

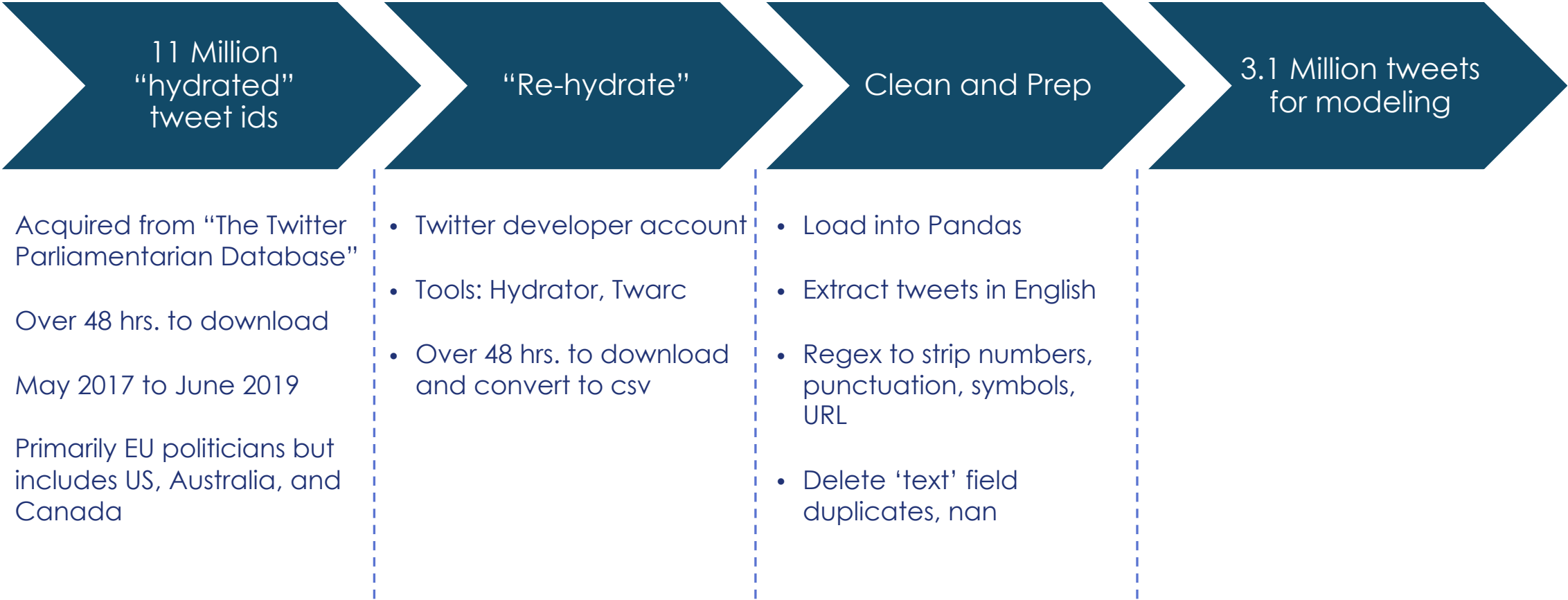


# What are politicians saying on twitter, anything meaningful?

Topic analysis of tweets by mostly EU  
politicians 2017 to 2019

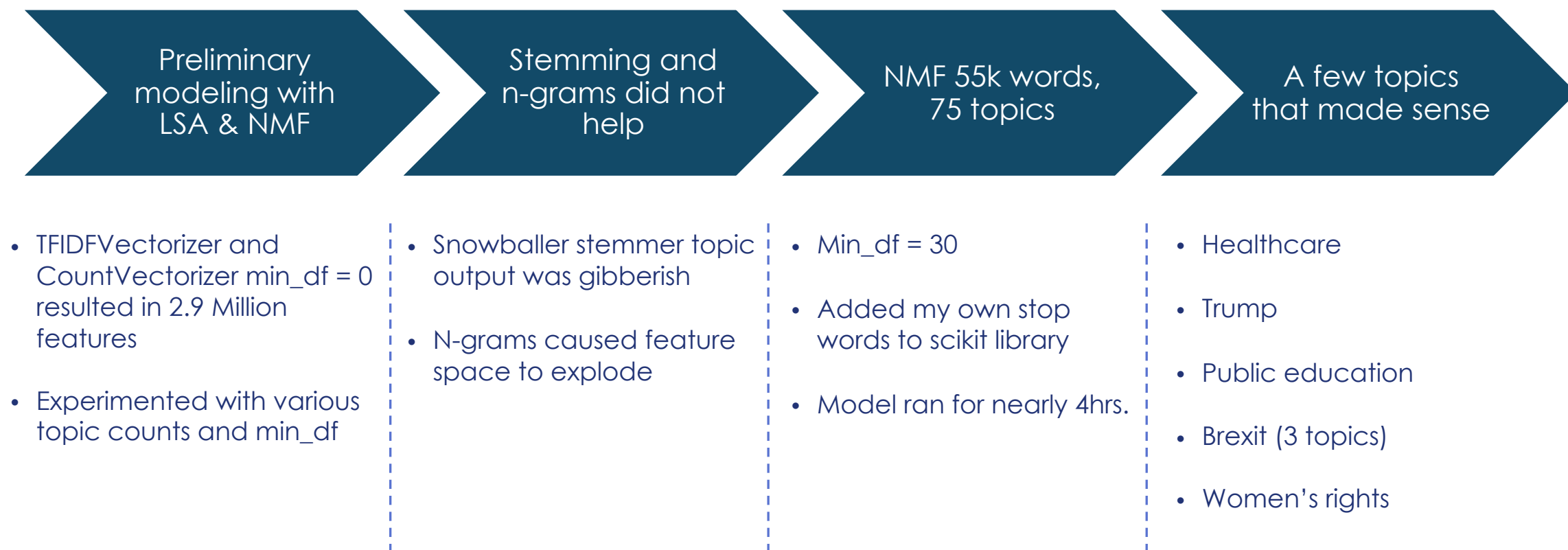
**Brexit and healthcare were the most consistent and persistent topics.**  
**Women’s rights, Trump, and climate also appeared.**  
**Immigration, religion, economy did not.**

Twitter data acquisition workflow



# Exploratory modeling: is it an elephant, a giraffe, or a T-rex?

## The surprise was what did not appear as a topic: immigration and religion



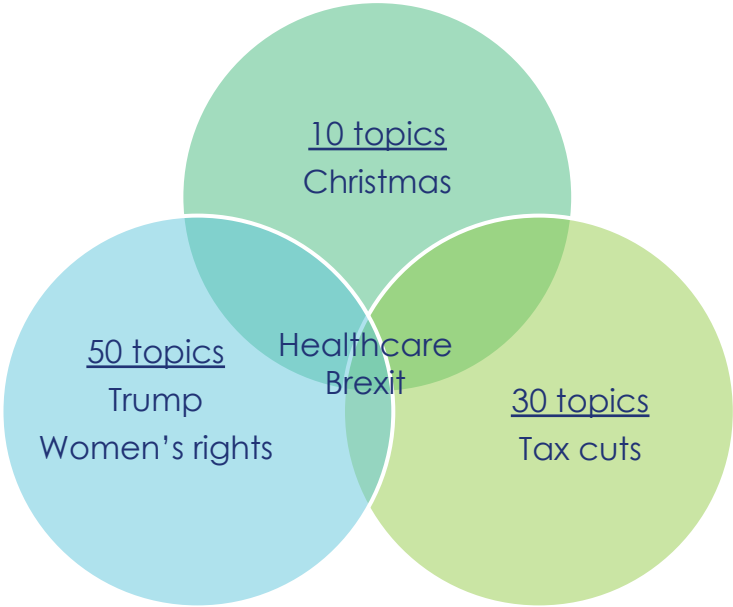
# NMF with different time periods and n\_component hyper\_parameters

## Healthcare and Brexit consistently appear as topics

NMF different time periods and n\_component = 10, 30, and 50  
( stop\_words = 'English' union my custom stop words)

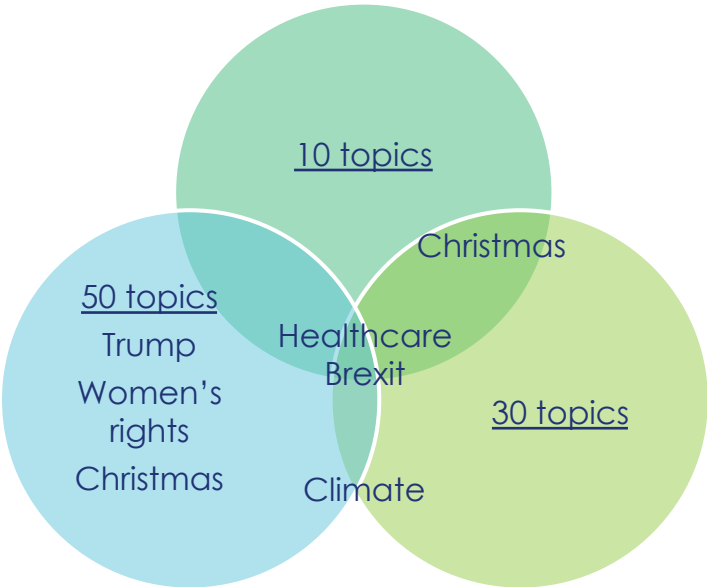
First 8 months

May 2017 – Feb 2018  
500k tweets  
min\_df = 30  
16.6K words



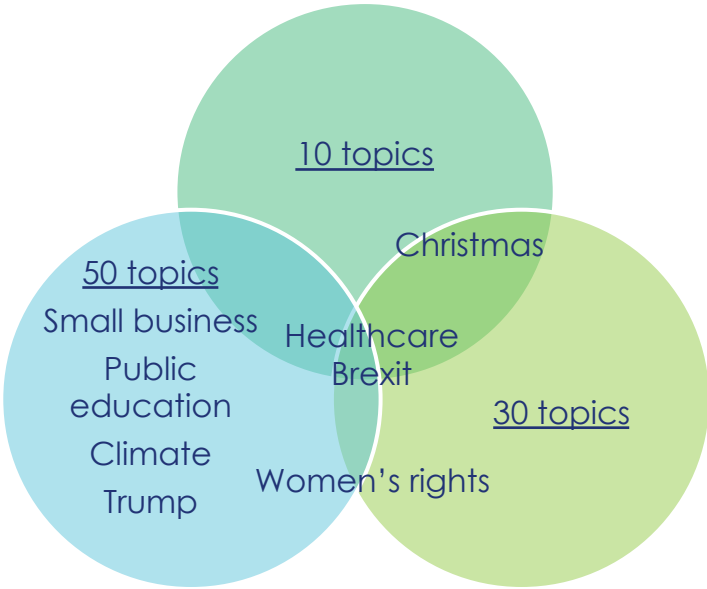
Last 10 months

July 2018 – May 2019  
500k tweets  
min\_df = 30  
15.9K words



All months

May 2017– June 2019  
3.1M tweets  
min\_df = 180  
17.4K words



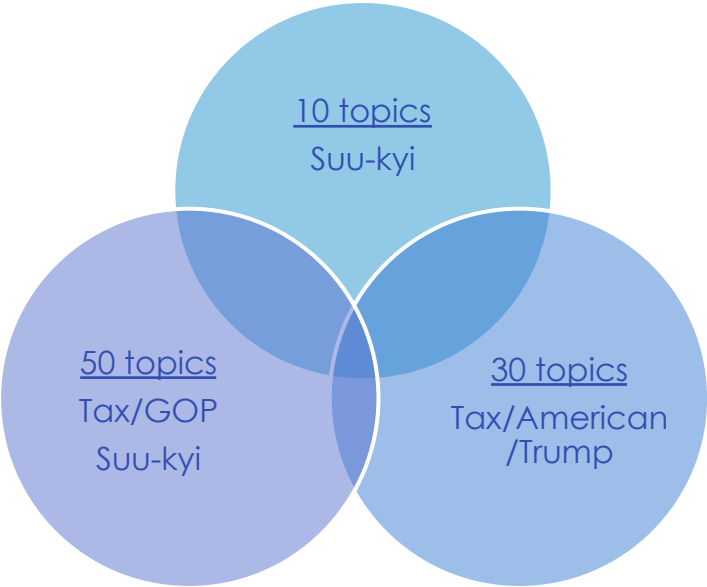
# Corex was less interpretable

## Brexit and Trump dominant topics for last 10 months and all month time periods

Corex unsupervised different time periods and n\_component = 10, 30, and 50  
( stop\_words = 'English' union my custom stop words)

First 8 months

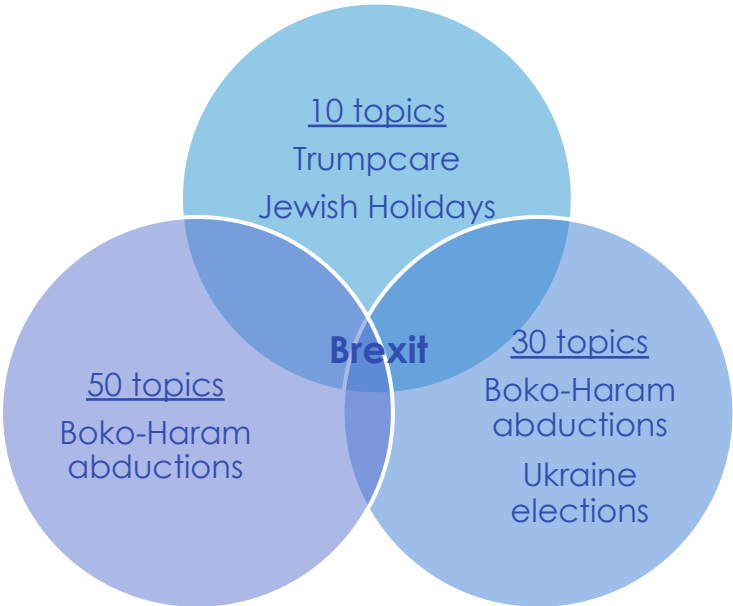
May 2017 – Feb 2018  
500k tweets  
min\_df = 30  
16.6K words



Topics difficult to interpret and diffuse

Last 10 months

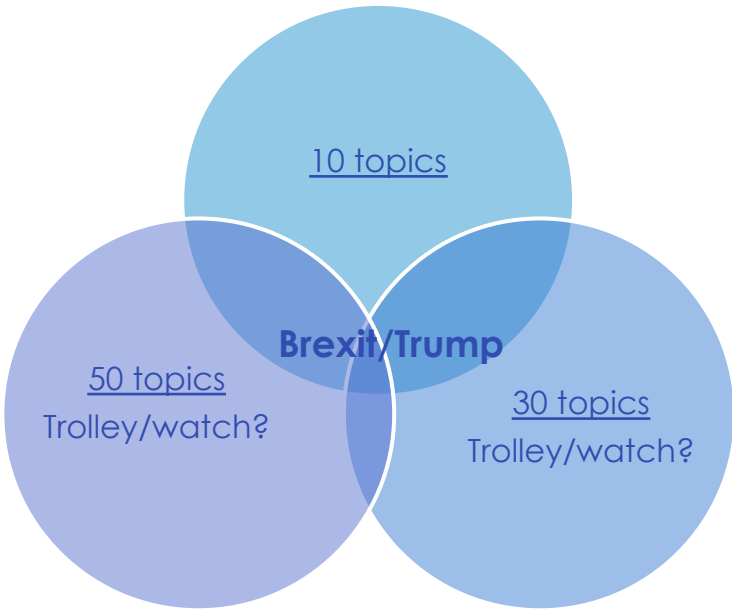
July 2018 – May 2019  
500k tweets  
min\_df = 30  
15.9K words



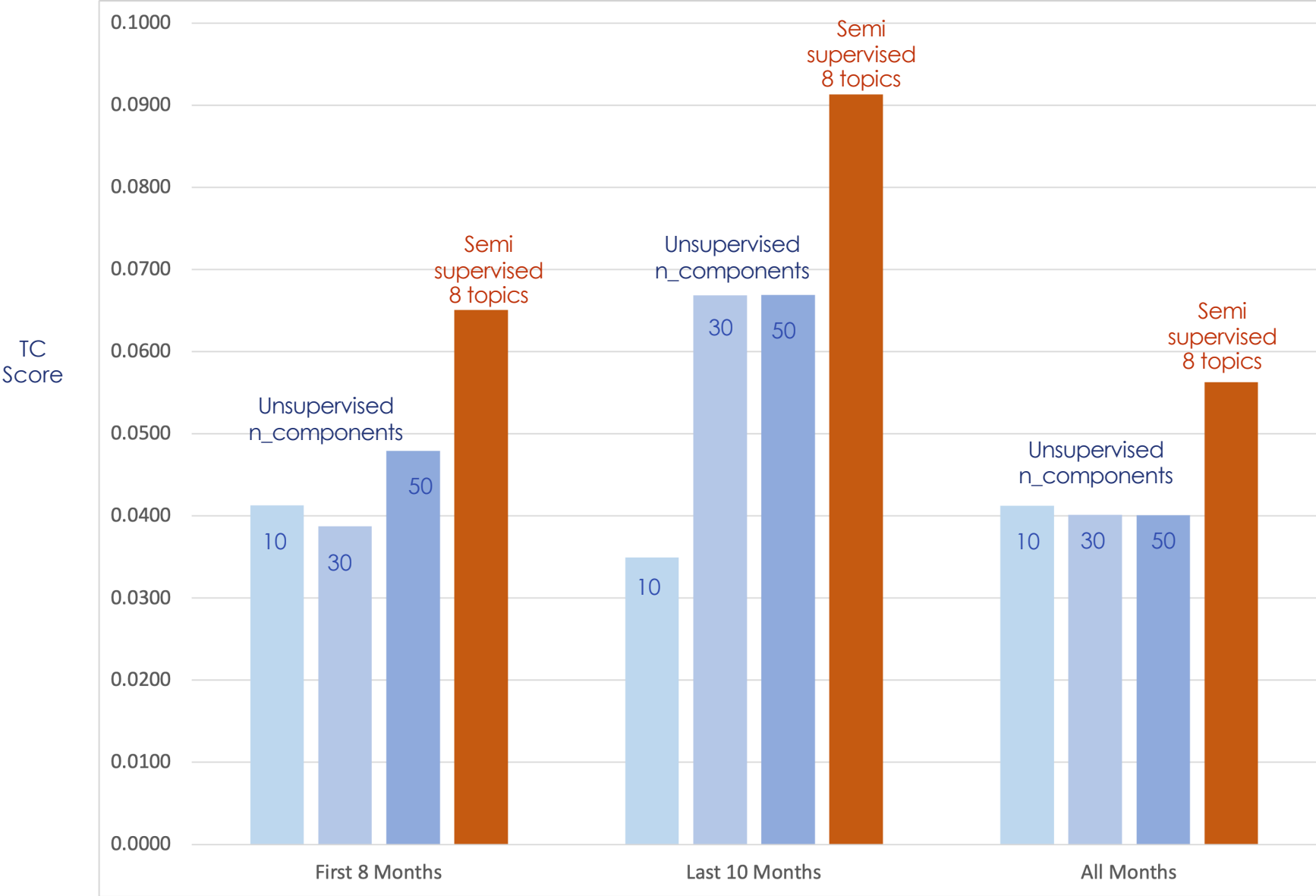
Topics more interpretable and defined

All months

May 2017– June 2019  
3.1M tweets  
min\_df = 180  
17.4K words



# Corex semi-supervised resulted in higher TC scores

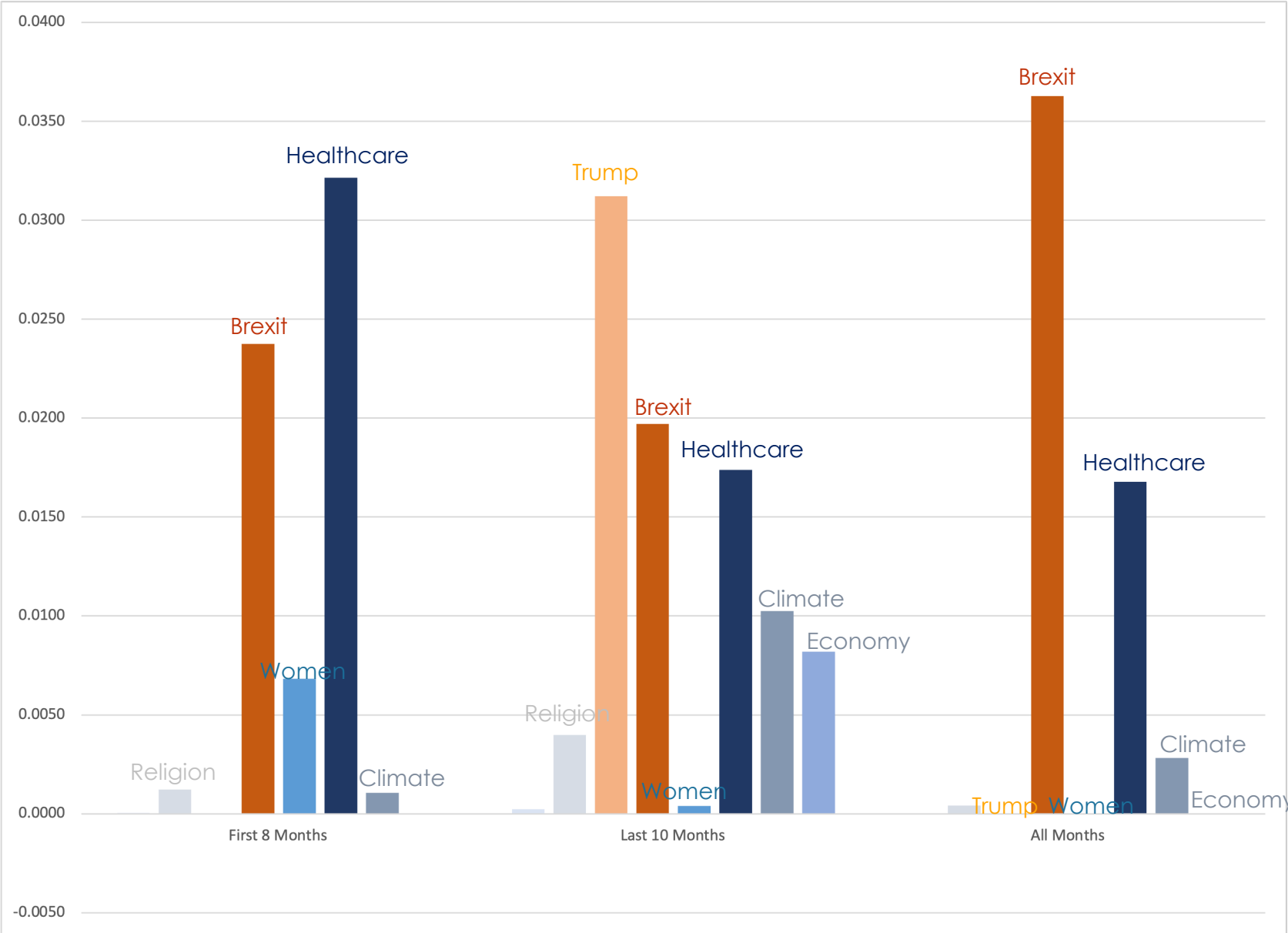


## Semi-supervised topic anchor words

immigration, immigrant  
Christian, Hindu, Muslim, Jewish, religion  
trump  
brexit, eu, leave  
women, violence, rights  
health, care, Insurance, mental  
climate, change  
jobs, economy, business

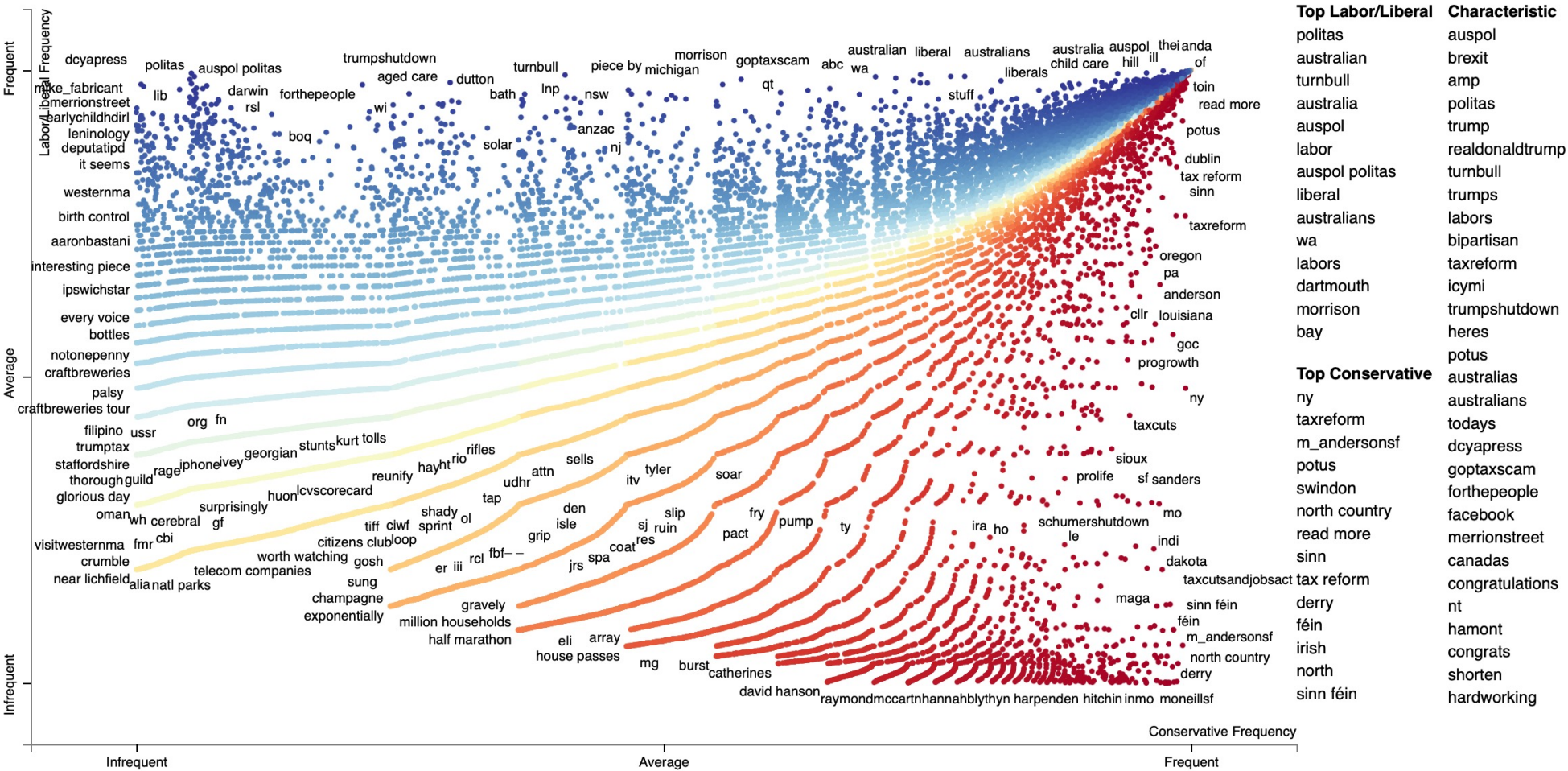
### Brexit and healthcare consistent key topics

TC Score



- Immigration and religion absent
- Last 10-month period sharper topic definition
- Is Brexit subsuming topics like economy?

# Finally used spacey and ScatterText to map overall word frequency and relationship on Liberal v Conservative axes



Search the chart

Note: Used code from a ScatterText tutorial to generate