

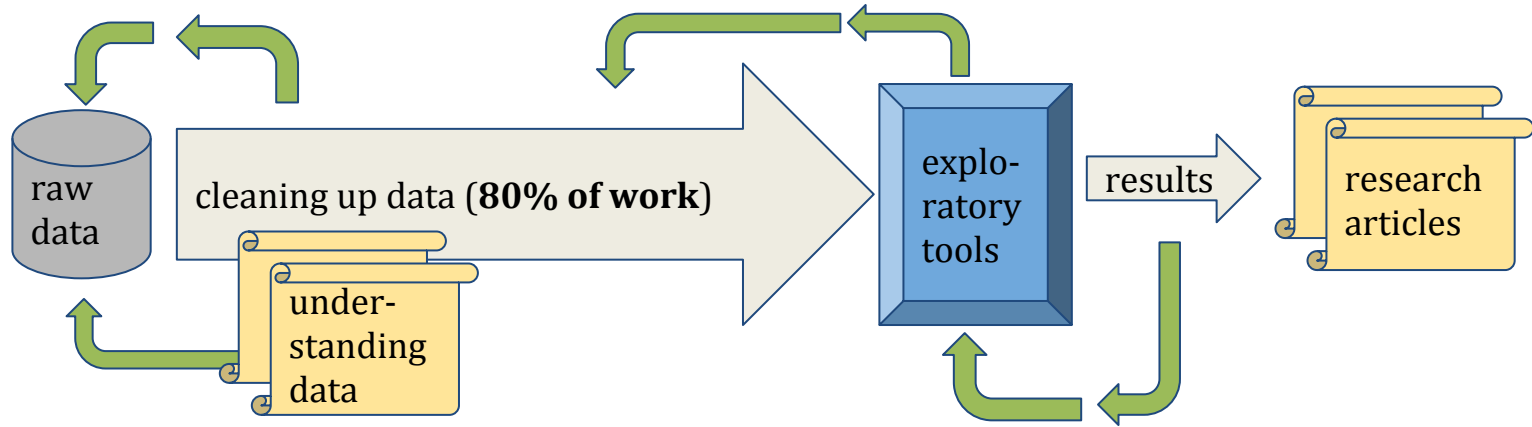
Exploring ECCO: Key moments in 18th-century philosophical literature

Comhis-collective, Helsinki

this presentation:

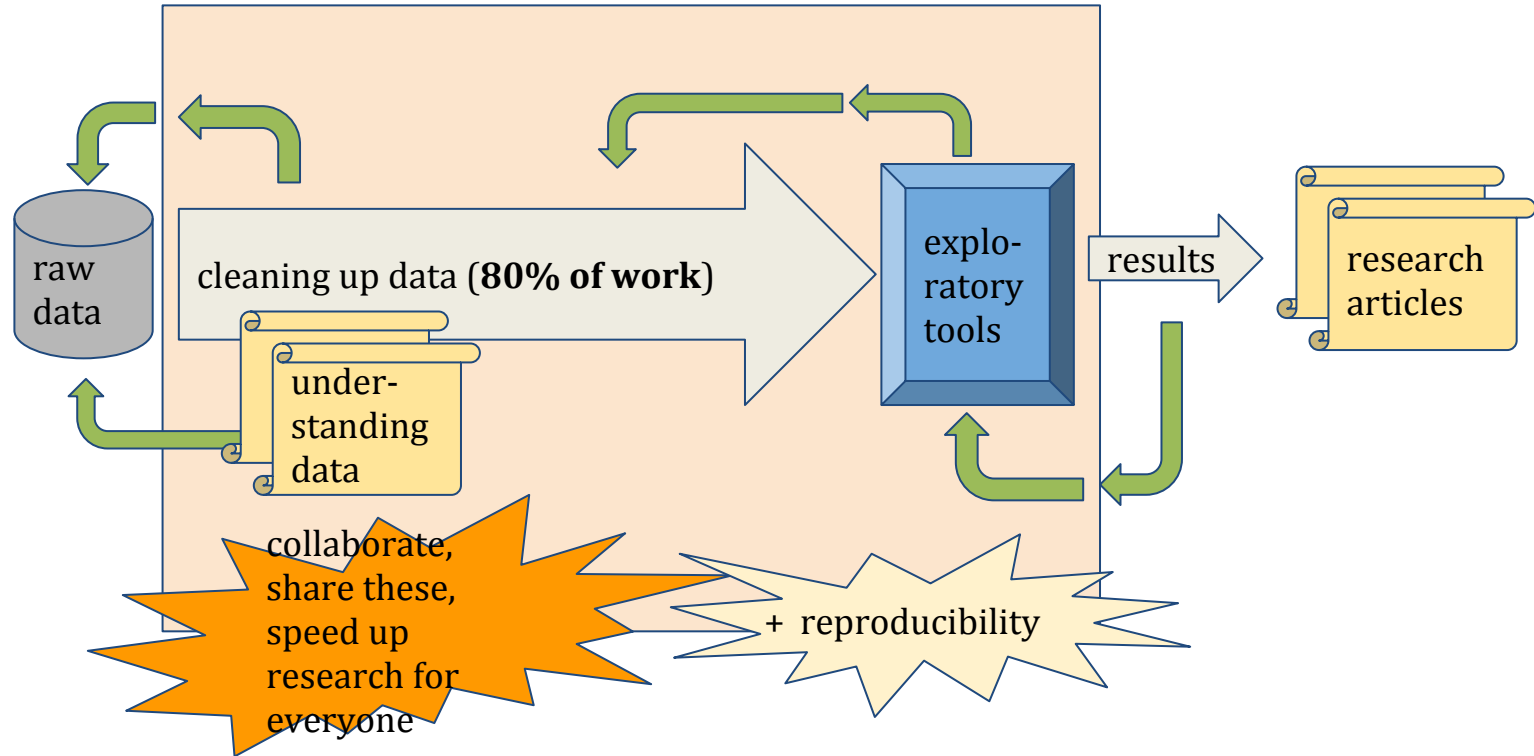
<http://j.mp/comhis-nvm>

Digital humanities research process



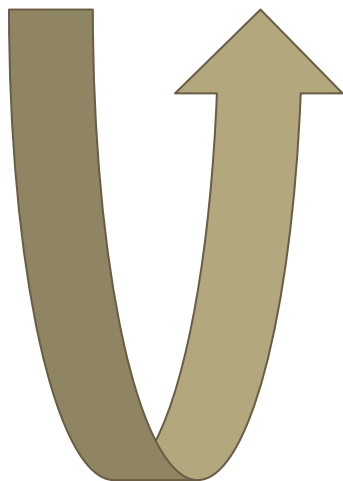
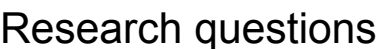
80% of your time for data cleanup, another
80% for algorithms, ...

Leverage collaboration, open science workflows to reduce individual workload



Comhis collective

- Computer scientists researching open workflows, algorithms and interfaces for humanities text and metadata
- Linguists exploring the relationship between words and concepts
- Historians interested in conceptual and actual historical processes



Project goals

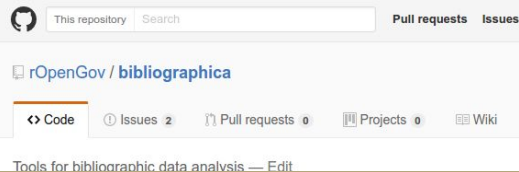
Research publications

Keywords: history publishing, short-title catalogue

How to Cite: Lahti, L., Ilomäki, N. & Tolonen, M., (2015). A Quantitative Study of the History of the English Short-Title Catalogue (ESTC), 1470–1800. *Library Review*, 25(2), pp.87–116. DOI: <http://doi.org/10.18352/lq.10112>

596	62	1
-----	----	---

Public tools, APIs, code



Refined data

publisher	latitude	longitude	publication
printed for J. Cooke; and sold by the booksellers of B...	51.50853	-0.125740	London
re-printed at the Old Post Office in Fishamble street	53.33306	-6.248980	Dublin
Sold at the Bible and Heart, in Cornhill	51.50853	-0.125740	London
Printed by J. Bew	51.50853	-0.125740	London
N/A	37.65996	14.282404	N/A
Printed by J. Evans, Long-lane, West-smithfield, Lon...	51.50853	-0.125740	London
printed for Thomas Ewing, and William Smith	53.33306	-6.248980	Dublin
sold by I. Williams, and all the booksellers in town...	51.50853	-0.125740	London

Evolving set of tools

Data refinement libs/tools

sing fields can be estimated with th

Gornier, Robert

```
polish_dimensions("1/2fo (37 cm)", Notes
:kable(res)
```

Notes

Notes

rings.original width.original

[illegible]

Match	alabel
-------	--------

0	None of the below	
---	-------------------	--

the total page count:

```
(polish_physical_extent(c("50 p.",
```

4	Bellarmino, Roberto	Italian Jesu
	Francesco Romolo, 1542-1621	Roberto Fr

APIs

```
"docFreq" : 6322,
"totalTermFreq" : 50,
"collocations" : {
  "conversation" : 0.00015,
  "politeness" : 0.0001,
  "style" : 0.00107,
  "eloquence" : 0.0001,
  "obliging" : 0.0001,
  "genteel" : 0.0001,
  "gallantry" : 0.0001,
  "breeding" : 0.0001,
  "polite" : 0.0034,
  "elegance" : 0.0001
```

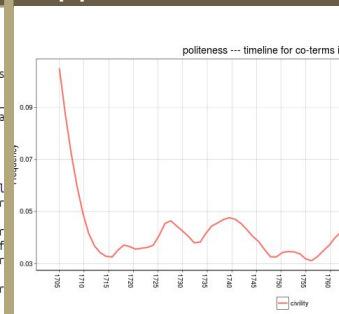
Analysis libs/tools

```

21   return(publications_yearly)
22 }
23
24
25 get_relative_hits_yearly <- function(hits_subsets,
26   averages_yearly <- hits_subset_yearly
27   averages_yearly["total_titles"] <- hits_all_
28   averages_yearly["frequency"] <- averages_yearly
29   return(averages_yearly)
30 }
31
32
33
34 get_query_set_with_freqs <- function(query_filenames,
35   pubs_yearly
36   filter_list <- list(id = get_query_ids_df,
37   idfiltered_data <- get_filtered_data(filter_list,
38   query_titles_yearly <- get_publications_yearly
39   query_titles_yearly_averages <-
40   get_relative_hits_yearly(query_titles_yearly_averages,
41   query_filename <- basename(query_file)

```

Applications



Data sources

ESTC

ECCO

EEBO

BLN

...

CERL

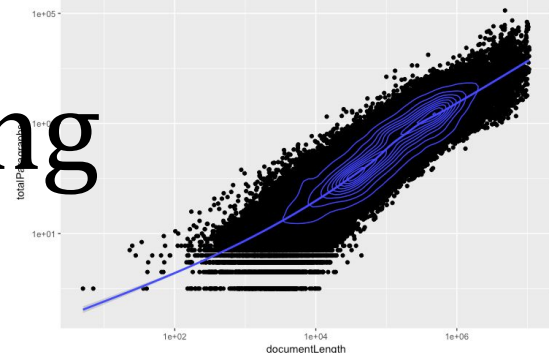
...

DFN

KUNGLIGA

FENNICA

Levels of processing

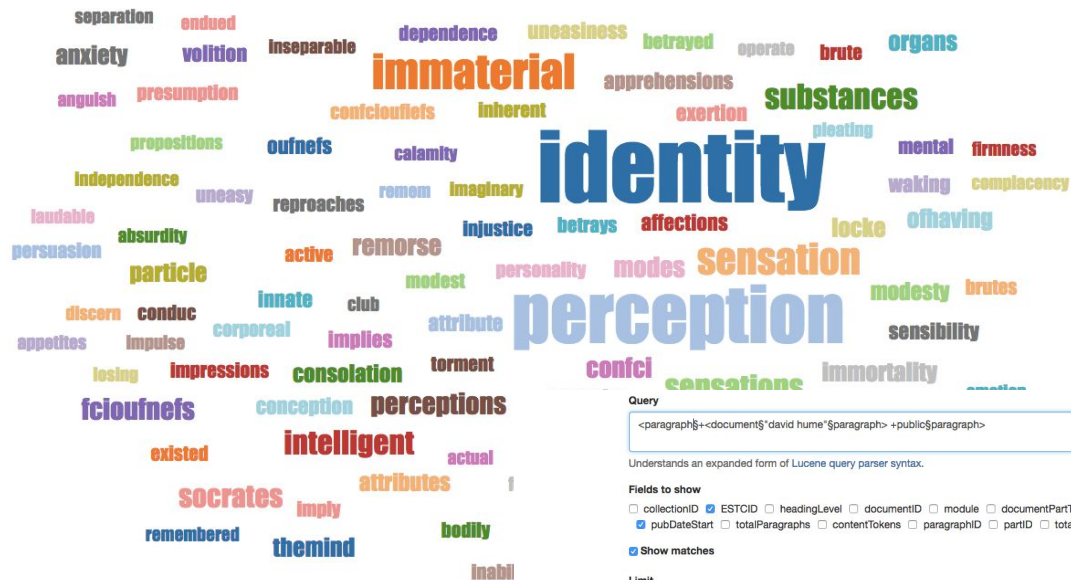


- Paragraph
 - “Uniform-size” chunk of highly related content
 - ECCO2 does not contain paragraph segmentation!
- Section
 - Non-uniform, but useful for constraining, e.g. PREFACE
- Document part
 - TITLEPAGE, FRONTMATTER, BODY, BACKMATTER, TOC, INDEX
- Complete Work
 - Not uniform size (unless so constrained)

Subcorpora formulation

“Human nature” ->

“Human nature” or “frailty” or “misanthropy” or “pravity” in 1) books between 100 and 200 pages long, that do not have the word “sermon” or “christ” in their PREFACE or TITLE and that also contain the word “philosophy”, against 2) books that DO have the word christ or sermon in their PREFACE or TITLE



Query

<paragraph>+<document>\$david hume\$paragraph> +public\$paragraph>

Understands an expanded form of Lucene query parser syntax.

Fields to show

☐ collectionID ☒ ESTCID ☐ headingLevel ☐ documentID ☐ module ☐ documentPartType ☐ language ☒ pubDateStart ☐ totalParagraphs ☐ contentTokens ☐ paragraphID ☐ partID ☐ totalPages ☐ pubDateEnd

☒ Show matches

Limit

20

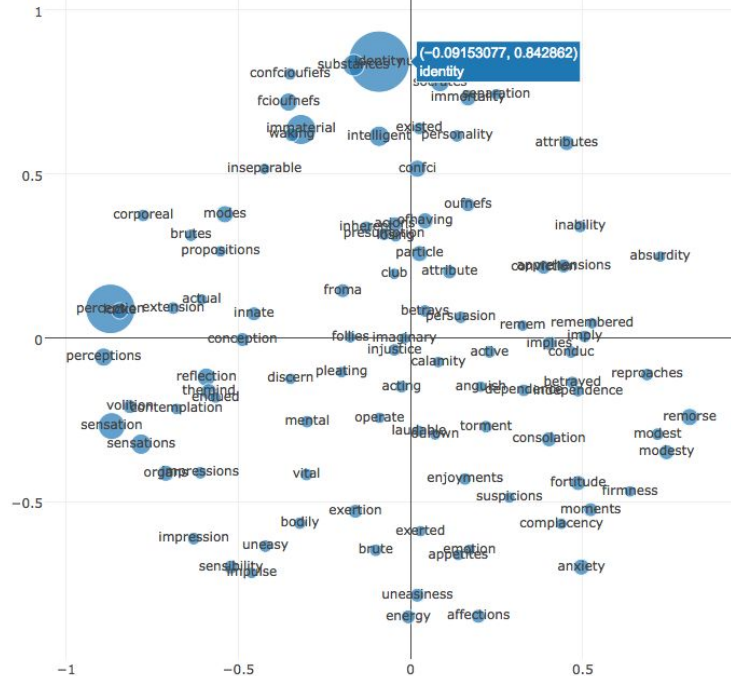
-1 for no limit

Further options

Query

First 20 of 49085 Results

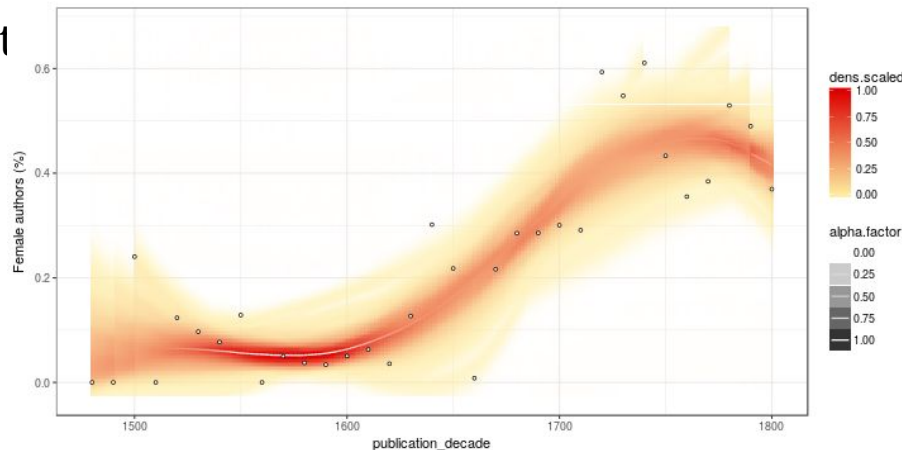
ESTCID	fullTitle	pubDateStart	pubDateEnd
P001934	The European magazine, and London review; containing the literature, history, politics, arts, manners and amusements of the age. By the Philological Society of London	17820000	18269999
Thornton presented a petition from the Bank of England against the Bill, as aviolation of public faith, an infringement of private, right, and establishing a dangerous precedent.			
Fox said, he had two different objections to the Bill, the first of which was, that it was unjust to the public creditors at large, and weakening to the vital strength of public credit; the second, that it was unjust to the Bank, as a trading company.			
-however, That the fact of such property was the indisputable right of the Bank, and that the revenue arising therefrom was as much the property of the Bank, as the principal was the property of the public creditor.			
He next proceeded to argue in support of his first objection to the Bill, namely, the injury it would occasion to public credit, and the injustice of it with respect to the public creditor.			
The contrast between the Public Creditor and Government dated particularly when, how, and where their Dividends were to be paid, and the Bank by that contrast was made trustee; this Bill however would break that contrast, and take from the Bank the trust before reposed in them; and it would bide to say that a better security was given. for even were a better security given, the contract ought not to be deviated from, unless with the consent of all parties.			
The Banker again said was the trustee for the public, and could not, without a breach of public faith, have the trust taken away.			



Research opportunities provided by metadata alone

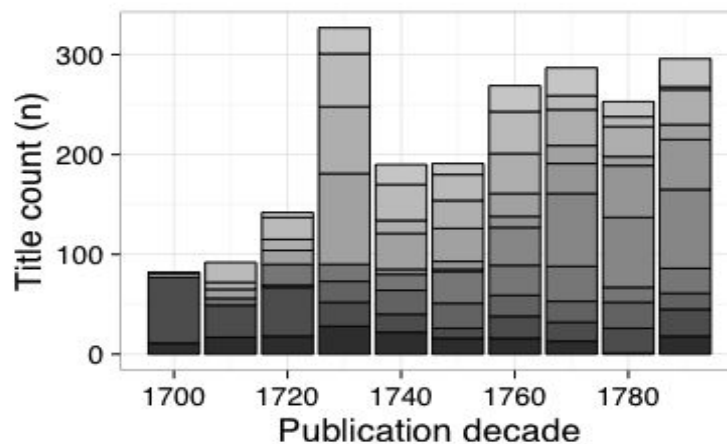
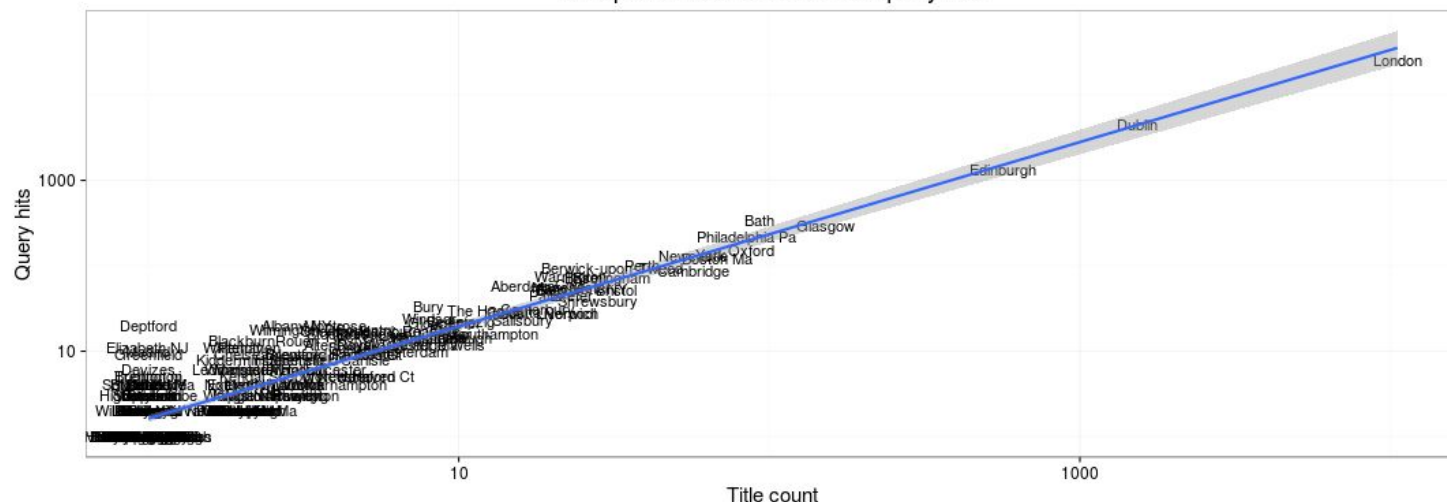
- Quantitative frame for qualitative research
- Understanding knowledge production
- Beyond counting titles (facts about volume)
- Analysis of "history", "philosophy", "religion" etc.
- Publishers and their networks, visualization
- Publication places, cultural transfer
- Individual authors and comparisons
- Histoire Croisée

Example: proportion of female authors over time



<https://github.com/rOpenGov/estc/blob/master/inst/examples/gender.md>

Comparison of title count vs. query hits



Author

- Thomson, James (1700–1748)
- Swift, Jonathan (1667–1745)
- Shakespeare, William (1564–1616)
- Pope, Alexander (1688–1744)
- Johnson, Samuel (1709–1784)
- Goldsmith, Oliver (1728–1774)
- Gay, John (1685–1732)
- Fielding, Henry (1707–1754)
- Defoe, Daniel (1661–1731)
- Cibber, Colley (1671–1757)

Cleaning up data

- 80 % of statistical analysis is tidying up of the data. Often neglected yet implicitly assumed by many tools.

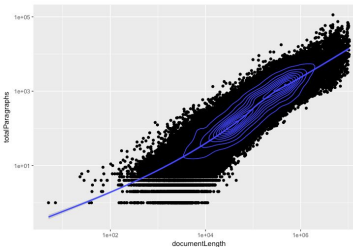
The missing fields can be estimated with the 'fill' argument:

```
res <- polish_dimensions("1/2fo (37 cm)", fill = TRUE)
knitr::kable(res)
```

gatherings.original	width.original	height.original	obl.original	original
2fo	NA	37	FALSE	2fo 37 cm

Estimate the total page count:

```
unlist(polish_physical_extent(cc('
```



Chastelier, Jean
Mersenne, Marin
Bredeau, Claude
Peiresc, Nicolas-Claude Fabri de
Cornier, Robert
Mydorge, Claude
Stanihurst, Henry de
Lefebvre
Unidentified sender

Persons		Date range	Plac
Cornier, Robert		1625-1628	Rou
Notes			
Match	alabel		
0	None of the below		
1	Cornier, Robert, fl. 1625-1628		
2	Corker, Robert (fl. 1700)		
3	Ball, Robert, fl. 1634-1691		
4	Bellarmino, Roberto		
	Italian Jesuit and a Cardinal, Bellarmine, Robert; Bellarmin; Bellarmino, Roberto Francesco Romolo; Belarminus, Robertus		

ECCO/ESTC page count comparison (n = 135847)

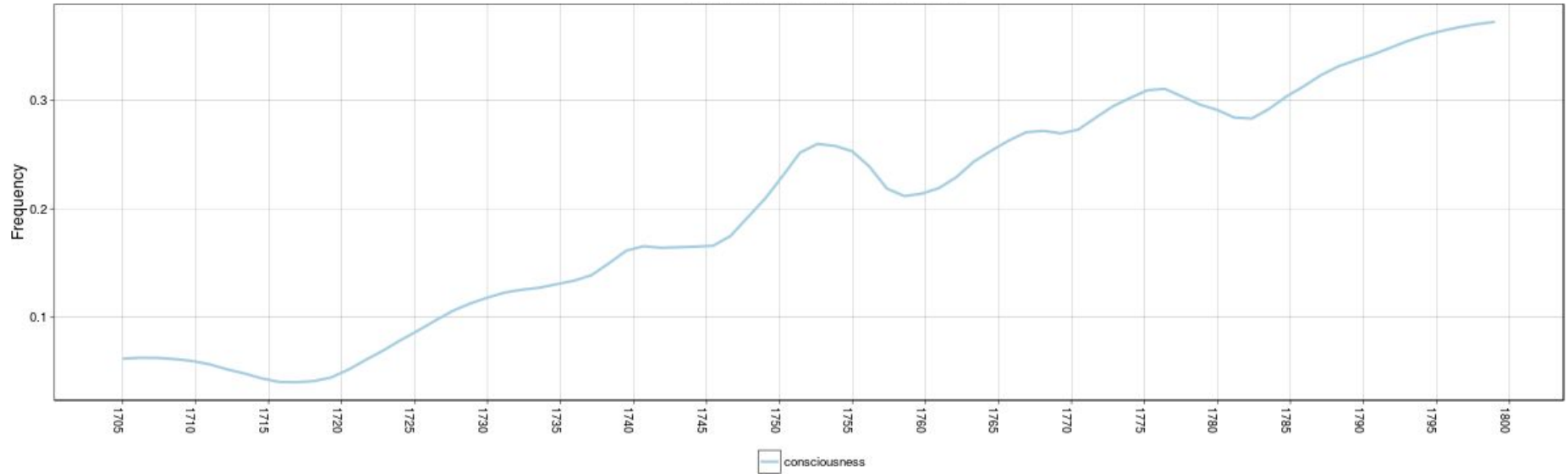


politeness

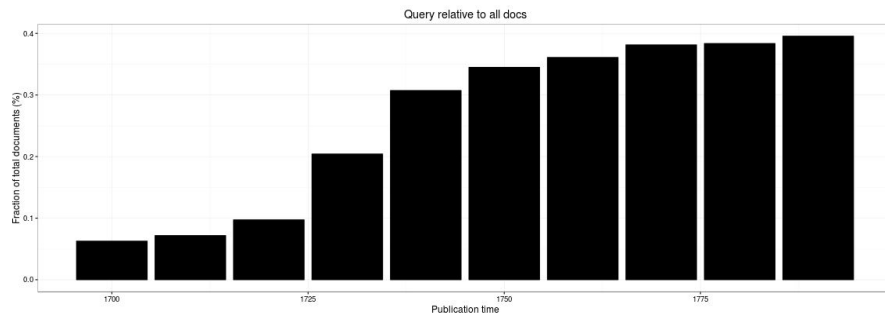
Keywords

- ☒ politeness (1) ☒ politeness (19035) ☒ politeness (1)
- ☒ politenehs (4) ☒ politeness (1) ☒ politeness (3)
- ☒ politeness (1) ☒ politeness (1) ☒ politeness (3)
- ☒ politeneós (1) ☒ politeness (3) ☒ politeness (2)
- ☒ politenehs (3) ☒ politeness (16) ☒ politeness (6)
- ☒ politeness (2) ☒ politeness (1) ☒ politeness (1)
- ☒ politenejs (61) ☒ politenejs (4) ☒ politene'ss (1)
- ☒ politene's (2) ☒ po'iteness (4) ☒ politenehs (1)
- ☒ politenes (1) ☒ politenes (1) ☒ politeness (2)
- ☒ politness (1) ☒ politeness (1) ☒ politeness (2)
- ☒ politeness (49) ☒ politeness (2) ☒ politeness (2)
- ☒ politness (2) ☒ politenehs (2) ☒ politenehs (585)
- ☒ politeness (1) ☒ politenehs (1) ☒ politenehs (1)
- ☒ politenejs (7) ☒ politenehs (4) ☒ politenehs (1)
- ☒ politenejs (1) ☒ politenehs (1) ☒ politenehs (7)
- ☒ politenehs (7) ☒ politenehs (1) ☒ politenehs (1)
- ☒ politenehs (1) ☒ politenehs (1) ☒ politenehs (2)
- ☒ politenehs (1) ☒ politenehs (6) ☒ politenehs (1)
- ☒ politenehs (6) ☒ politenehs (3) ☒ politenehs (10)
- ☒ politenehs (904) ☒ politenehs (1) ☒ politenehs (1)
- ☒ politenehs (4) ☒ politenehs (1) ☒ politenehs (9)
- ☒ politenehs (1) ☒ politenehs (1) ☒ politenehs (1)
- ☒ po.iteness (1) ☒ politenehs (2) ☒ politenehs (1)
- ☒ politenehs (2) ☒ politenehs (159) ☒ politenehs (5)
- ☒ pol'teness (1) ☒ politenehs (3) ☒ politenehs (311)
- ☒ politenehs (3) ☒ politenehs (1) ☒ politenehs (3)

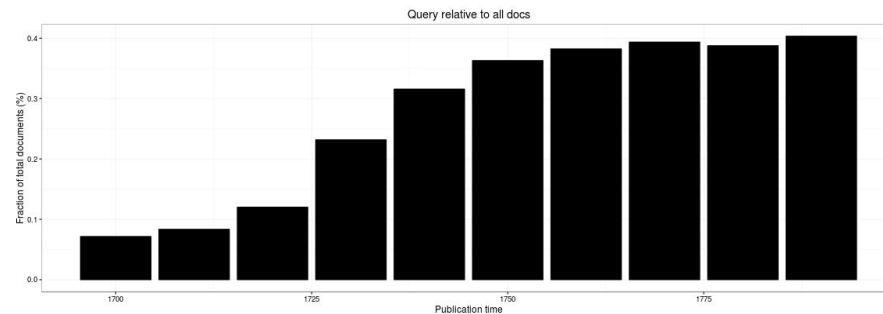
“Consciousness”



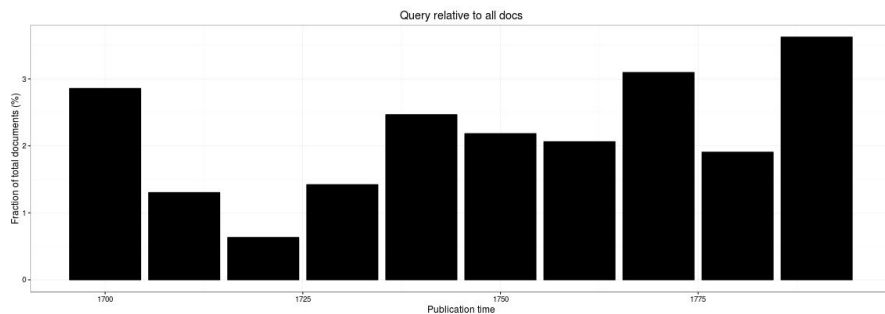
Frequency growth of “consciousness” (Locke’s works excluded)



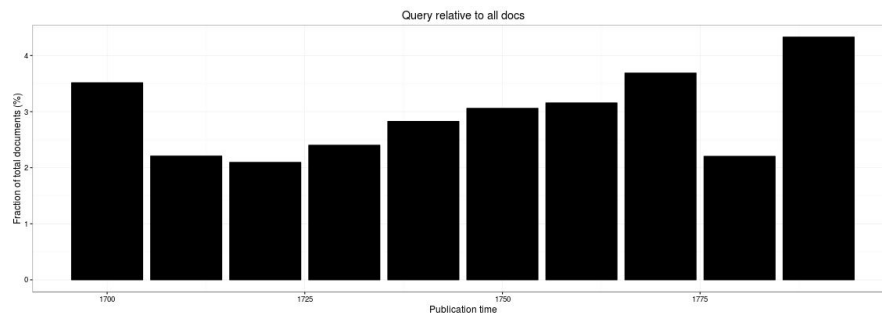
Frequency growth of “consciousness” (Locke’s works included)



“Consciousness” among philosophers, no Locke



“Consciousness” among philosophers, with Locke

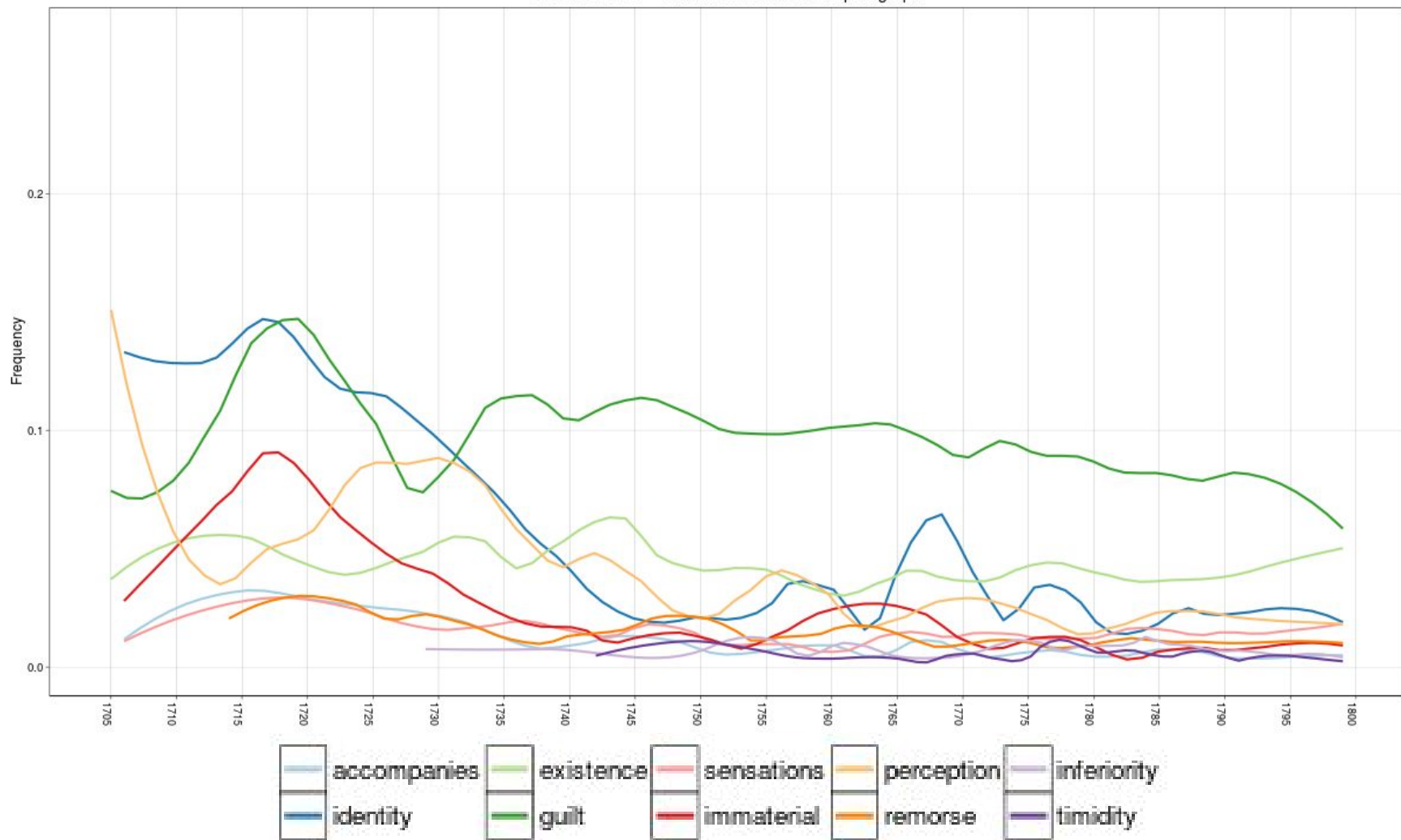


Terms occurring in paragraphs that contain “consciousness”

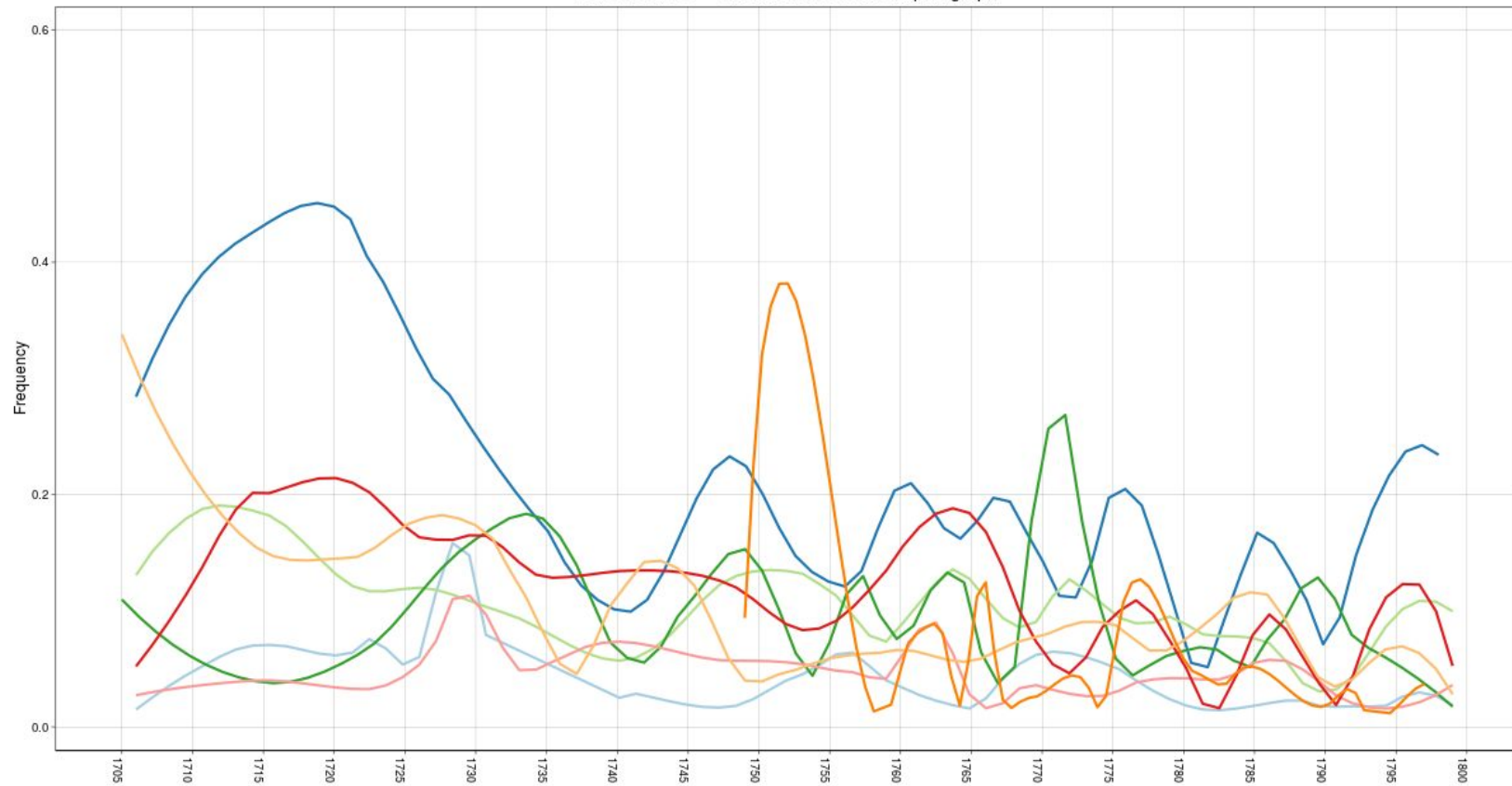
1700-1749	identity	17,09%
	perception	5,62%
	immaterial	5,60%
	perceptions	5,16%
	conscious	4,28%
	waking	2,58%
	subflance	2,34%
	sensitive	2,16%
	forgetfulness	2,01%
	knowledg	1,94%
	atom	1,79%
	dreaming	1,78%
	substances	1,71%
	suppositions	1,69%
	innate	1,63%
	atheist	1,62%
	thinking	1,60%
	remembred	1,53%
	intelligent	1,51%
	atheists	1,46%

1750-1799	identity	7,46%
	perceptions	2,39%
	perception	2,35%
	inferiority	2,02%
	immaterial	1,84%
	conscious	1,82%
	sensations	1,19%
	rectitude	1,08%
	sensation	1,06%
	existence	1,03%
	betrays	0,96%
	guilt	0,94%
	accompanies	0,93%
	exult	0,90%
	affectation	0,89%
	unworthiness	0,88%
	implanted	0,87%
	complacency	0,87%
	oufnefs	0,86%
	inability	0,82%

consciousness --- timeline for co-terms in paragraph



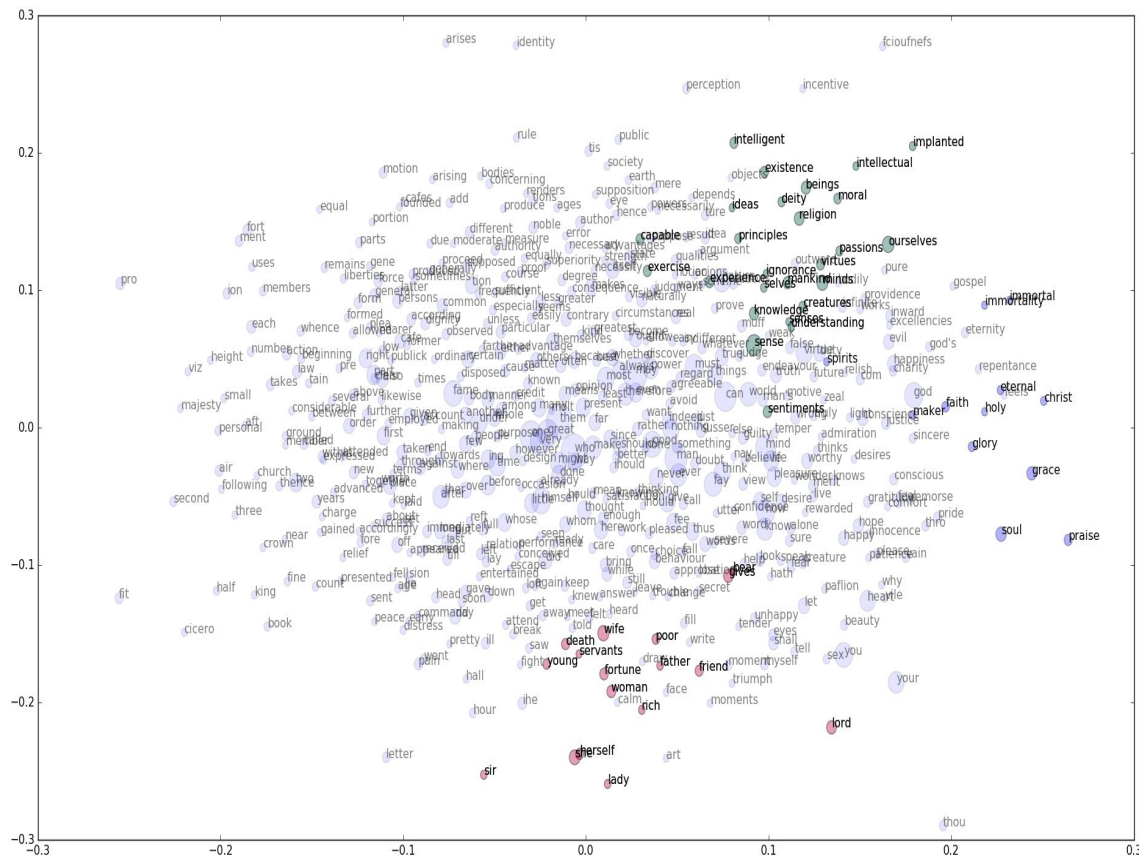
consciousness --- timeline for co-terms in paragraph



Every dot corresponds to a word occurring in the subcorpus: ‘paragraphs containing “consciousness” in 1750’.

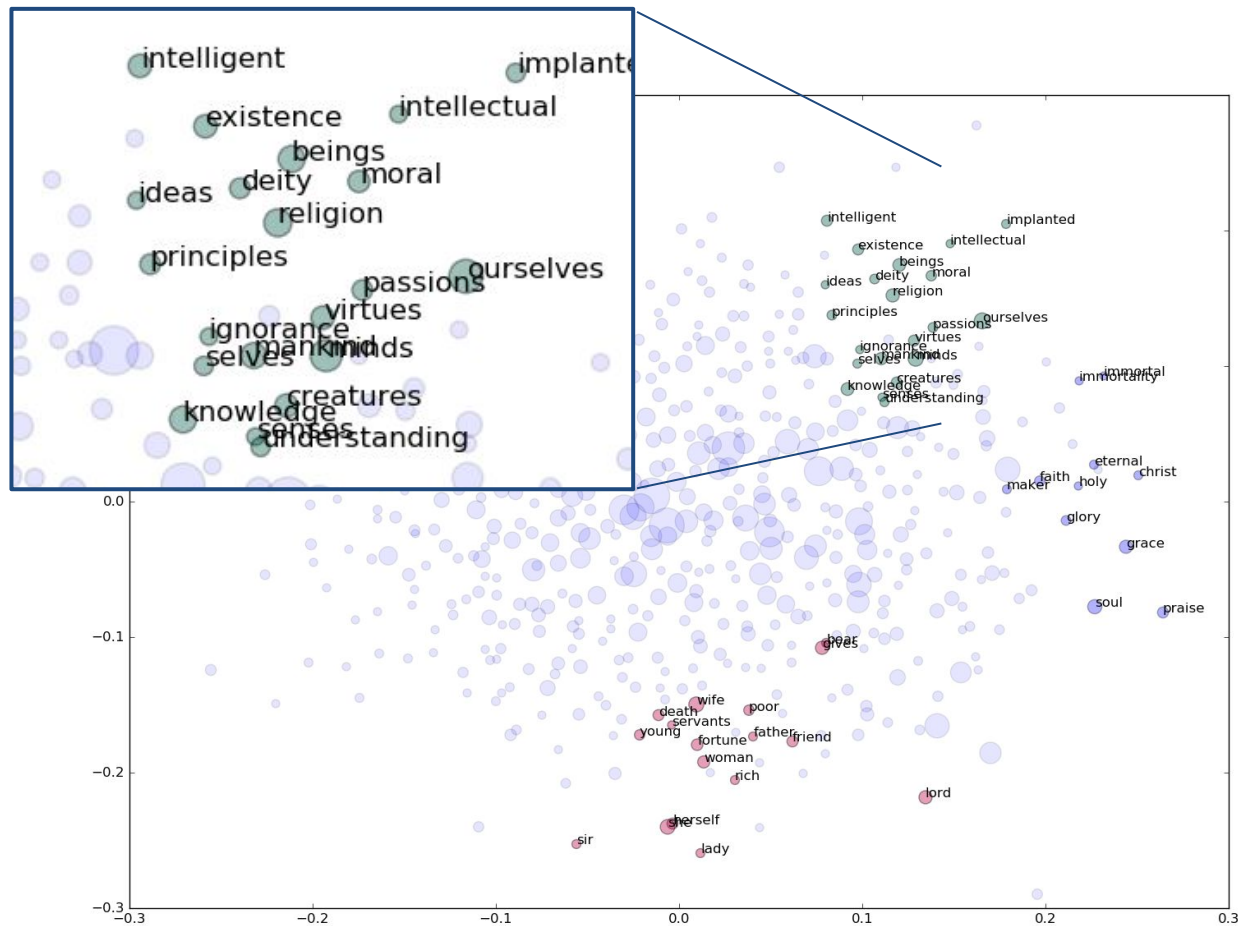
Distance between two dots is an approximation of the similarities between the contexts in which these words occur.

Sizes of the dots indicate relative frequencies of the words in the subcorpus.

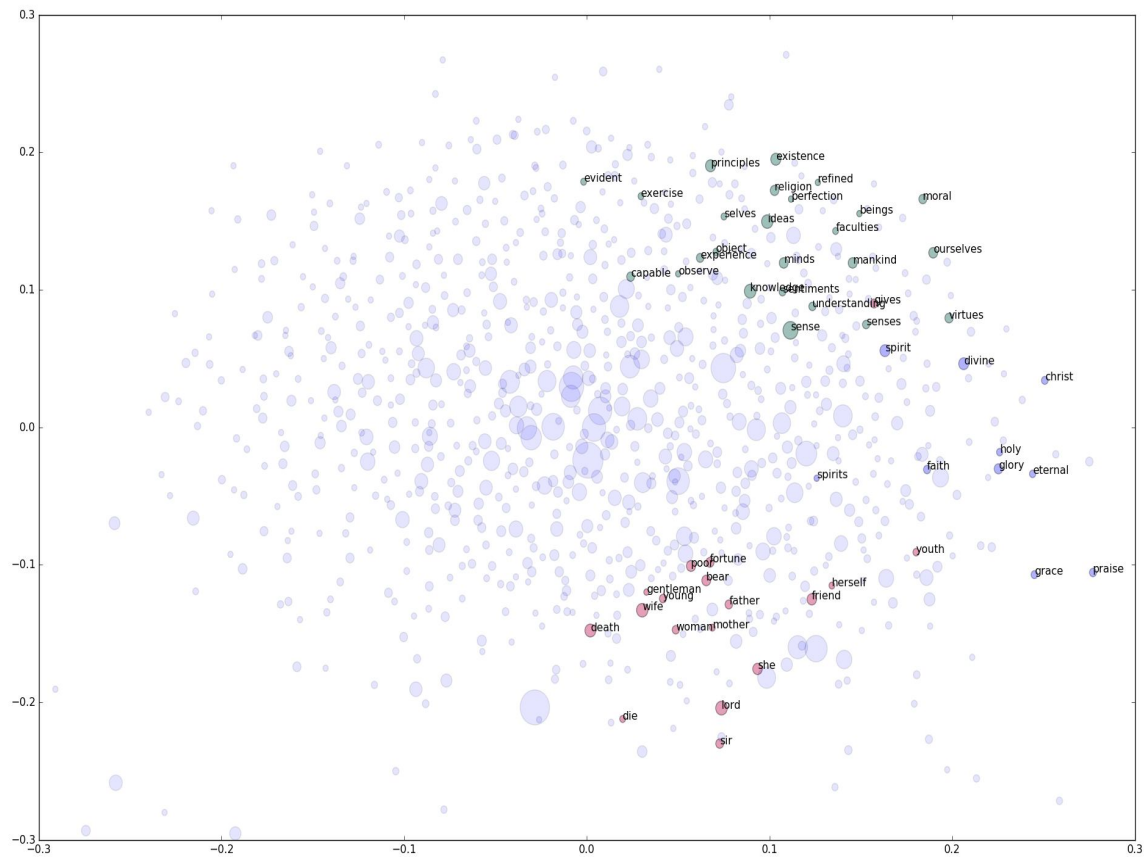


This allows to identify semantic fields.

Tracking the contents and densities of these fields over time can be used to make observations of the contextual changes within the subcorpus.



1780



```

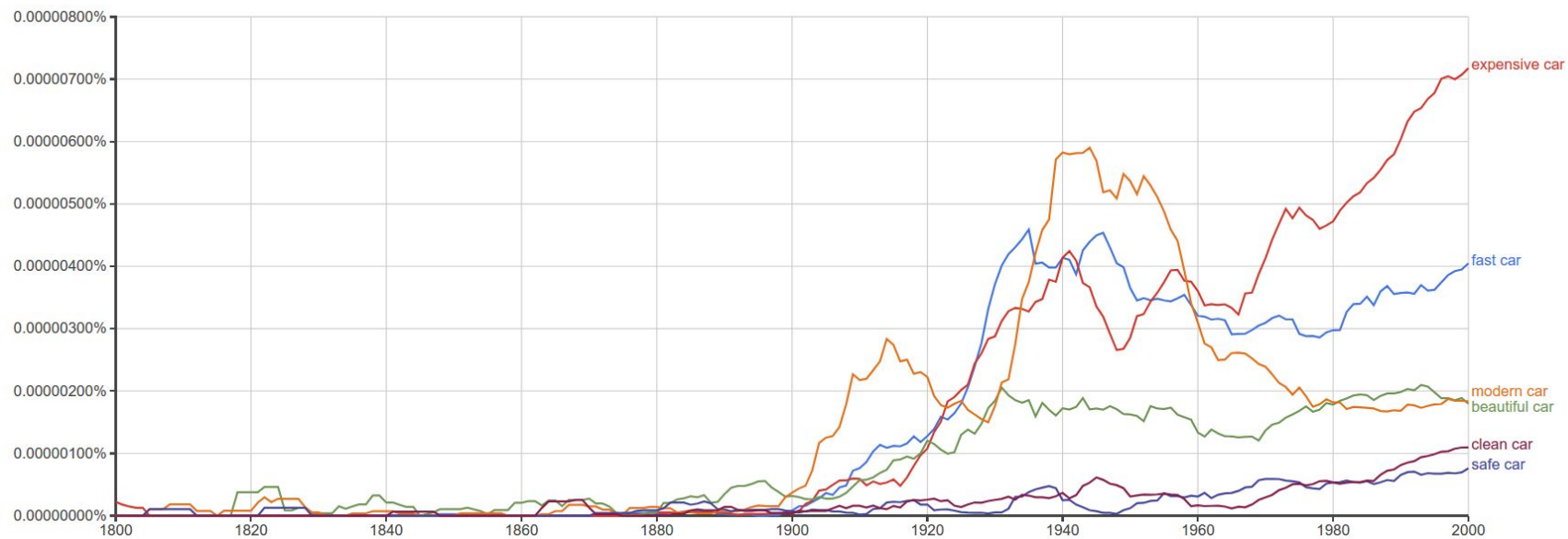
}, {
  "matches" : ".; 7, 40. .I aniw.er, This is. oneMean wifely appointed by God for our RefÃ³r-matnic ; but -hat
whatever it prq;ee4d frqA,it fttpofes nothing tike his Liberty: For, whe-ther it arise in a vicious Man, from his.A
tingcontrary to his :moral Sense or Confcieñc ,: (aSense that God has implanted in every 'Man,iby which we
hneeeffarily approve; good AAtions in <b>ourselves</b> apd others, and disapprove .il-ones) or froam a
<b>Consciousness</b> that henhasinw.arr'd the Resentment, Diapprobataiqñ, .nd-Ãfontempt of others, ( the necessary
Etf& ofS;elfdefence in them ) which no Man can en,:dure the Want, of, even in Thought, he- mustof Consequence be
afham'd, grjev', and haveRemoerf ;'and this Remorse must follo, w.with-put supposing the Agent free in his Sense
ofthe Word. , } ;; 'He concludesi withz faying, \"C ~If thze few4: plain Observations had been adverted .to,,y' Mr.
",
  "pubDateStart" : 17340000,
  "ESTCID" : "T169583",
  "score" : 2
}, {
  "matches" : "not so much from national prejudice, as fromna <b>consciousness</b> of the purity of your
own:intentions in your conduct towards this re-public ;:from having felt a general friendship:to her: cause, and not
perceiving a sufficientground of complaint on which her resent-ment can be founded. But you are not tolearn that
jealousy is one of the strongest and:blindest of the human <b>passions</b>; and I believeyou will be convinced that
the facts hintedat in my Letter, viewed through the mist ofjealousy that had constantly surrounded theleaders in the
revolution, could not fail ofproducing effects similat to what' we nowdeplore.No:!",
  "pubDateStart" : 17990000,
  "ESTCID" : "N055117",
  "score" : 2
}, {
  "matches" : "We cannot easily divest ourselves of the <b>consciousness</b> of our native dignity to such a
degree, asnot to perceive upon the slightest refleaión, that those fatis-fadions which the brute may enjoy as
perfectly as the man, andthe ideot, as those of the best capacities, cannot be worthy ofan eager and intense
perfruit.The transitory nature of most kinds of pleasure is anothercircumstance that ought to moderate our attachment
to them:their duration is, in general, to short, as to bear no proportioñto the time we must spend in order to
obtain them. ... The immoderate perfruit of themmust be a toil and fatigue, instead of a pleasure ; and, if we arenót
sparing in the use of them, they must very soon becomefamiliar, unanimated, and insipid.Nor can we long receive any
considerable fatisfa&ion fromthe most exquisite entertainments of a refined and elegant taute,if we devote
<b>ourselves</b> to them merely as matter of amusement,and without a view to some superior and more substantial
utility.Though we are by nature firongly attached to pleasure, yet thehuman mind cannot but resent such an excessive
application toamusements of any k'nd, as renders them the principal businessof life. ",
  "pubDateStart" : 17610000,
  "ESTCID" : "P6032",
  "score" : 2
}, {

```

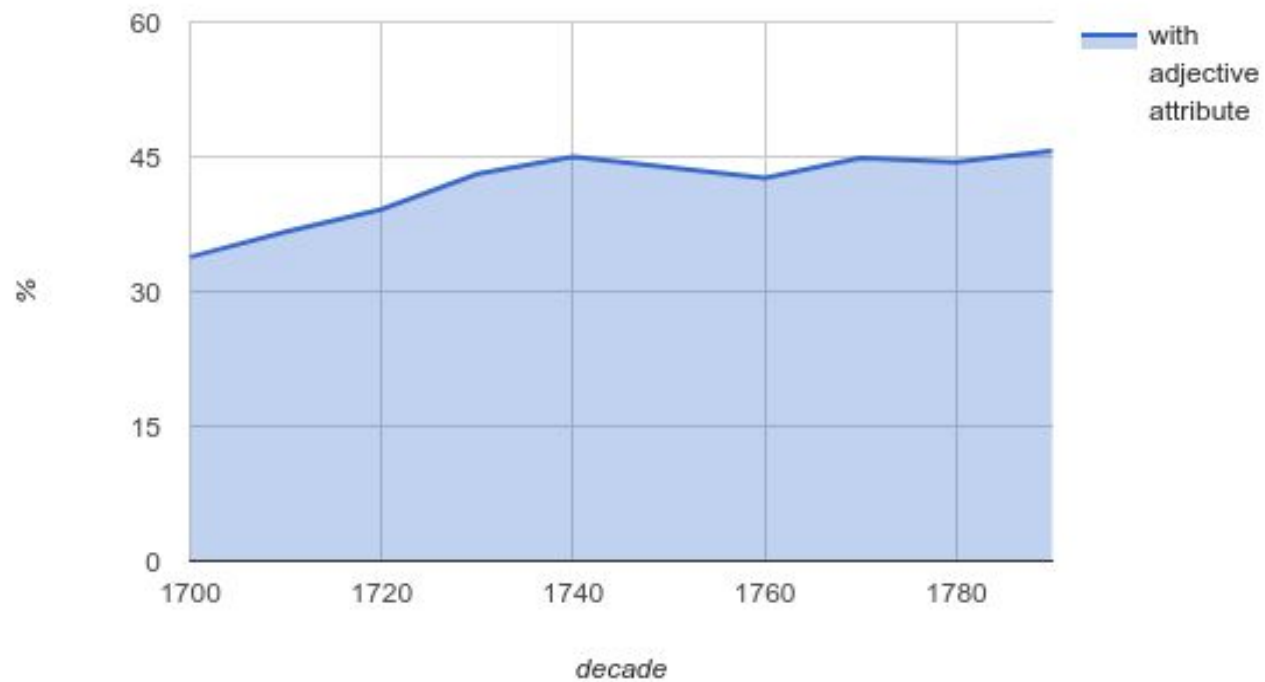
By choosing a set of terms in one of the groups, one gets directly to the actual texts, i.e paragraphs containing “consciousness” and at least one of the chosen co-terms.

Semantic features in evaluation of *philosophy*

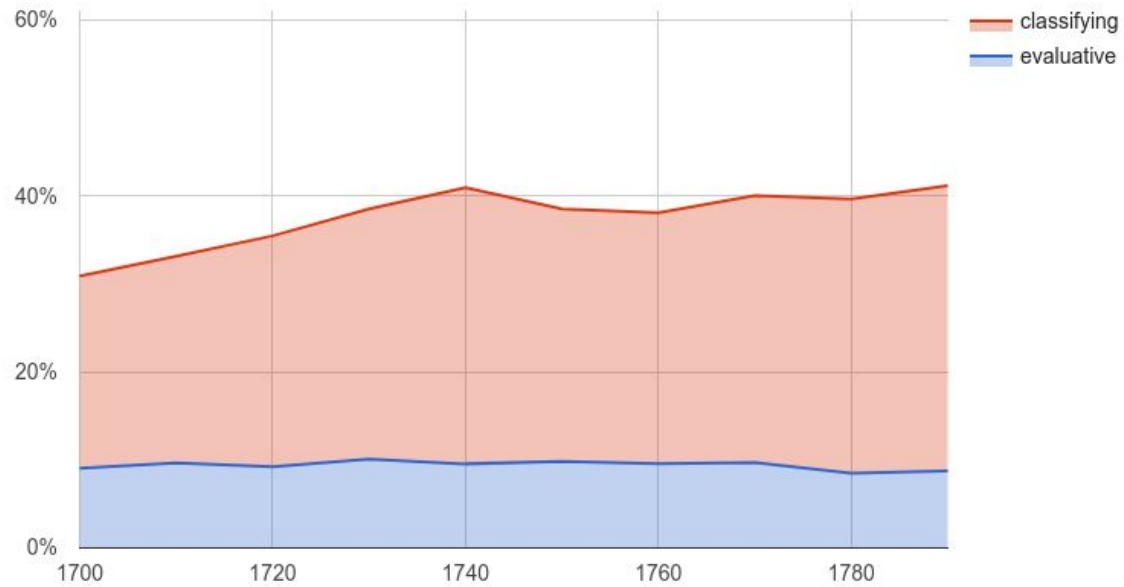
- Fluctuations in meaning of the *philosophy* are studied by looking at semantic features of evaluative adjectives used in its immediate syntactic context.
- List of adjectives attributing *philosophy* are extracted from the corpus.
- These are then grouped by their shared semantic features



PHILOSOPHY with adjective attribute vs decade



PHILOSOPHY with classifying and evaluative adjectives

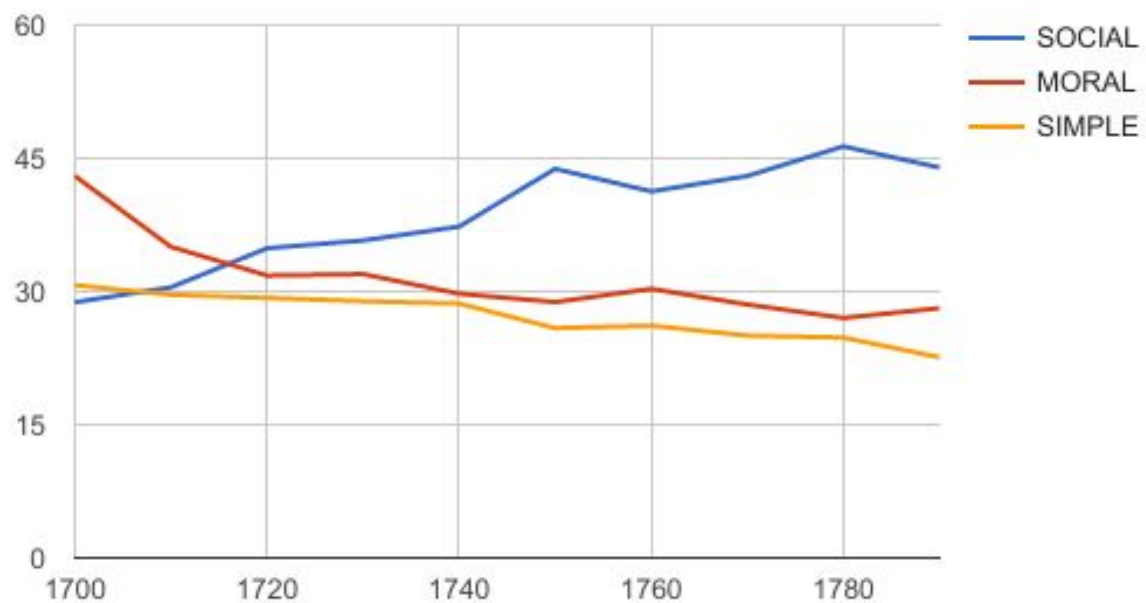


		referenced						simple				
	pos/neg	evaluation	barren	original	friendly	active	status	quality	embodied	true	spiritual	moral
referenced												
evaluation	0.26	1.00	0.48	0.74	0.11	0.66	0.33	-0.73	-0.33	-0.46	-0.37	-0.31
original	0.31	0.74	0.53	1.00	0.45	0.62	0.44	-0.61	0.03	-0.03	-0.71	-0.74
active	0.72	0.66	0.45	0.62	0.53	1.00	0.81	-0.41	-0.02	0.16	-0.26	-0.13
status	0.82	0.33	0.62	0.44	0.66	0.81	1.00	-0.04	0.27	0.53	-0.01	-0.04
barren	0.65	0.48	1.00	0.53	0.53	0.45	0.62	-0.01	-0.05	0.19	-0.27	-0.30
controlled	0.05	-0.14	0.28	-0.29	0.03	-0.09	0.34	0.10	-0.14	0.17	0.66	0.27
friendly	0.76	0.11	0.53	0.45	1.00	0.53	0.66	0.14	0.26	0.76	-0.38	-0.32
pos/neg	1.00	0.26	0.65	0.31	0.76	0.72	0.82	0.27	0.34	0.56	-0.10	0.13
spiritual	-0.10	-0.37	-0.27	-0.71	-0.38	-0.26	-0.01	0.35	0.07	-0.09	1.00	0.83
true	0.56	-0.46	0.19	-0.03	0.76	0.16	0.53	0.50	0.52	1.00	-0.09	-0.05
moral	0.13	-0.31	-0.30	-0.74	-0.32	-0.13	-0.04	0.51	0.19	-0.05	0.83	1.00
embodied	0.34	-0.33	-0.05	0.03	0.26	-0.02	0.27	0.55	1.00	0.52	0.07	0.19
poetic	0.21	-0.47	-0.19	-0.07	0.30	-0.20	0.00	0.65	0.89	0.51	-0.04	0.12
simple quality	0.27	-0.73	-0.01	-0.61	0.14	-0.41	-0.04	1.00	0.55	0.50	0.35	0.51

		referenced						simple						
	pos/neg	evaluation	barren	original	friendly	active	status	quality	embodied	true	spiritual	moral	controlled	poetic
referenced evaluation	0.26	1.00	0.48	0.74	0.11	0.66	0.33	-0.73	-0.33	-0.46	-0.37	-0.31	-0.14	-0.47
original	0.31	0.74	0.53	1.00	0.45	0.62	0.44	-0.61	0.03	-0.03	-0.71	-0.74	-0.29	-0.07
active	0.72	0.66	0.45	0.62	0.53	1.00	0.81	-0.41	-0.02	0.16	-0.26	-0.13	-0.09	-0.20
status	0.82	0.33	0.62	0.44	0.66	0.81	1.00	-0.04	0.27	0.53	-0.01	-0.04	0.34	0.00
barren	0.65	0.48	1.00	0.53	0.53	0.45	0.62	-0.01	-0.05	0.19	-0.27	-0.30	0.28	-0.19
controlled	0.05	-0.14	0.28	-0.29	0.03	-0.09	0.34	0.10	-0.14	0.17	0.66	0.27	1.00	-0.33
friendly	0.76	0.11	0.53	0.45	1.00	0.53	0.66	0.14	0.26	0.76	-0.38	-0.32	0.03	0.30
pos/neg	1.00	0.26	0.65	0.31	0.76	0.72	0.82	0.27	0.34	0.56	-0.10	0.13	0.05	0.21
spiritual	-0.10	-0.37	-0.27	-0.71	-0.38	-0.26	-0.01	0.35	0.07	-0.09	1.00	0.83	0.66	-0.04
true	0.56	-0.46	0.19	-0.03	0.76	0.16	0.53	0.50	0.52	1.00	-0.09	-0.05	0.17	0.51
moral	0.13	-0.31	-0.30	-0.74	-0.32	-0.13	-0.04	0.51	0.19	-0.05	0.83	1.00	0.27	0.12
embodied	0.34	-0.33	-0.05	0.03	0.26	-0.02	0.27	0.55	1.00	0.52	0.07	0.19	-0.14	0.89
poetic	0.21	-0.47	-0.19	-0.07	0.30	-0.20	0.00	0.65	0.89	0.51	-0.04	0.12	-0.33	1.00
simple quality	0.27	-0.73	-0.01	-0.61	0.14	-0.41	-0.04	1.00	0.55	0.50	0.35	0.51	0.10	0.65

		referenced						simple						
	pos/neg	evaluation	barren	original	friendly	active	status	quality	embodied	true	spiritual	moral	controlled	poetic
referenced														
evaluation	0.26	1.00	0.48	0.74	0.11	0.66	0.33	-0.73	-0.33	-0.46	-0.37	-0.31	-0.14	-0.47
original	0.31	0.74	0.53	1.00	0.45	0.62	0.44	-0.61	0.03	-0.03	-0.71	-0.74	-0.29	-0.07
active	0.72	0.66	0.45	0.62	0.53	1.00	0.81	-0.41	-0.02	0.16	-0.26	-0.13	-0.09	-0.20
status	0.82	0.33	0.62	0.44	0.66	0.81	1.00	-0.04	0.27	0.53	-0.01	-0.04	0.34	0.00
barren	0.65	0.48	1.00	0.53	0.53	0.45	0.62	-0.01	-0.05	0.19	-0.27	-0.30	0.28	-0.19
controlled	0.05	-0.14	0.28	-0.29	0.03	-0.09	0.34	0.10	-0.14	0.17	0.66	0.27	1.00	-0.33
friendly	0.76	0.11	0.53	0.45	1.00	0.53	0.66	0.14	0.26	0.76	-0.38	-0.32	0.03	0.30
pos/neg	1.00	0.26	0.65	0.31	0.76	0.72	0.82	0.27	0.34	0.56	-0.10	0.13	0.05	0.21
spiritual	-0.10	-0.37	-0.27	-0.71	-0.38	-0.26	-0.01	0.35	0.07	-0.09	1.00	0.83	0.66	-0.04
true	0.56	-0.46	0.19	-0.03	0.76	0.16	0.53	0.50	0.52	1.00	-0.09	-0.05	0.17	0.51
moral	0.13	-0.31	-0.30	-0.74	-0.32	-0.13	-0.04	0.51	0.19	-0.05	0.83	1.00	0.27	0.12
embodied	0.34	-0.33	-0.05	0.03	0.26	-0.02	0.27	0.55	1.00	0.52	0.07	0.19	-0.14	0.89
poetic	0.21	-0.47	-0.19	-0.07	0.30	-0.20	0.00	0.65	0.89	0.51	-0.04	0.12	-0.33	1.00
simple quality	0.27	-0.73	-0.01	-0.61	0.14	-0.41	-0.04	1.00	0.55	0.50	0.35	0.51	0.10	0.65

Adjectives with SOCIAL, MORAL and SIMPLE features



Text reuse in ECCO

Detecting 70 million clusters of text reuse in ECCO

- Method: NCBI BLAST
 - similarity search of text fragments
 - originally for comparing similarities in biological sequences
 - adapted to historical text corpora by Turku NLP Group
- Type of reuse detected: Syntactic
 - Similar, quotation-like passages of text
 - Somewhat long: 200 characters and more
 - character recognition errors accounted for
 - Some variation in phrasing accounted for
 - Complete rephrasing (semantic similarity) not detected
- Largest “clusters”
 - Fragments of legal text
 - Religious text
 - Quotations from classical authors
 - Lists of book titles (advertisements)

Example: Hume's History of England

Hume, David: The history of Great Britain (1754)

Ormond, who was entirely devoted to him, to fend over considerable bodies of it to England. Most of them continued in his service: **But a small part of them, having foitered** in Ireland a high animosity against the catholics, and hearing the King's party universally reproached with popery, soon after deserted to the parliament. **SOME Irish catholics came over, along with there troops, and joined the King's army,** where they continued the same cruelties and disorders, to which they had been accustomed. The parliament voted, that no quarter, in any action, should ever be granted them: But Prince Rupert, by using some reprisals, soon repressed .this inhumanity. * See farther Cartes Ormond, Vol ii

Ashburton, Charles Alfred: A new and complete history of England (1795)

land. The king ordered Ormond, who was entirely devoted to him, to fend over considerable bodies of it to England. Most of them continued in his service: **but a small part having imbibed** in Ireland a strong animosity against the catholics, and hearing the king's party universally reproached with popery, soon after deserted to the parliament. **Some Irish catholics came over with there troops, and joined the royal army,** where they continued the same cruelties and disorders to which they had been accustomed. The parliament voted, that no quarter, in any action, should be given them: but prince Rupert, by making some reprisals, soon repressed this in

Approaches

- Wide field of possibilities
 - General picture of text reuse in 18th century
 - More focused cases: *single author, work, genre, fragment*
- General research directions
 - Mapping reception, trends
 - Mapping out undocumented sources and influences
 - Mapping publisher networks
 - Well known cases of text reuse: *history, dictionaries*

Tracing impact & reception

- Quantifying “virality” of a publication
 - # of reuse cases with 1, 3, 5, 10, 20, 50, ... years
 - within single publication: reuse cases by chapter
- Test case - Mandeville: Fable of Bees (1714)



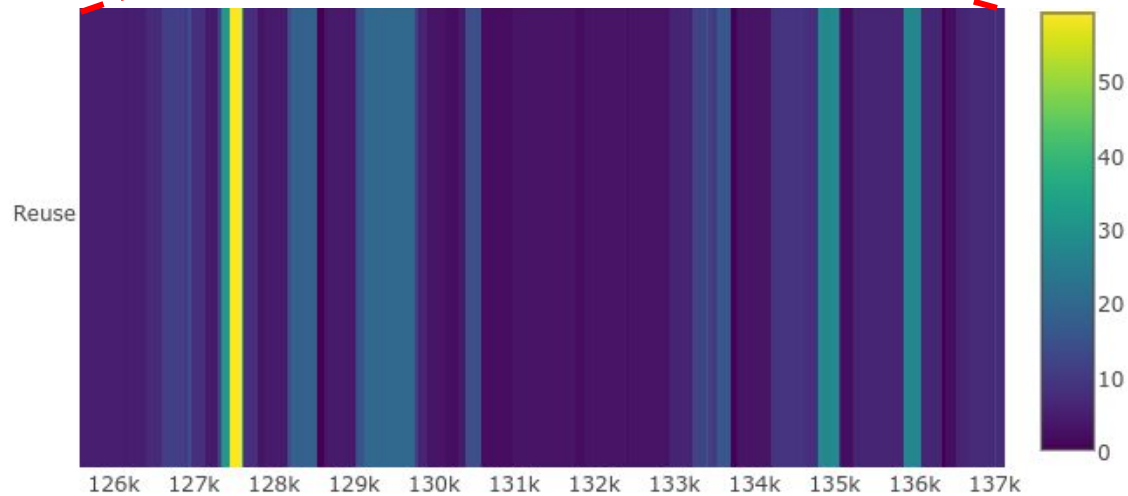
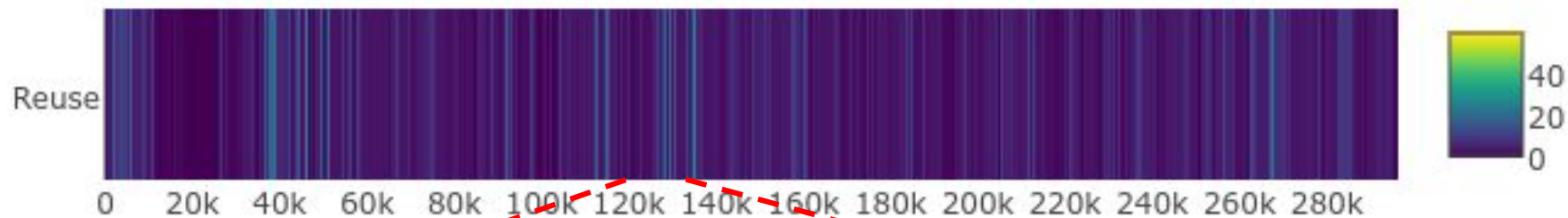
THE PREFACE

LAWS and Government are to the Political Bodies of Civil Societies, what the Vital Spirits and Life it self are to the Natural Bodies of Animated Creatures; and as those that study the Anatomy of Dead Carcasses may see, that the chief Organs and nicest Springs more immediately required to continue the Motion of our Machine, are not hard Bones, strong
A 2 Muscles

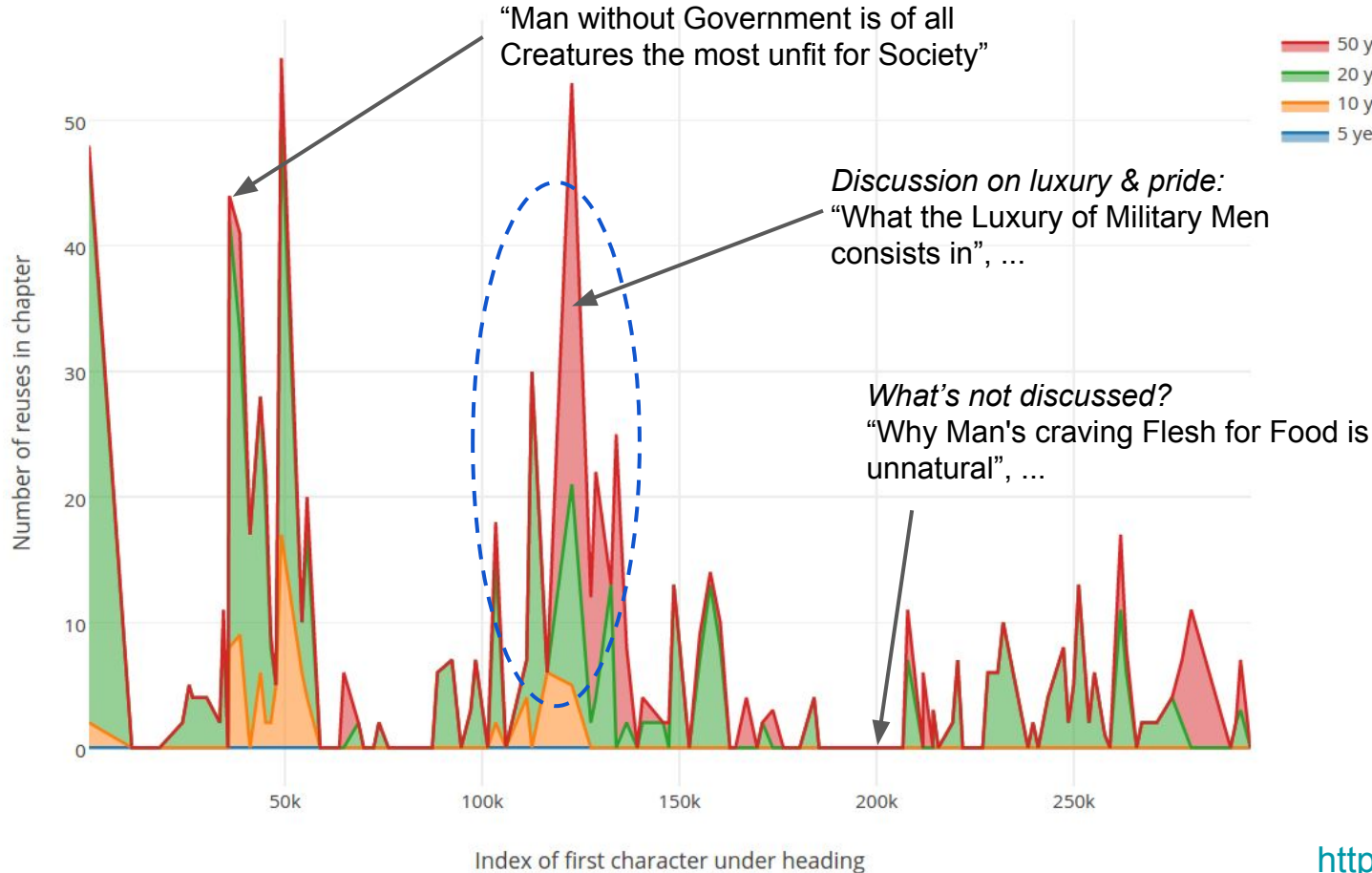
Top references for Fable of Bees

<i>title</i>	<i>author</i>	<i>year</i>	<i>references</i>
The true meaning of The fable of the bees	?	1726	134
An enquiry whether a general practice of virtue tends to the wealth or poverty, benefit or disadvantage of a people?	Blewitt, George	1725	86
Aretē-logia, An enquiry into the original of moral virtue	Campbell, Archibald (1691-1756)	1728	66
A general dictionary, historical and critical	Bayle, Pierre (1647-1706)	1734	51
A short examination of the notions advanc'd in a (late) book, intituled, The fable of the bees or private vices, publick, benefits. By John Thorold Esquire	Thorold, John (1703-1775)	1726	48
Remarks upon a late book, entituled, The fable of the bees, or private vices, publick benefits	Law, William (1686-1761)	1726	36

Mandeville: Fable (1714)



Fable (1714) reuses by chapter heading



CONTENTS.

<i>Man pretends to have for his Species,</i>	153
<i>Why Man’s craving Flesh for Food is unnatural,</i>	ibid.
<i>We ought not to judge of Nature’s design, but from the effects she shews,</i>	155
<i>Man never acknowledges Superiority without Power,</i>	156
<i>The feeling of Brutes proved from several concurring Symptoms,</i>	157
<i>A Definition of Frugality,</i>	158
<i>What the Lavishness or Frugality of Nations depend upon,</i>	159
<i>Maxims to make a People great and flourishing.</i>	162
<i>To make a Society good and honest,</i>	162
<i>The present Grandeur of the Dutch is not owing to the Virtue and Frugality of their Ancestors,</i>	164
<i>The Hardships and Calamities they have suffered</i>	164
<i>Their natural Wants,</i>	165
<i>The Dutch not frugal by Principle</i>	168
<i>’Tis Policy and not Virtue that makes the Dutch encourage Frugality,</i>	169
<i>How they promote Lavishness when it suits with their Interest,</i>	170
<i>What</i>	

David Hume's History of England

- Hume used long unmodified quotes from earlier works
 - Typical practice for historians in the period
 - Hypothesis: Hume considered these passages less important than his own original work
 - Hume's (some presumably previously unknown) sources can be identified
 - Hume's influences can be traced
- Analysis of structure of the work
 - Amount of reused vs. original text can be traced
 - Nature of reused and original text can be compared
 - can clarify Hume's intentions by highlighting his original work
- Applied to other works on history
 - Tracing influences
 - Tracing debates by analyzing text surrounding reused section

Conclusions

- **Tools** - Octavo, a developing ecosystem for studying ECCO
 - web tools, full text search tools, open code repositories, cleaned ESTC metadata ...
- **Organization** - loose collective:
 - <https://comhis.github.io/>
 - Antti Kanner, Leo Lahti, Viivi Lähteenoja, Jani Marjanen, Eetu Mäkelä, Hege Roivainen, Laura Tarkka, Mikko Tolonen, Ville Vaara
 - Partners: Filip Ginter, Hannu Salmi & al. in Turku, Vili Lähteenmäki
- **Research** - Study of key moments of conceptual change in 18th century