

How Does the State Speak of Globalisation? A Quantitative Text-Mining Approach

Justin Murphy

University of Southampton

j.murphy@soton.ac.uk

Abstract

Scholars argue that globalisation is strategically deployed by governments (Hay and Rosamond 2011). This article is the first large-scale quantitative assessment of this argument, using text-mining and machine learning techniques to analyze how the government of the United Kingdom has deployed the term 'globalisation.' Specifically, this article exploits the newly released United Kingdom Government Web Archive to analyze a random sample of more than 60,000 web pages published across the entire UK government web system.

I requested 150,000 web pages, received 67k and about 1k were errors.
Thus, the final sample consists of a corpus of dtm

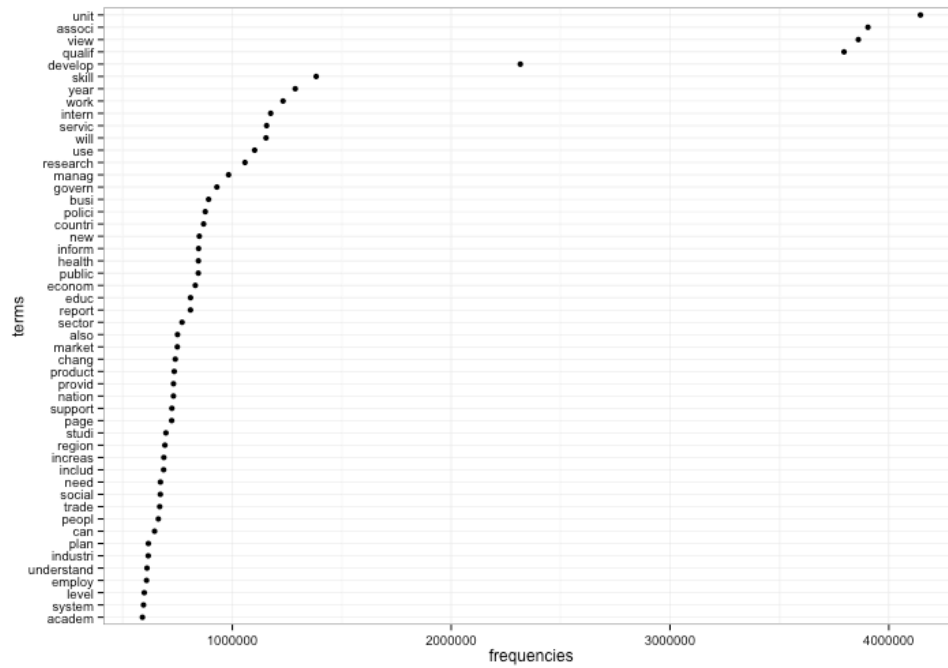


Figure 1: Most frequent terms

Descriptive Statistics

Terms	Globalis
vision	0.9
emphasi	0.87
secondari	0.84
context	0.83
legaci	0.82
implic	0.81
amongst	0.8
educ	0.79
vocat	0.78
perceiv	0.76

Correlated terms

Cluster analysis

[1] Cluster 1

[1] will report can committe increas level offic

[8] also school need invest busi includ compani

[15] govern transport use industri number respons

[1] Cluster 2

[1] journal intern manag technolog system engin comput

[8] titl publish onlin euro print intellig busi

[15] innov enterpris electron scienc sustain price

[1] Cluster 3

[1] health public care journal nhs diseas find

[8] polici author avail articl sourc impact london

[15] research exclud refin similar add clinic

[1] Cluster 4

[1] research skip accesskey esrc social search page

[8] studi public econom news scienc impact societi

[15] contact output fund grant guidanc work

[1] Cluster 5

[1] eldi websit globalis topic home develop countri

[8] inform guid what announc profil publish site

[15] job map event communiti visit new

[1] Cluster 6

```

## [1] unit      qualif      associ      view      understand skill
## [7] provid      design      introduct  perform    listen     write
## [13] sport      leisur      treatment  music      physic     read
## [19] techniqu  wale
## [1] Cluster 7
## [1] eldi      group      develop    space      communiti member
## [7] interest  friend     help       blog       search     opinion
## [13] profil    home       comment    livelihood famili    secur
## [19] name      aid
## [1] Cluster 8
## [1] oper      countri    procur     organis    non        specif    servic
## [8] tel      profit     project    search     india      result    programm
## [15] africa   support    govern     china      fund       aid
## [1] Cluster 9
## [1] academ    year      breakdown  geograph   coverag    articl    region
## [8] pdf      labour     econom     market     review     countri    statist
## [15] filter    economi    trend      releas     book       sep
## [1] "Cluster Membership"
##
##      1      2      3      4      5      6      7      8      9
## 46337  907  2235  5459  3535  2345  2121  779  2696
## [1] "Within cluster sum of squares by cluster"
## [1] 383.976905  0.009232 18.591247 36.403308 28.291523 4.239123
## [7] 14.269600 21.928006 55.741371

```

Appendix

This is the appendix.

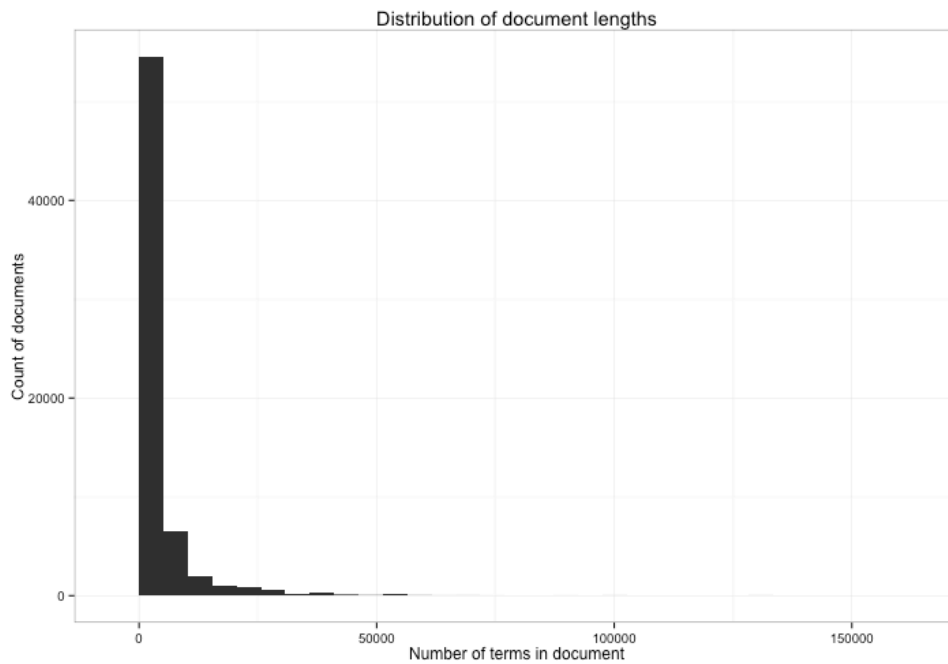


Figure 2: Document lengths

##	Length	Class	Mode
## cluster	66414	-none-	numeric
## centers	17640	-none-	numeric
## totss	1	-none-	numeric
## withinss	9	-none-	numeric
## tot.withinss	1	-none-	numeric
## betweenss	1	-none-	numeric
## size	9	-none-	numeric

```
## iter          1 -none- numeric
## ifault        1 -none- numeric
```

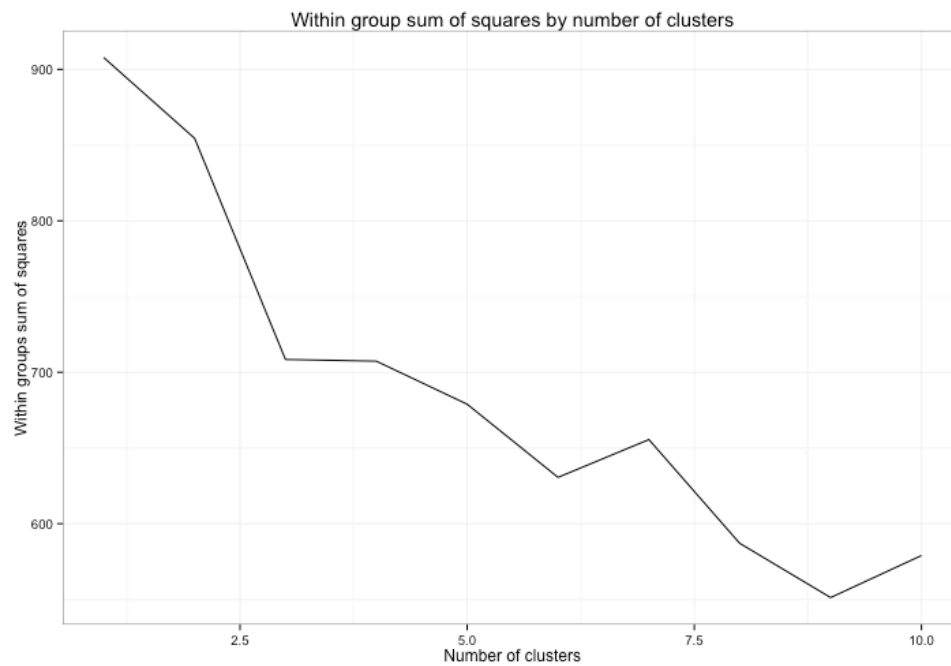


Figure 3: Binwidth

##	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
## [1,]	"skill"	"school"	"globalis"	"financi"	"product"
## [2,]	"develop"	"use"	"websit"	"market"	"per"
## [3,]	"educ"	"educ"	"global"	"intern"	"year"
## [4,]	"globalis"	"teacher"	"eldi"	"globalis"	"countri"
## [5,]	"countri"	"ict"	"polici"	"capit"	"price"
## [6,]	"dfid"	"can"	"new"	"bank"	"world"
## [7,]	"also"	"learn"	"home"	"develop"	"cost"

##	[8,]	"train"	"dfid"	"unit"	"economy"	"market"
##	[9,]	"sector"	"project"	"europa"	"global"	"industry"
##	[10,]	"research"	"student"	"relationship"	"economy"	"export"
##	[11,]	"need"	"technology"	"intern"	"risk"	"increase"
##	[12,]	"culture"	"people"	"business"	"country"	"trade"
##	[13,]	"nation"	"work"	"information"	"increase"	"cent"
##	[14,]	"economy"	"community"	"trade"	"growth"	"import"
##	[15,]	"world"	"activity"	"country"	"system"	"development"
##	[16,]	"global"	"experience"	"world"	"will"	"govern"
##	[17,]	"poverty"	"computer"	"state"	"policy"	"period"
##	[18,]	"women"	"development"	"tel"	"service"	"even"
##	[19,]	"increase"	"also"	"work"	"benefit"	"exchange"
##	[20,]	"private"	"teach"	"security"	"invest"	"however"
##		Topic 6	Topic 7	Topic 8	Topic 9	
##	[1,]	"pay"	"development"	"development"	"development"	
##	[2,]	"plus"	"relation"	"country"	"impact"	
##	[3,]	"fair"	"skill"	"trade"	"research"	
##	[4,]	"intern"	"strategy"	"intern"	"agriculture"	
##	[5,]	"institute"	"level"	"govern"	"increase"	
##	[6,]	"concern"	"govern"	"work"	"service"	
##	[7,]	"commerce"	"globalisation"	"will"	"report"	
##	[8,]	"encourage"	"country"	"policy"	"product"	
##	[9,]	"certain"	"term"	"world"	"rural"	
##	[10,]	"population"	"policy"	"people"	"food"	
##	[11,]	"import"	"nation"	"import"	"use"	


```
## [12,] "cent"      "market"    "poverti"  "also"
## [13,] "constitu" "requir"    "poor"     "may"
## [14,] "assur"     "key"       "globalis" "health"
## [15,] "programm"  "global"    "effect"   "popul"
## [16,] "public"   "educ"      "invest"   "water"
## [17,] "accord"   "two"       "price"    "migrat"
## [18,] "social"   "differ"    "inflat"   "area"
## [19,] "parti"    "state"     "can"      "growth"
## [20,] "term"     "competit" "increas"  "poverti"
```

References

Hay, Colin, and Ben Rosamond. 2011. "Globalization, European Integration and the Discursive Construction of Economic Imperatives." *dx.doi.org.libproxy.temple.edu* 9(2): 147–67. <http://www.tandfonline.com/doi/abs/10.1080/13501760110120192>.