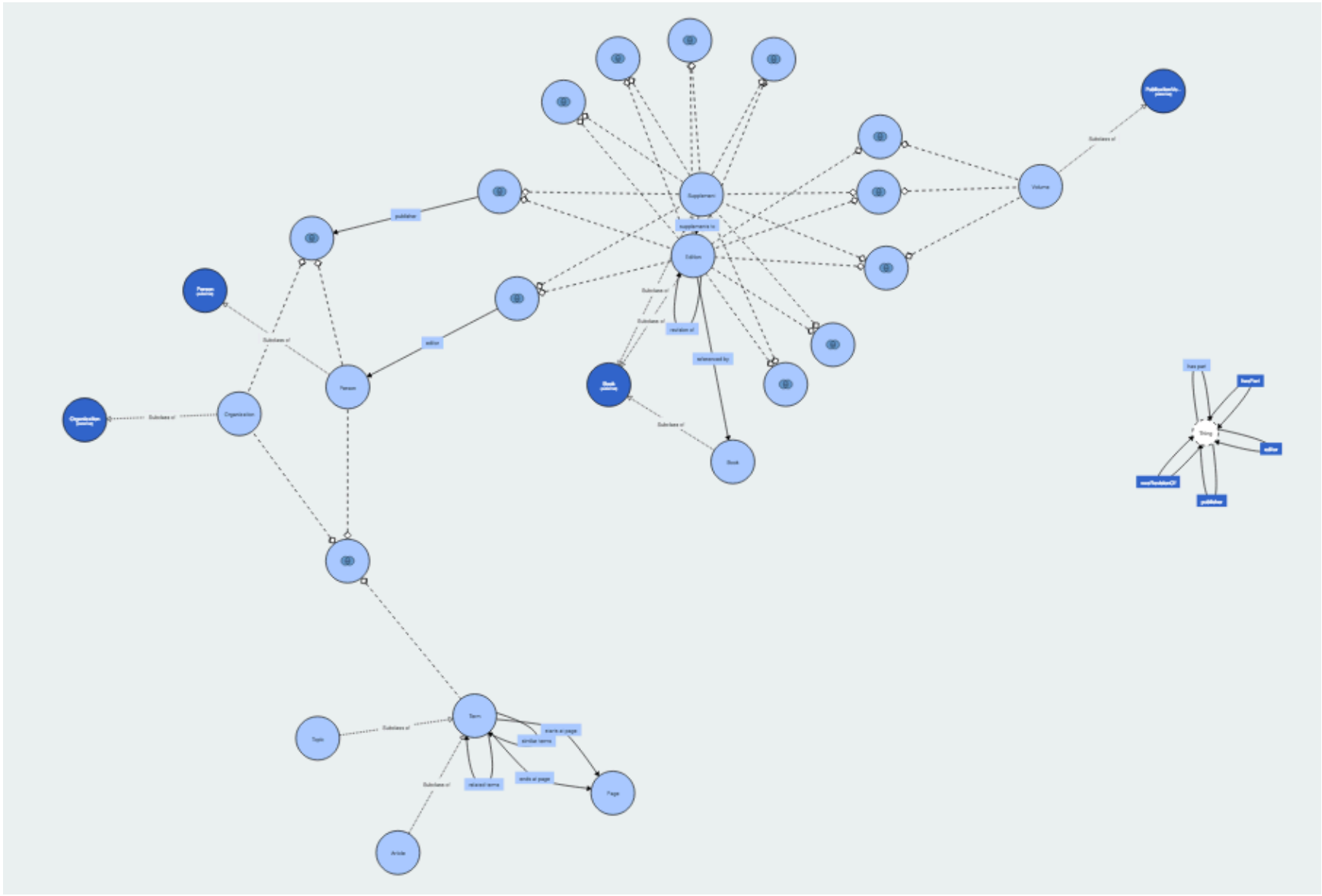
2.2 . Created **EB Knowledge Graph 1.0:** Populated the post-processed information (extracted **Terms & Metadata v.2**) into a RDF triplestore using the EB Ontology

- 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

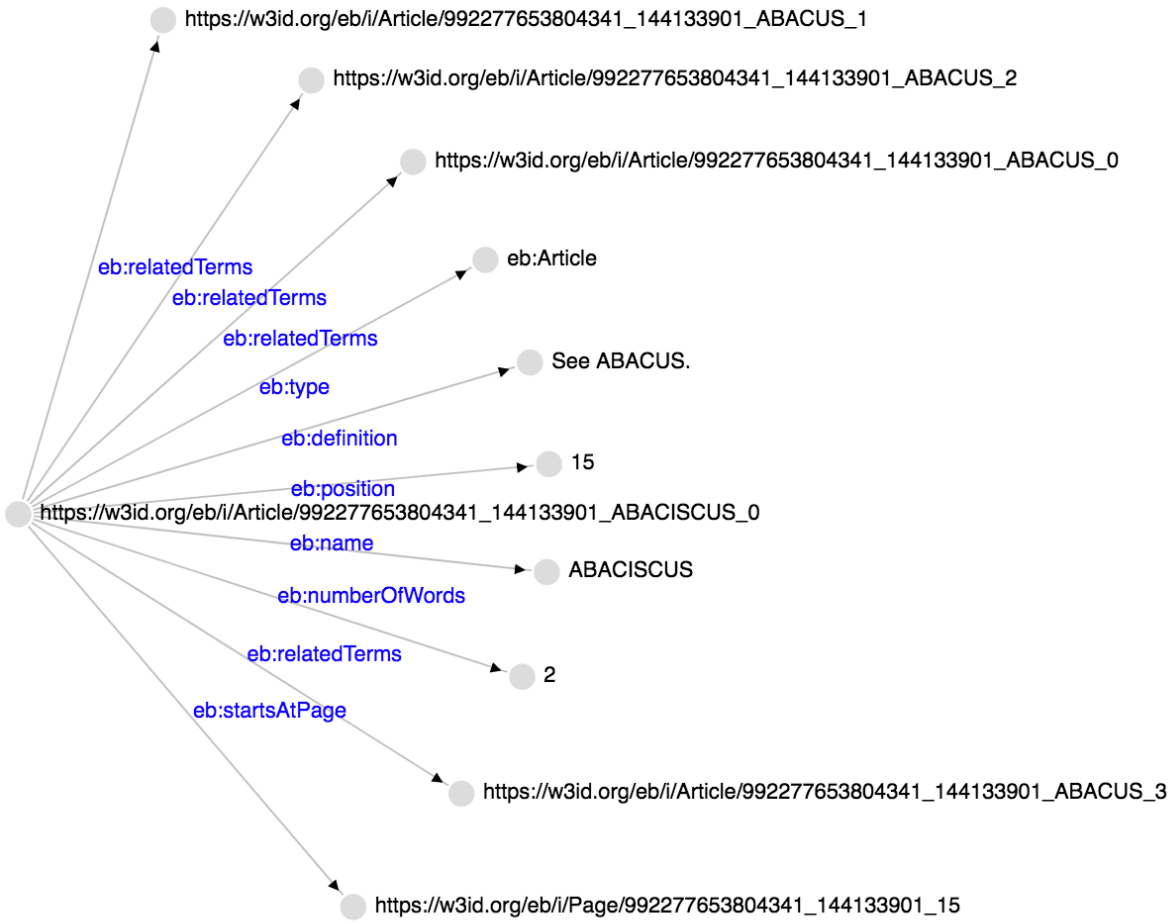
NOTE:

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

EB Knowledge Graph

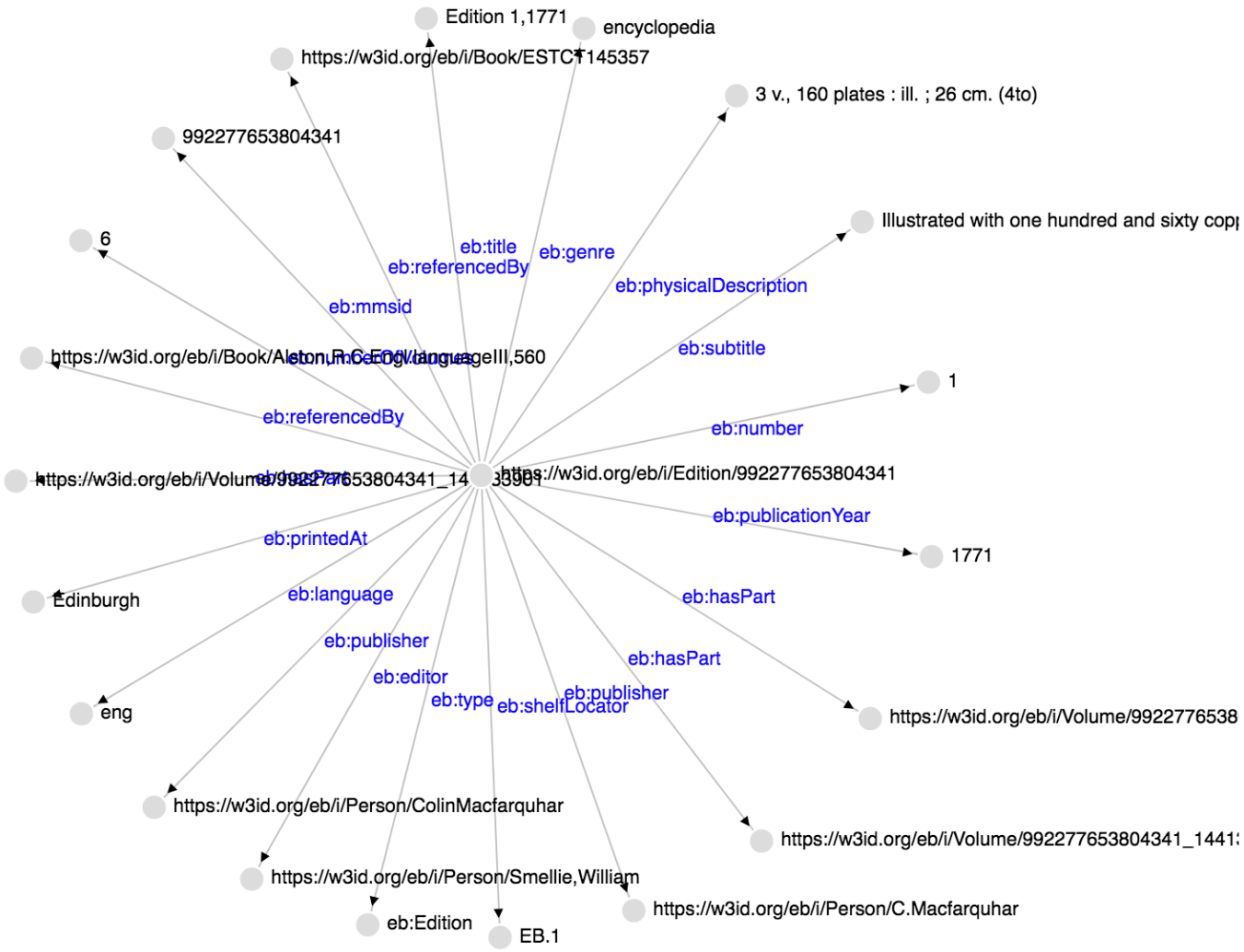


EB Knowledge Graph



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.

```
SELECT  ?article ?definition
    WHERE {
        ?article a eb:Article .
        ?article eb:definition ?definition
        FILTER (CONTAINS(?definition, "Scotland"))
        FILTER (CONTAINS(?definition, "Glasgow"))
        OPTIONAL{FILTER CONTAINS(?definition, "Edinburgh") }
} LIMIT 10
```

Fuseki RDF triplestore

QUERY RESULTS

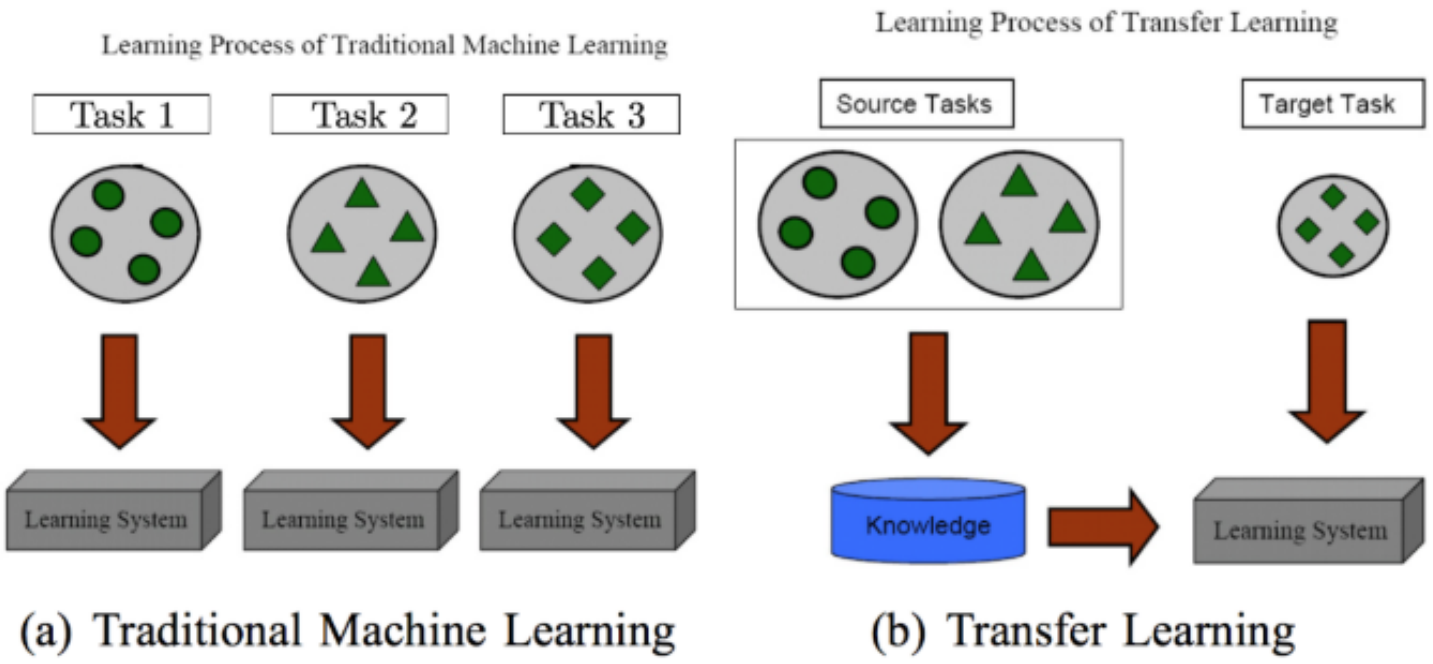
Table | Raw Response

Showing 1 to 8 of 8 entries

Search: Show 50 entries

	article	definition
1	<https://w3id.org/eb/i/Article/9929192893804340_144850368_PAISLEY_0>	"a town of Scotland, in the county of Renfrew, six miles west of Glasgow."
2	<https://w3id.org/eb/i/Article/9929192893804340_144850366_ARGYLESHIRE_0>	"a county of Scotland, lying westward of Glasgow, and comprehending the countries of Lorn, Cowal, Knapdale, Kintyfe, together with the islands Mull, Jura, Ifla, foe. It gives the title of duke to the noble family of Campbell."
3	<https://w3id.org/eb/i/Article/9929192893804340_144850367_INVERARY_0>	"a parliament town of Scotland, in the county of Argyle, of which it is the capital, fituatedin Lochhn, forty five miles north-west of Glasgow : W. long. 5 0 , N. lat 36° 28-."
4	<https://w3id.org/eb/i/Article/9929192893804340_144850367_HAMILTON_0>	"a town of Scotland, in the county of Clydefdale, situated on the river Cl^de, eleven miles south-east of Glasgow : W. long. 3 0 jo', N. lat. 55 0"

Phase 3: Augmented Knowledge Graph with Deep Transfer Learning



Phase 3: Augmented Knowledge Graph with Deep Transfer Learning

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

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 + ElmosclstmChecker*
- **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 .

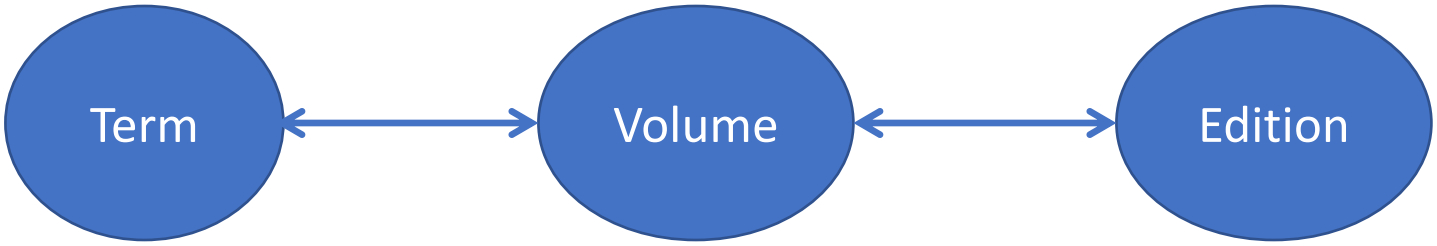
↓ Compute Difference

the morst northerly of any of the w efastern islands of Scotland , lying in 8\u00b0 and odd minutes W. long , and between 58\u00b0 and 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.



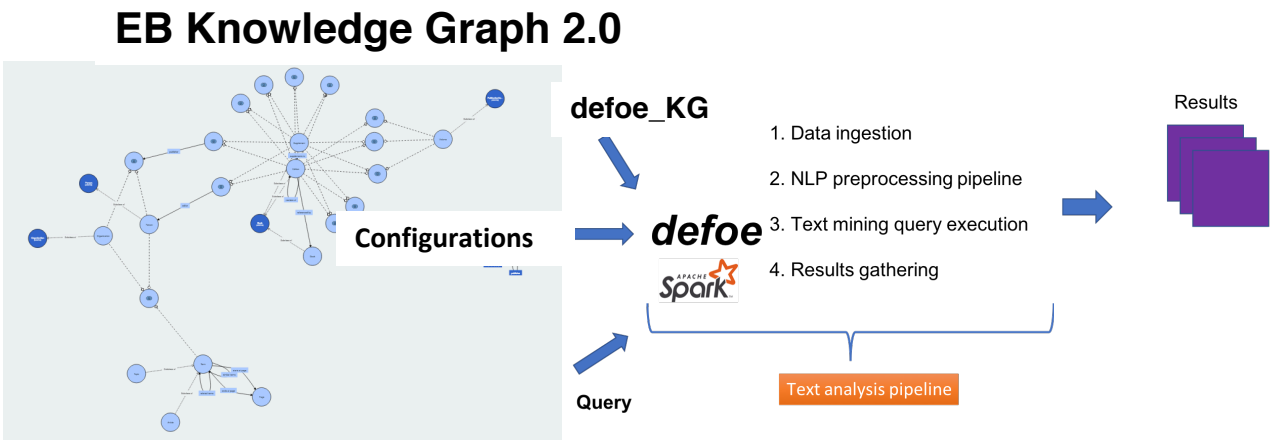
Example: Given a **Term** (e.g., *Edinburgh*), I can get the **Edition** Information (e.g., *Edition Title*)

- 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**

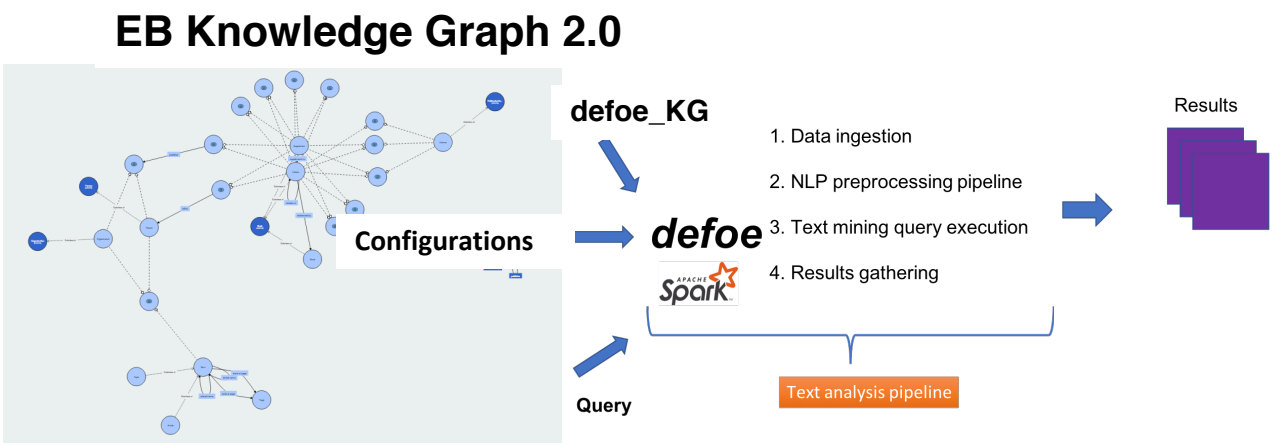
4.2. Improved **defoe queries to be fully configurable**: different filtering options, target, lexicon, etc.



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 snippet of definition** in which appear keywords or keysentences and group by years. Several filtering options, including the snippet size.
- *publication_normalization:* **Extract 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.

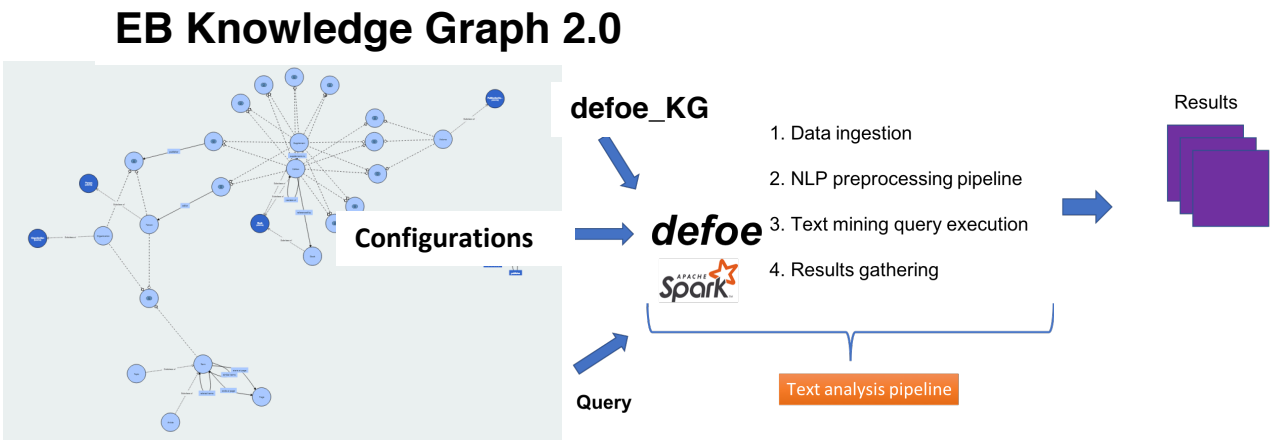


Phase 4: Defoe and Knowledge Graph

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)



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 not term is introduced, it will search for the first term in the Encyclopaedia.

Search term

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\u00b0. We ffiall... More		15_scotland_county_edinburgh_frith	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\u00b0. We (hall ... More		15_scotland_county_edinburgh_frith	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 not term is introduced, it will search for the first term in the Encyclopaedia.

Search term

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\u00b0. We ffiall... More		15_scotland_county_edinburgh_frith	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\u00b0. We (hall ... More		15_scotland_county_edinburgh_frith	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).

frances: Architecture