The Wittgenstein Project

Data Semantics – 2022/23

Davide Giardini

Chapters

01.
Presentation

Presentation of the main idea of the project

02.
Past Works

Past Works and Inspirations

03.
The Corpus

Building and cleaning the text

04.
The Work

Word2Vec, CADE and SWEAT

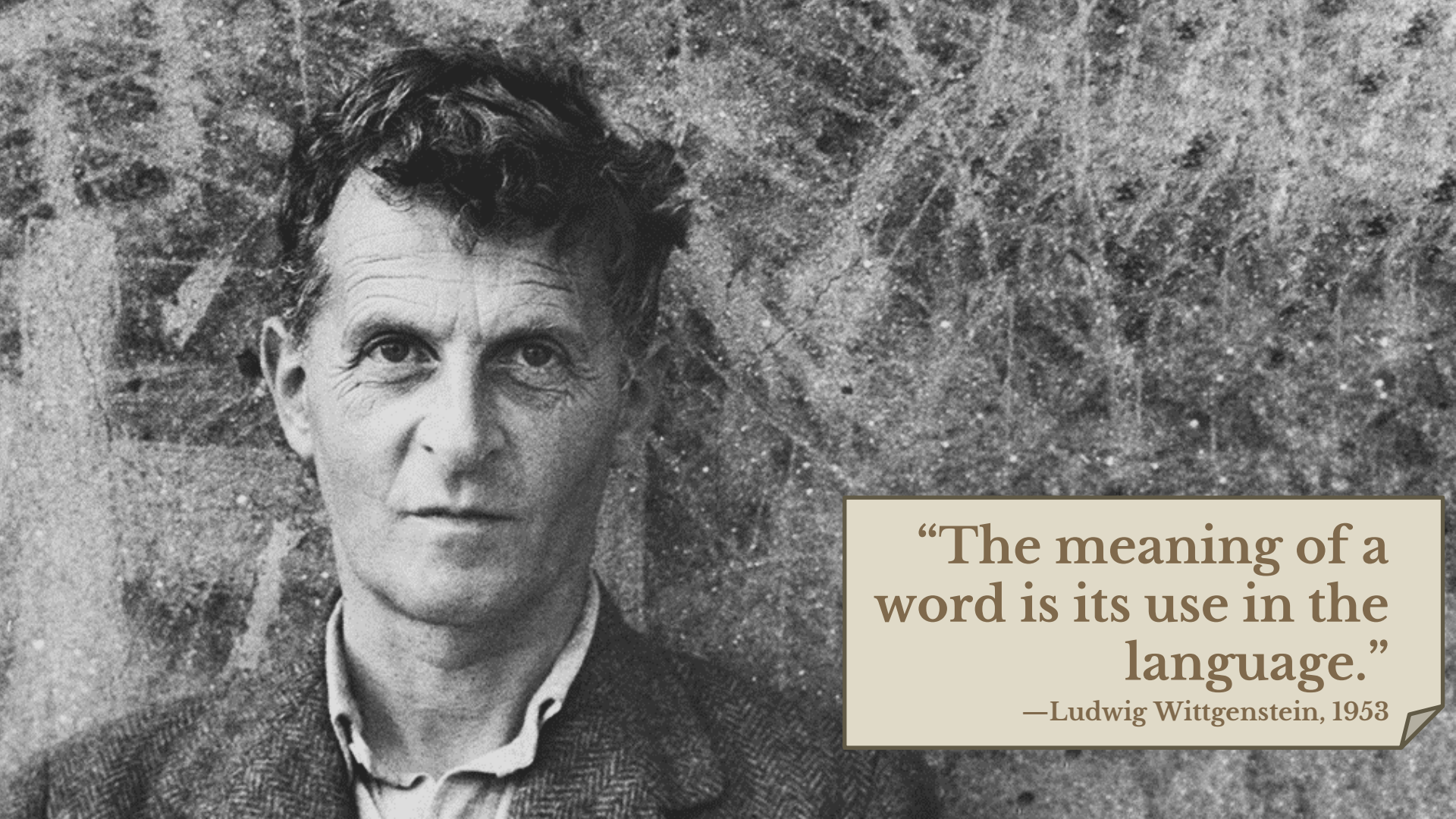
05.
Results

Conclusion and future developments

01

Presentation

Presentation of the main idea of the project


“The meaning of a word is its use in the language.”
—Ludwig Wittgenstein, 1953

Questions

What are the differences in the way the various philosophical schools approach topics such as life, death, humanity, ideas, the soul, reason, God, nature ...? **What is closer**, for one philosophy that does not assume the existence of a soul, **to the concept** that a second philosophy instead ascribes to it? Is it possible that, for some thinkers, reason and logic assume the same characteristics that others place on ideas such as God? How does the **polarity** relative to these philosophers' perception of man, existence, or morality change?

Can word embeddings algorithms identify the differences that delineate the thinking of each school of philosophers? Or does the complexity of the subject matter mean that, for now, man's thoughts remain interpretable only by humans themselves and **inaccessible to machines?**

Past Works

Past Works and Inspirations

Papers

Word Embedding Driven Concept Detection in Philosophical Corpora

Dylan Hayton-Ruffner

Distributional techniques for philosophical enquiry

A. Herbelot, E. Von Redecker, J. Muller

Challenging Distributional Models with a Conceptual Network of Philosophical Terms

Y. Oortwijn et. al.

The Philosophy Data Project

WordClouds

Reading Difficulty

Citations

philosophydata.com

«applies modern data analysis techniques to great texts in the history of philosophy.»

Kourosh Alizadeh

Sentiment Analysis

Classification

Word2Vec

The Corpus

Building and cleaning
the text

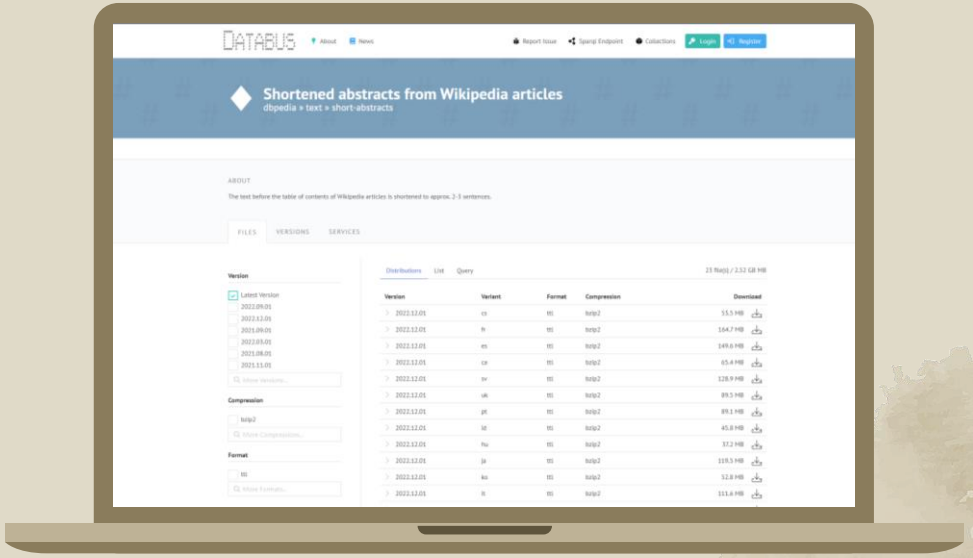
Data Acquisition

School	Author
Analytic	Kripke, Lewis, Moore, Popper, Quine, Russel, Wittgenstein
Aristotle	Aristotle
Empiricism	Berkeley, Hume, Locke
German Idealism	Fichte, Hegel, Kant
Nihilism	Nietzsche, Kierkegaard
Plato	Plato
Rationalism	Descartes, Leibniz, Malebranche, Spinoza

The Neutral Slice

An **eight slice** has been implemented from **Wikipedia's abstract**. Since this corpus would have been way bigger than the others, a **downsampling** technique has been used (3 mln tokens).
The usefulness of this additional slice is:

- **More data** to train the compass.
- Possibility to draw **comparisons** with a «**neutral**» slice.



Pre-processing

Separation in Sentences

With the “Natural Language ToolKit” (NLTK)

1

Capital Letters removal

Set all the text to lower-case

2

Punctuation removal

Remove punctuation, digits, and other non-alphabethic characters

3

Pre-processing

Stop-Words removal

Removing the words that occur commonly across all the corpus (mostly articles and pronouns)

4

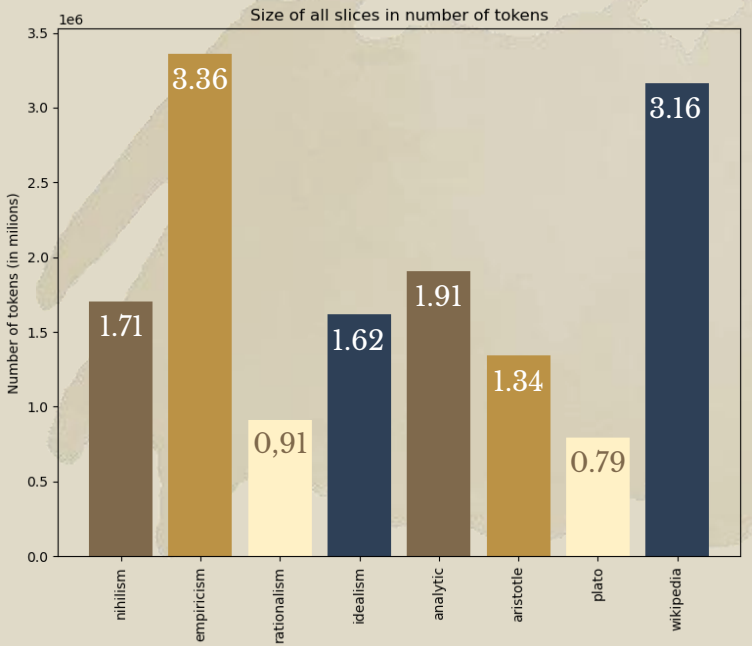
Lemmatization

Grouping inflected forms together as a single base form

5

2 Corpus Ready

Stats

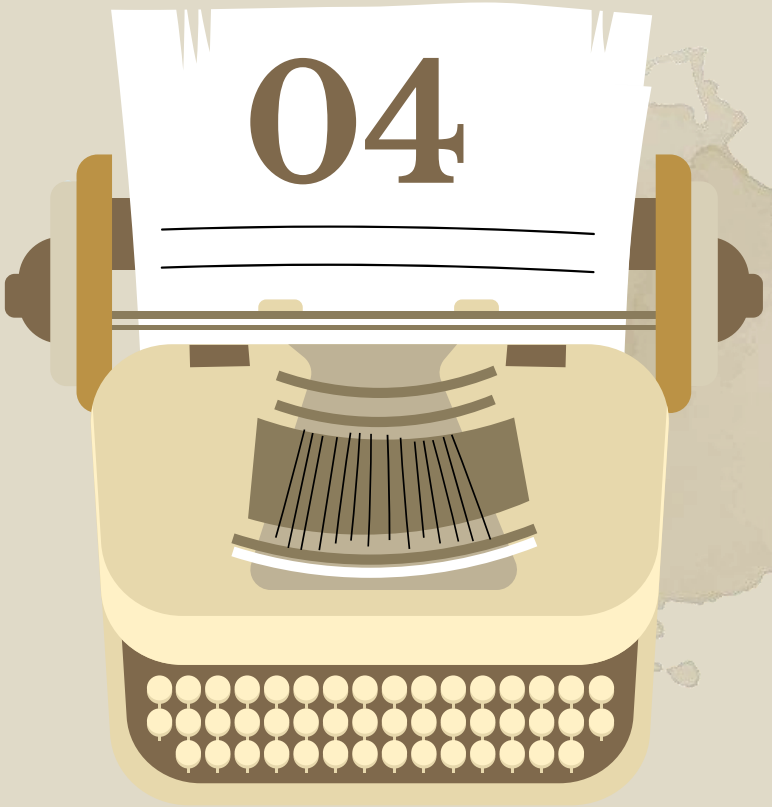


In total:

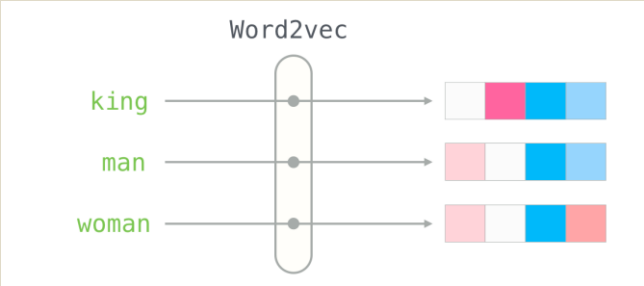
14.7 mln tokens
620 thousand sentences

The Work

Word2Vec, CADE and SWEAT

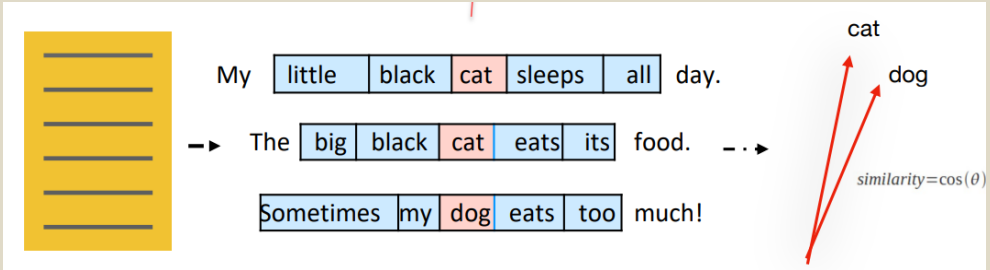


Word2vec



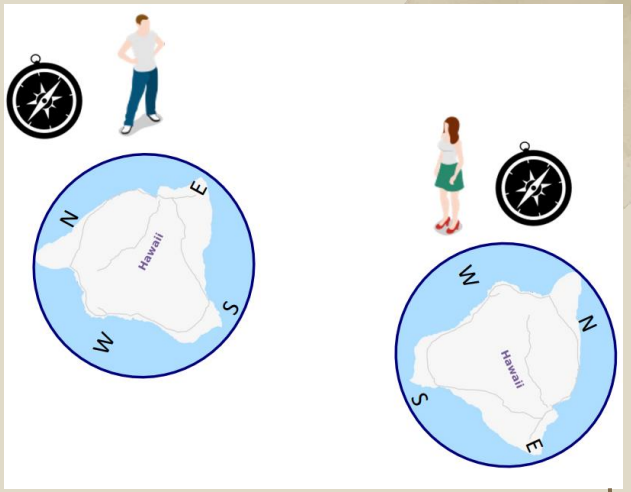
The word2vec algorithm uses a **neural network** model to learn word associations from a large corpus of text.
As the name implies, word2vec **represents each distinct word with a vector**.

The vectors **capture the semantic and syntactic qualities of words**.
In this way, a simple mathematical function such as the **cosine similarity** can indicate the level of semantic similarity between the words represented by those vectors.



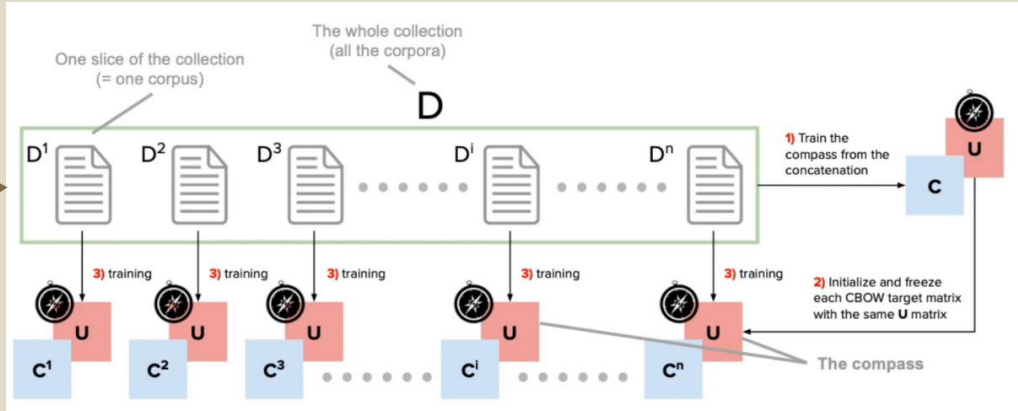
Visualizations from the slides of Matteo Palmonari – Federico Bianchi & Jay Alammar

CADE



The Compass Alligned Distributional Embedding (CADE) method helps us the resolve the **allignment problem** and **compare** distributional models derived from **different corpora**.

To do so, it trains the compass from the total corpus (the concatenation of all the slices) and **uses the resulting target matrix U to train all of the different slices**.

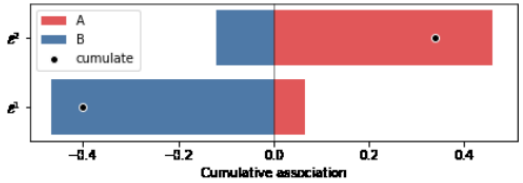


Visualizations from the slides of Matteo Palmonari – Federico Bianchi

SWEAT

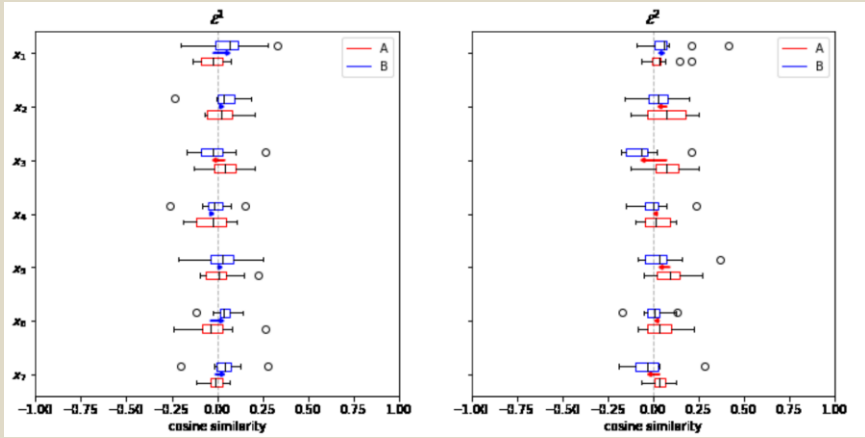
$$s(x, \mathcal{E}, A, B) = \frac{1}{|A|}\sum_{a \in A} \cos\big(\mathcal{E}(x), \mathcal{E}(a)\big) - \frac{1}{|B|}\sum_{b \in B} \cos\big(\mathcal{E}(x), \mathcal{E}(b)\big)$$

$$S(X, \mathcal{E}^1, \mathcal{E}^2, A, B) = \sum_{x \in X} s(x, \mathcal{E}^1, A, B) - \sum_{x \in X} s(x, \mathcal{E}^2, A, B)$$



The SWEAT algorithm compares the same embedding (vector/word) in different corpora to **detect meaning differences**.

In this way, given two corpora, we can understand if they have a significant relative **polarization on a given topic.**



Visualizations from the slides of Paolo Nicoli's Master Thesis

The four tools

```
compare('logic')
```

	1	2	3	4	5
analytic	mathematics	semantics	modal	quantified	syntax
aristotle	oratorical	professors	architectonic	selection	episteme
empiricism	criticism	metaphysics	mathematics	ethics	metaphysical
idealism	method	theory	ethics	physics	metaphysic
nihilism	causality	logical	criterion	concepts	interpretation
plato	electorate	paradigms	sophistication	resemblances	commitment
rationalism	philosophic	medicine	mechanics	fundamental	morals
wikipedia	theory	mathematical	theorem	differential	geometry

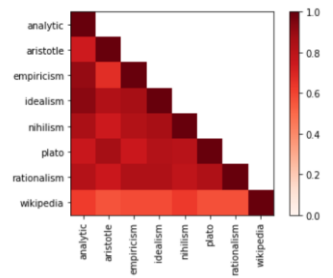
Compare

Compares the embedding of the same word in the 8 different slices.

Similarity

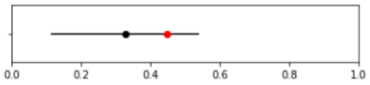
Plots an heatmap of the similarities of the embeddings of the same word across the 8 slices.

```
similarities('knowledge')
```




```
neutral('thinking', 'remembering', 'plato')
```

```
(0.113,0.541)
0.447
```



Neutrality

Defines if an embedding is significatively more similar to another embedding in the same slice with respect to the other slices.

Equivalence

Given a word and one slice, finds the equivalent words in the other 7 slices

```
equivalence('god', 'aristotle')
```

	1	2	3	4	5
analytic	god	divine	cassio	desdemona	goodness
empiricism	maker	god	providence	gods	goodness
idealism	god	trinity	wisdom	revelation	divine
nihilism	god	lord	knight	divine	truly
plato	god	wise	wisdom	supreme	honor
rationalism	wisdom	providence	blessed	goodness	justice
wikipedia	grace	jesus	shiva	titular	god

Results

Conclusion and future developments

Evaluation

	write : wrote = tell : ? do : did = say : ?	man : men = person : ? child : children = tree : ?	justice : injustice =good : ? right : wrong = happy : ?	hard : harder = good : ? high : higher = wide : ?	man : woman=son : ? father : mother=husband : ?	fire : heat = water : ? sky : air = sea : ?	death : end = birth : ? man : home = citizen : ?
	past tense	plurals	opposites	comparatives	sex	knowledge	similarities
Analytic							
Aristotle							
Empiricism							
Idealism							
Nihilism							
Plato							
Rationalism							
Wikipedia							

Analytics

logic as the foundation for understanding the universe

Most_similar(«logic»):
Analytic: [('mathematics', 0.866),
('**semantics**', 0.767),
('modal', 0.750),
('quantified', 0.738),
('syntax', 0.734),
('**foundations**', 0.725),
('arithmetic', 0.704),
('geometry', 0.703),
('philosophy', 0.700),
('metaphysics', 0.699)]

Most_similar(«mathematics»):
Analytic: [('logic', 0.866),
('**metaphysics**', 0.775),
('arithmetic', 0.762),
('science', 0.757),
('**quantum**', 0.751),
('**physics**', 0.749),
('**foundations**', 0.742),
('classical', 0.736),
('metageometry', 0.710),
('philosophy', 0.707)]

*For analytics a picture of the **universe** can be constructed by expressing facts in the form of atomic propositions and linking them using **logical** operators.*

equivalence('mathematics', 'analytic'):

	1	2	3	4	5	6	7	8	9	10
aristotle	dialectic	geometry	rhetoric	demonstrative	philosophy	theoretical	speculative	politics	truths	optics
empiricism	theology	mathematics	science	philosophy	ethics	metaphysical	metaphysics	sciences	criticism	logic
idealism	mathematics	physics	metaphysics	metaphysic	theology	logic	science	philosophy	theory	metaphysical
nihilism	philosophy	science	metaphysics	ethics	criticism	religion	psychology	optimism	investigation	logic
plato	astronomy	theory	geometry	statesmanship	rhetoric	sufficiently	kingship	sciences	dialectic	calculation
rationalism	philosophy	metaphysics	geometry	interpretation	algebra	theology	physics	truths	principles	medicine
wikipedia	astronomy	mathematics	physics	optics	logic	mechanics	prolog	buddhism	textbook	philosophy

Most_similar(«logic»):
Wikipedia: theory(1),
mathematical (2),
theorem (3), geometry (5)

Equivalent(«logic»):
Wikipedia: prolog(1),
astronomy(2), catholic(5),
buddhism(7)

Analytic, Aristotle and Idealism

on language

Most_similar(«language»):
Analytic: [('**usage**', 0.701),
('**use**', 0.629),
('notation', 0.627),
('grammar', 0.621),
('words', 0.614),
('**semantics**', 0.612),
('game', 0.611),
('technique', 0.608),
('sentences', 0.602),
('**learning**', 0.598)]

*Since its beginning, a basic goal of **analytic** philosophy has been **conceptual clarity**, in the name of which Moore and Russell rejected **Hegelianism** for being **obscure.***

Most_similar(«language»):
Idealism: [('speech', 0.763),
('**composition**', 0.762),
('**style**', 0.755),
('grammar', 0.750),
('**latin**', 0.748),
('text', 0.745),
('**treatment**', 0.732),
('languages', 0.731),
('english', 0.731),
('dutch', 0.724)]

*Dialectic in Aristotle is instead simply the technique used to **emerge victorious from an argument**. This success, which does not preclude an actual attainment of truth, comes from **prevailing** with one's own thesis over that held by one's **opponent**, within the framework of premises on which both have agreed before the beginning of the **confrontation**.*

Most_similar(«language»):
Aristotle: [('poetry', 0.810),
('**persuasion**', 0.758),
('words', 0.755),
('expression', 0.752),
('word', 0.751),
('**speech**', 0.746),
('**style**', 0.738),
('**rhetoric**', 0.709),
('english', 0.698),
('letters', 0.695)]

Idealism

*Idealism posits that mind or **consciousness** is the fundamental aspect of **reality**. It suggests that the external world, including objects and events, is either created or shaped by the mind or is inseparable from mental concepts.*

*For idealist, most fundamentally, **reality is equivalent to mind**, spirit, or consciousness. Depending on the philosopher, reality could be entirely a **mental construct**; or it could be that **ideas are the highest form of reality** or have the greatest claim to being considered "real".*

Most_similar(«consciousness»):
Idealism: [('mind', 0.850),
('thought', 0.804),
('**existence**', 0.790),
('selfconsciousness', 0.787),
('conception', 0.780),
('content', 0.768),
('actuality', 0.767),
('knowledge', 0.759),
('**reality**', 0.749),
('activity', 0.744)]

compare('idea'):

analytic	object	entity	assertion	expression	statement	intuition	proposition	belief	sense	conception
aristotle	attribute	accident	analysis	arrangement	examination	image	addition	intelligence	activity	admission
empiricism	notion	conception	ideas	perception	substance	impression	feeling	consciousness	modification	object
idealism	notion	conception	thought	reality	content	significance	form	essence	truth	consciousness
nihilism	conception	meaning	assumption	phenomenon	notion	antithesis	concept	hypothesis	substance	interpretation
plato	audience	task	importance	opportunity	explanation	advantage	answer	goal	consideration	definition
rationalism	conception	perception	thought	thing	mode	essence	modification	substance	intellect	object
wikipedia	question	thinking	text	thoughts	message	belief	fact	topic	speech	truth

equivalence('idea', 'idealism'):

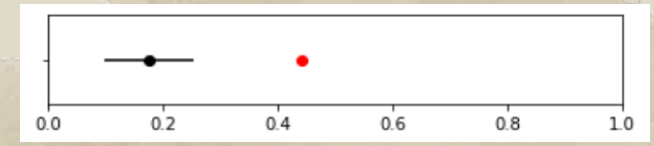
analytic	concept	idea	notion	conception	reality	statement	proposition	form	definition	usage
aristotle	matter	principle	substance	definition	element	demonstration	actuality	explanatory	universe	nature
empiricism	notion	idea	conception	revelation	reality	substance	proposition	deity	matter	essence
nihilism	conception	idea	meaning	ideal	world	phenomenon	reality	reason	concept	antithesis
plato	reason	nature	grasp	knowledge	understanding	constitution	belief	world	name	sign
rationalism	intellect	understanding	idea	conception	knowledge	thought	essence	mind	nature	ideas
wikipedia	symbol	idea	entity	region	principle	constitution	acronym	designation	name	scope

Aristotle , Plato...

Most_similar(«mover»):
Aristotle: [('**unmoved**', 0.740),
('moved', 0.626),
('changing', 0.596),
('motion', 0.546),
('moment', 0.543),
('simultaneously', 0.539),
('**eternally**', 0.534),
('constraint', 0.526),
('potentially', 0.492),
('mediate', 0.487)]

*The **unmoved mover** is a concept advanced from Aristotle as a primary cause (mover) of all the motion in the universe. The unmoved mover moves other things but is not itself moved by any prior action.*

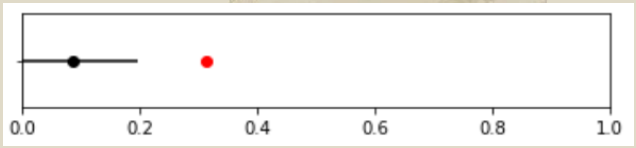
neutral('idea', 'innate'): Plato [(0.099,0.253) ; 0.441]



...and Empiricism

*Historically, empiricism was associated with the **blank slate** concept, according to which the **human mind** is "blank" at birth and develops its thoughts only through later experience.*

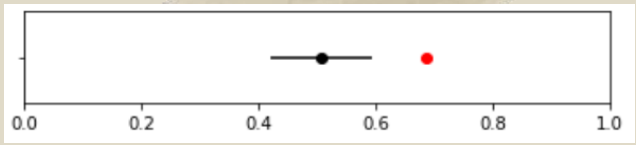
neutral('mind', 'empty'): Empiricism [(0,0.198) ; 0.314]



Most_similar(«consciousness»):
Plato: [('badness', 0.716),
('wickedness', 0.700),
('profit', 0.695),
('**ignorance**', 0.688),
('bad', 0.678), ...)]

*The ability to act justly presupposes, Socratically, the knowledge of what is good. Only this knowledge distinguishes the philosopher as such, since **those who do evil do so out of ignorance**.*

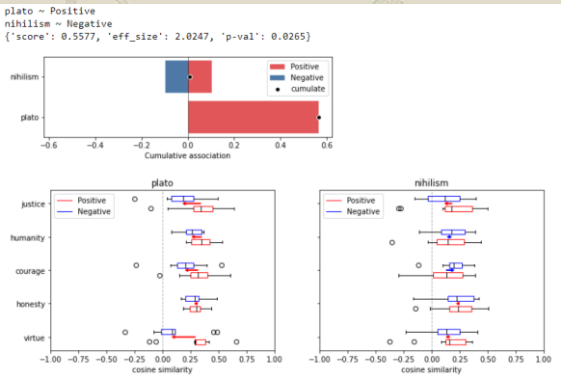
neutral('evil', 'ignorance'): Plato [(0.421,0.594) ; 0.688]

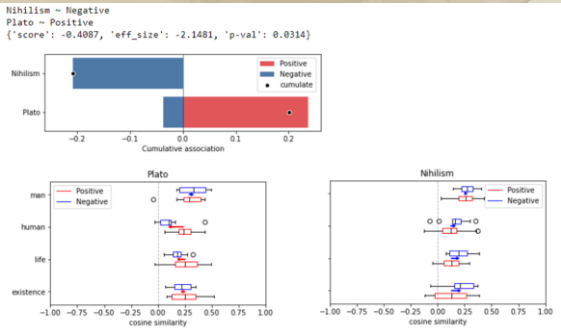


Nihilism and Plato

on virtues and existence

Lemmatized

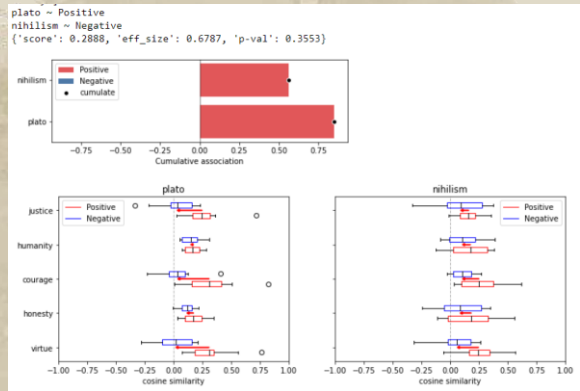


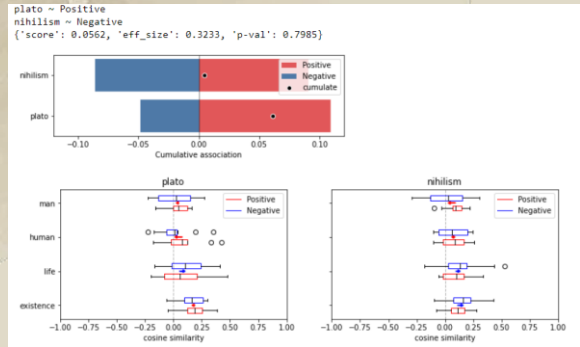


on virtues

on existence

Not Lemmatized





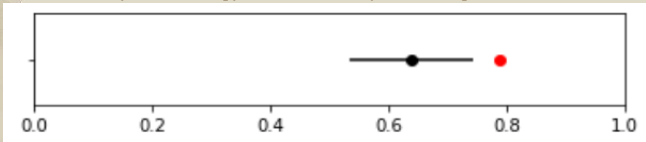
Empiricism and Rationalism

on knowledge

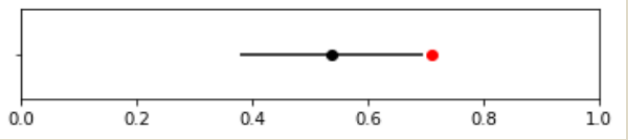
Most_similar(«knowledge»):
[('**experience**', 0.788),
('comprehension', 0.785),
('certainty', 0.785),
('truth', 0.733),
('faculties', 0.708),
('intuitive', 0.700),
('**observation**', 0.697),
('**discovery**', 0.694),
('understanding', 0.673),
('reason', 0.666)]

Most_similar(«knowledge»):
[('understanding', 0.741),
('notion', 0.720),
('**intelligence**', 0.710),
('faculty', 0.697),
('perfection', 0.697),
('idea', 0.695),
('**intellect**', 0.683),
('thought', 0.680),
('certainty', 0.677),
('perception', 0.675)]

neutral('knowledge', 'experience'):
Empiricism [(0.534,0.743) ; 0.788]



neutral('knowledge', 'intelligence'):
Rationalism [(0.380,0.696) ; 0.710]



Rationalists believed that knowledge is in-born and **the intellect**, the inner world of the mind, can **directly grasp logical truths**. The empiricists, instead, believed that knowledge can only be gained through **studying or observing** the physical world outside the mind: through **experiences**.

But also a lot of missed concepts

- Aristotle's «unmoved mover» is not identified as **God**.
- No differences in the embedding of "**ideas**".
  - For Aristotle our mind is empty but we have **innate** capabilities of ordering knowledge into **categories**.
  - According to Kant (german idealism) our experience of the world is mediated by the **a priori categories and concepts** that are inherent in the human mind.
  - For Plato ideas have a lot to do with "**remembering**".
- No interesting results for "**soul**" in Plato or in any other philosophies.

Cosine **similarity**:
(god, cause) = 0,180
(god, mover) = 0,047
(divine, cause) = 0,330

Cosine **similarity**:
(idea, memory) = 0,458
(thinking, remembering) = 0,447
(thought, memory) = 0,351

Conclusions

The algorithm definitely worked, and picked up some **interesting differences in meaning**, but there are also a lot of **missed concepts**.

From one point of view, evaluating the quality of the embeddings with a **qualitative analysis** is **imprecise**. On the other side, a quantitative approach would be nearly impossible considering the subject matter.

The **neutral slice**, even though being one of the biggest, was full of **imprecision**. This could be due to the variety of topics covered by the corpus.

An **expert of the field** would definitely be more capable of understanding what concepts the embeddings actually grasped.

Future Developments

An interesting question to address could related to the corpus. Better embeddings are generated:

- With a **larger, quantity-focused**, corpus (like the one used in this project)
- With a **smaller, quality-focused**, one (like the one used by The Philosophy Data Project), composed of only the books that better represent the thoughts of the philosophical school.

While word embeddings techniques are fundamental for understanding the meaning of single words, **the way humans share theoretical concepts does not rely on continuous repetition**.

«It is generally accepted that in order to learn 'a good vector' for a word, a model must have sufficient examples of its usage. ***This contradicts the fact that humans can guess the meaning of a word from a few occurrences only****. In this paper, we show that a neural language model such as Word2Vec only necessitates minor modifications to its standard architecture to learn new terms from tiny data, using background knowledge from a previously learnt semantic space.»*
«Our main conclusion is that the combination of a ***heightened learning rate*** *and* ***greedy processing*** *results in very reasonable oneshot learning, but that some safeguards must be in place to mitigate the high risks associated with this strategy.»*

High-risk learning: acquiring new word vectors from tiny data

Aurélie Herbelot, Marco Baroni

«Our results confirm that SVD representations are superior to Word2Vec for small data and show that ***Nonce2Vec outperforms Word2Vec*** *and, in most cases, SVD. However, results are* ***currently not accurate enough*** *for providing evidence or new insights to philosophers. Nevertheless, we are hopeful that better results can be obtained in the future by* ***optimizing Nonce2Vec to deal with small rather than tiny data****.»*

Challenging Distributional Models with a Conceptual Network of Philosophical Terms

Y. Oortwijn et. al.

Thanks for
your attention