

A Tale of Two Nations



William Baum

“Crossing the Swamp”

<http://jonmcnaughton.com/crossing-the-swamp-10x15-litho/>

<https://www.weeklystandard.com/hannah-yoest/angry-kitsch>

<http://jonmcnaughton.com/crossing-the-swamp-10x15-litho/>

Русский vs English

- ◆ *Project goal:*
 - ◆ *Compare Russian twitter topics to English ones...*
 - ◆ *Regarding international relations in light of the recent collusion investigation, Russia's alleged election meddling*
 - ◆ *Search terms: USA, Russia, bot, troll, Putin, Trump*
 - ◆ *США, Россия, Русский, бот, тролль, Путин, Трамп*

Which topics concern people?

Let's take a look at twitter activity in the two countries!

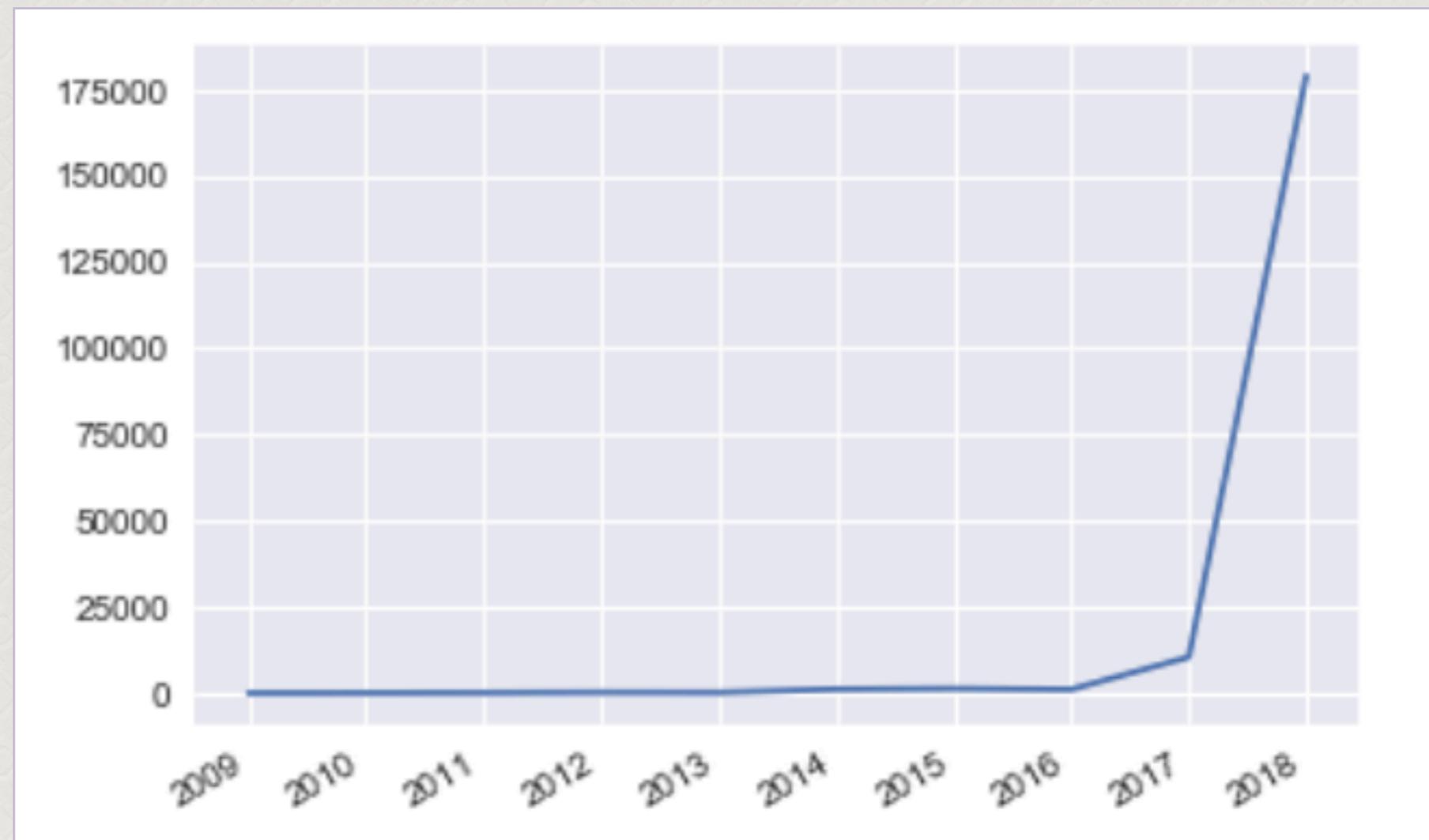
*English in the USA
Russian in Russia*



media.istockphoto.com/crowd-of-people-walking-on-street

Recent Tweets

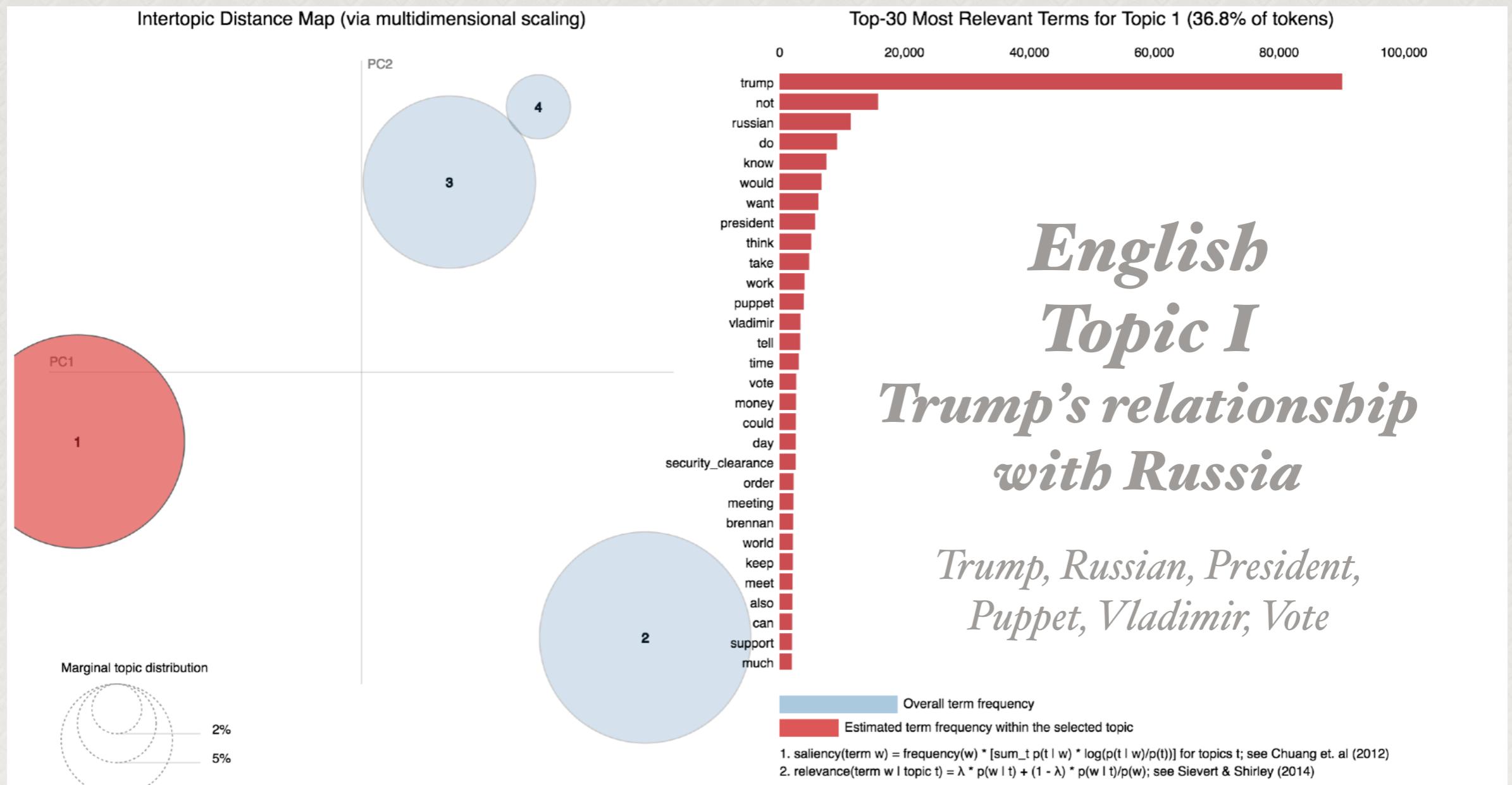
- ◆ 300,000 tweets scraped
- ◆ Balanced Corpus
- ◆ Equal English & Russian Tweets
- ◆ Still... Lots of Bots!



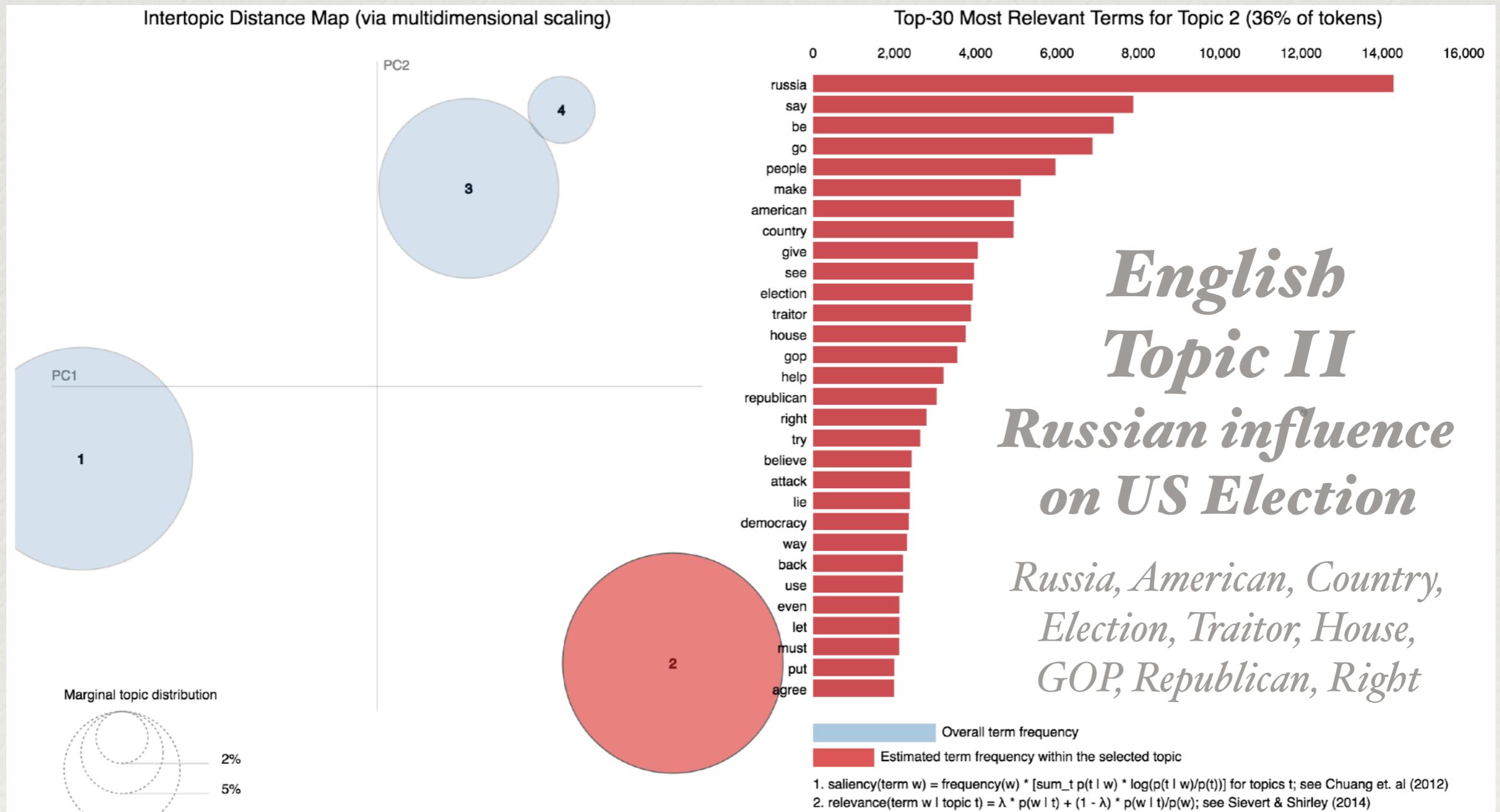
Process

- ◆ *Tokenization*
- ◆ *Removed: duplicates, links, hashes, mentions*
- ◆ *Stop-words: English, Russian, custom list for Russian language*
- ◆ *Punctuation*
- ◆ *NMF*
- ◆ *LDA - gensim*
- ◆ *MDS (t-SNE option)*
- ◆ *Shepard plot - pyLDAvis*

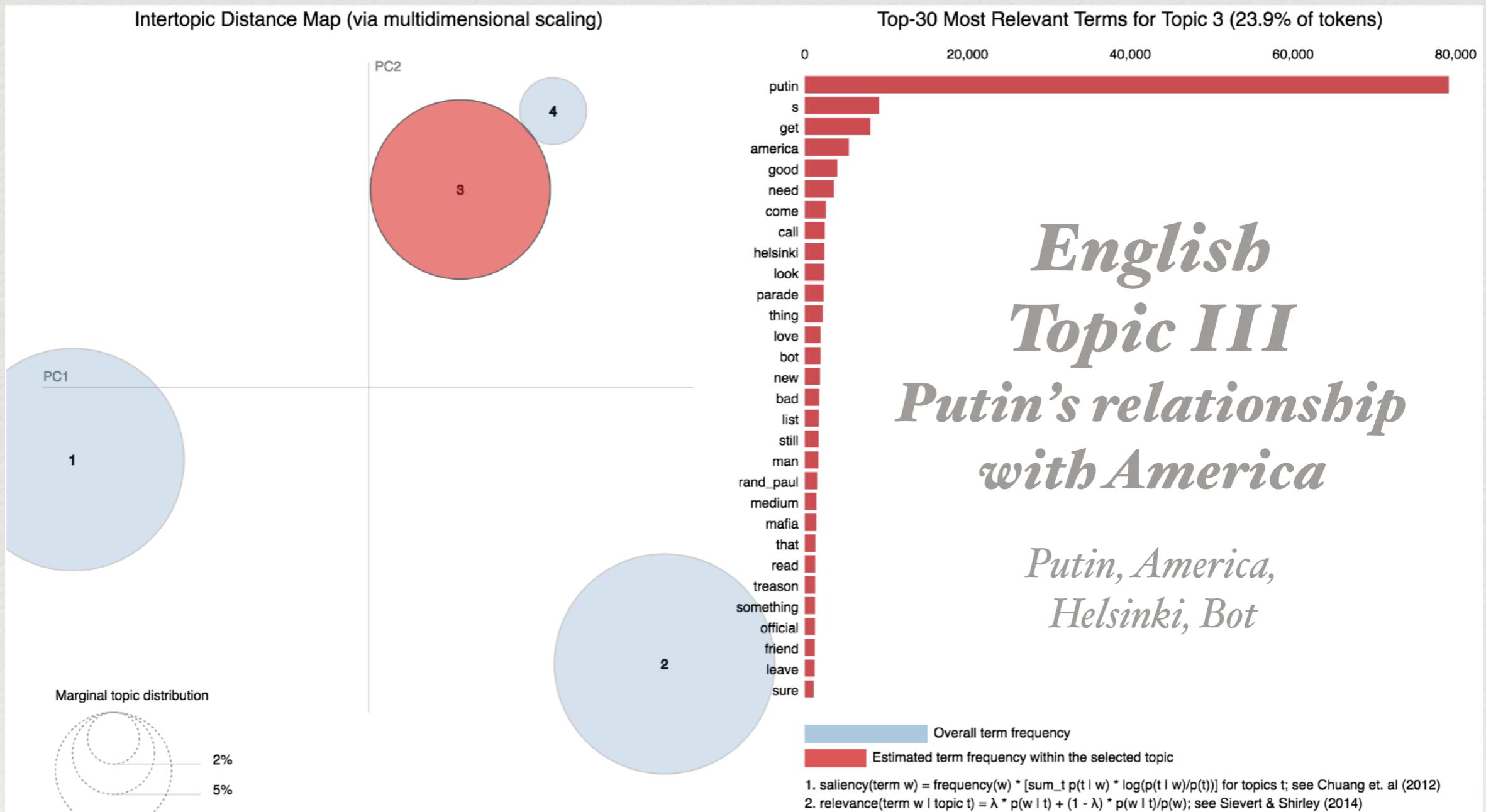
MDS



MDS



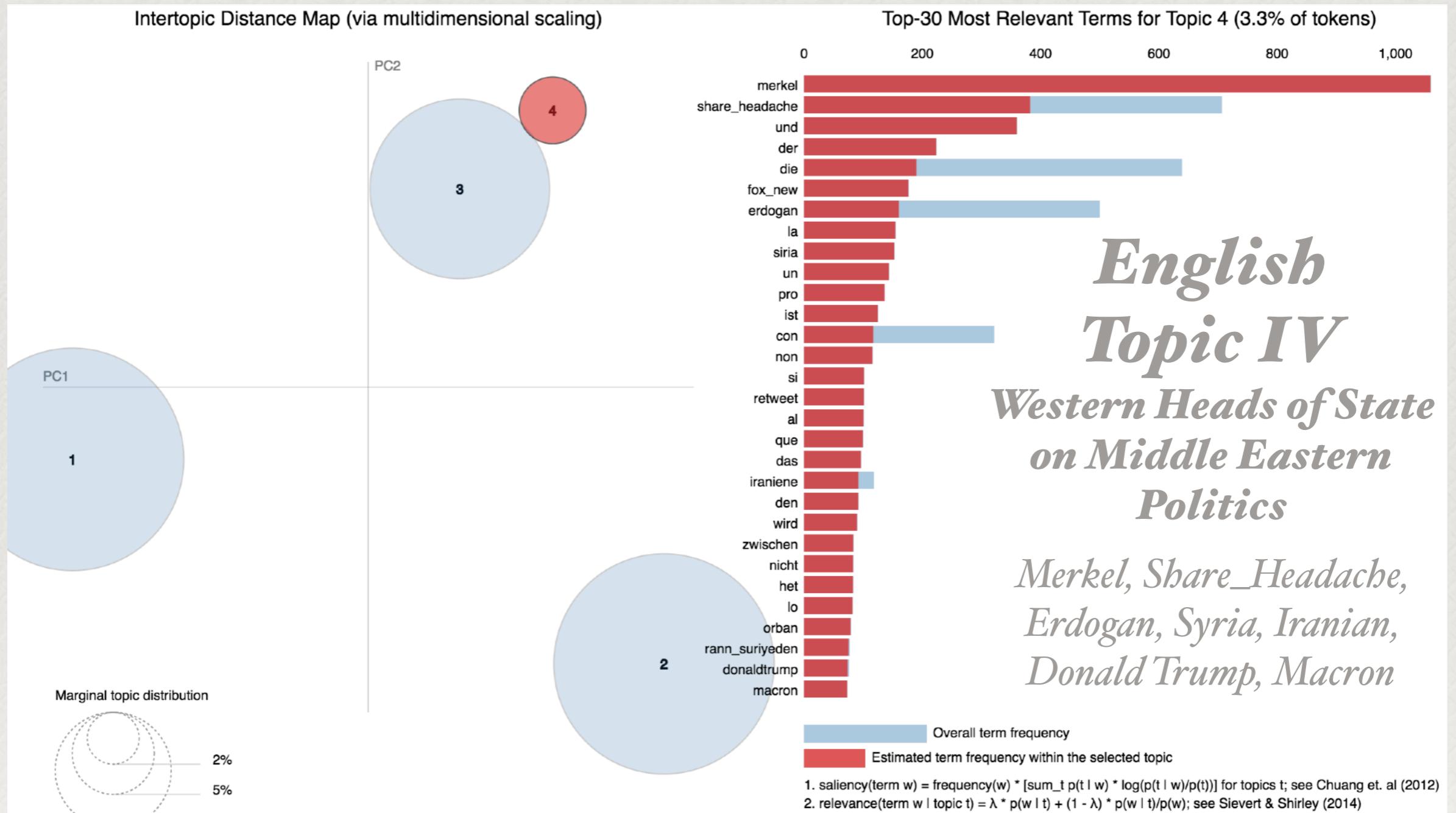
MDS



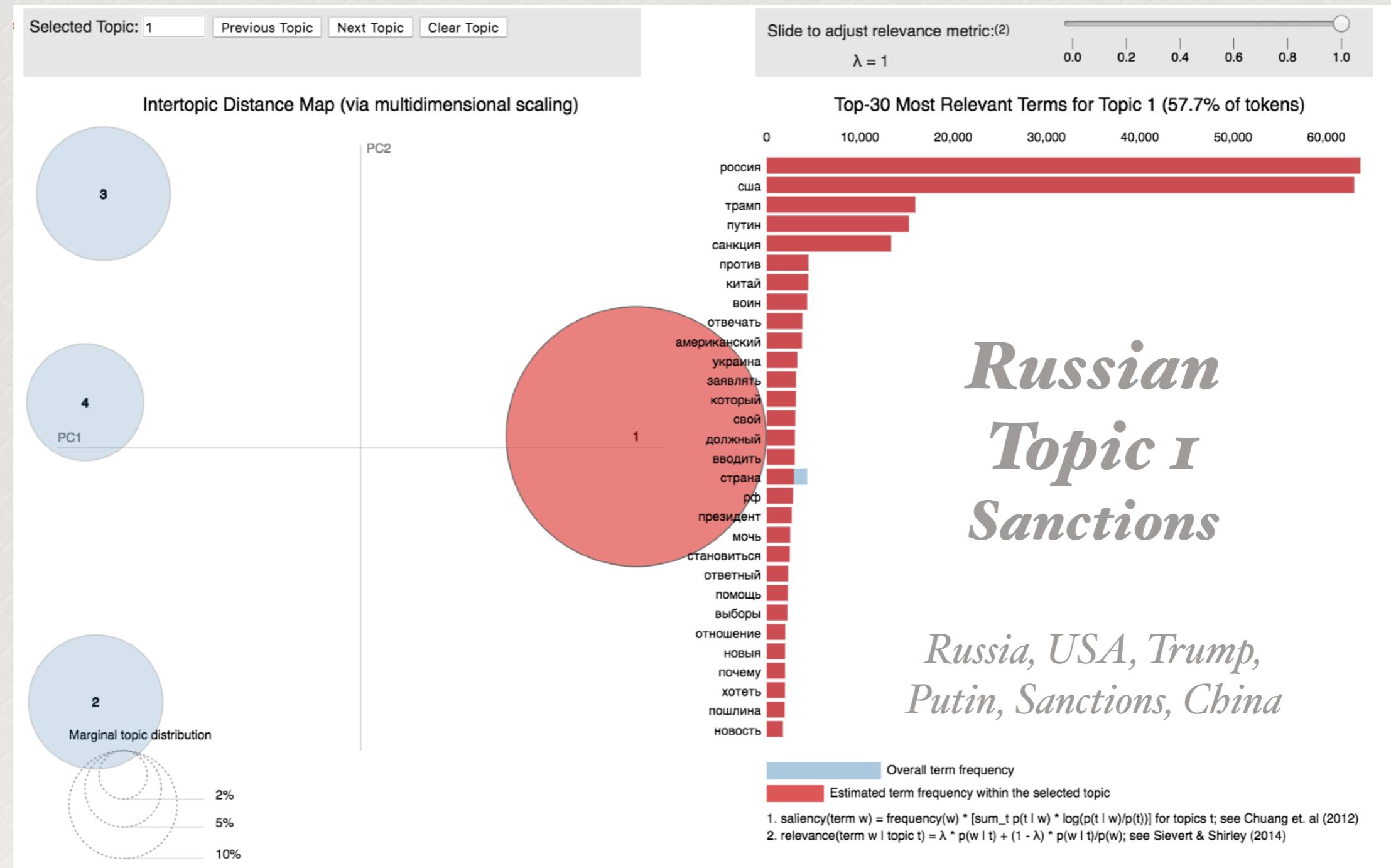
English Topic III *Putin's relationship with America*

*Putin, America,
Helsinki, Bot*

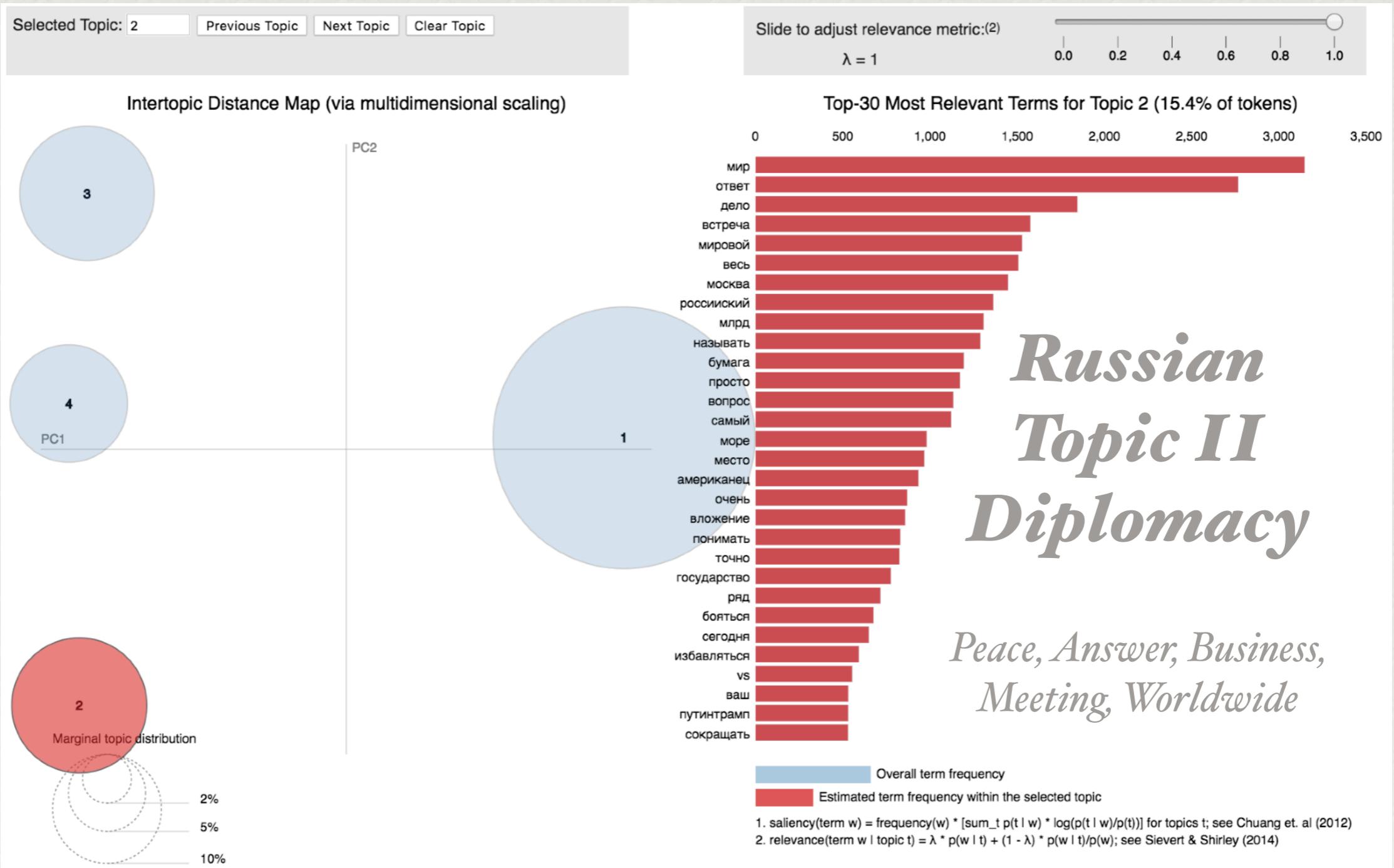
MDS



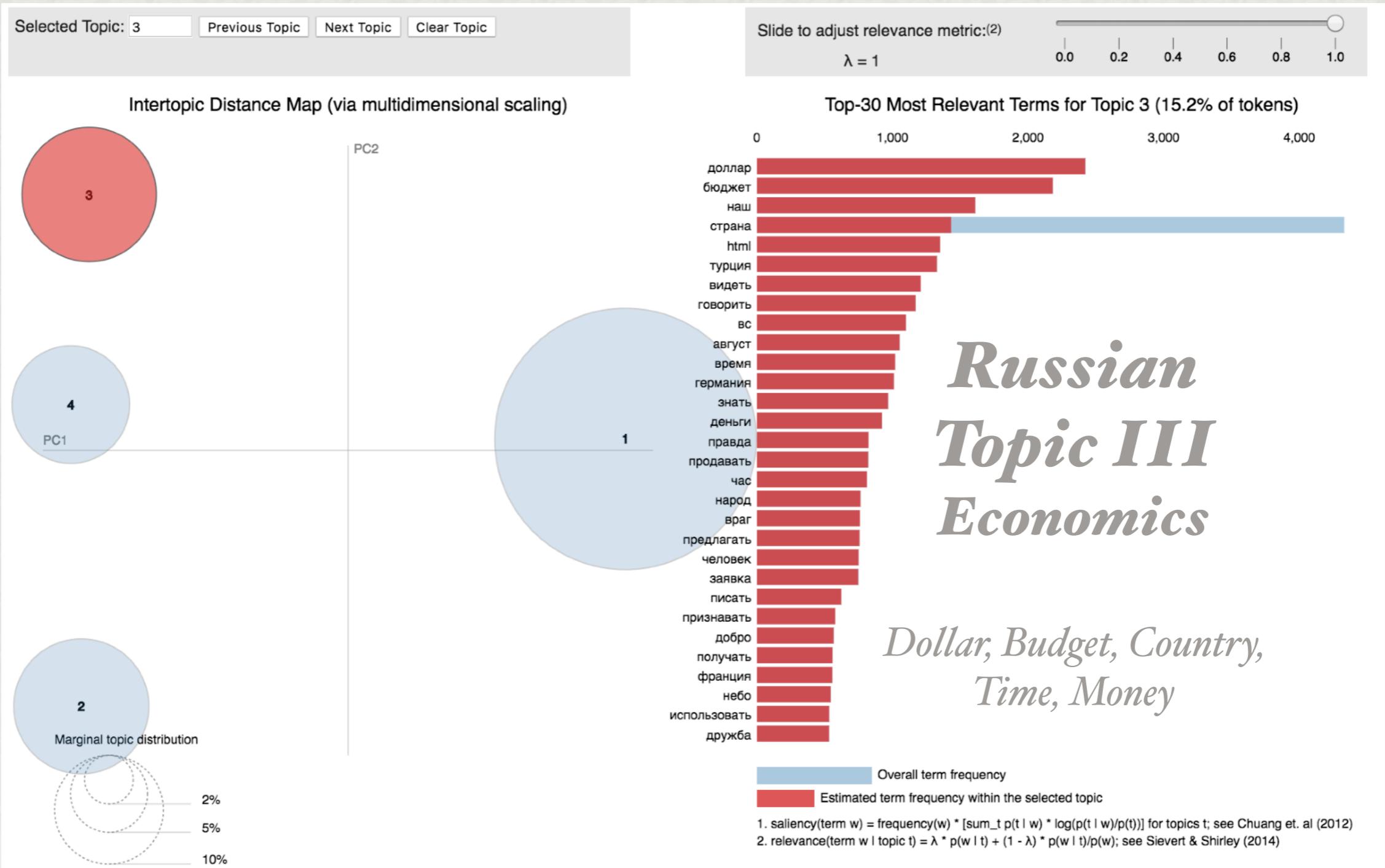
MDS



MDS



MDS



MDS

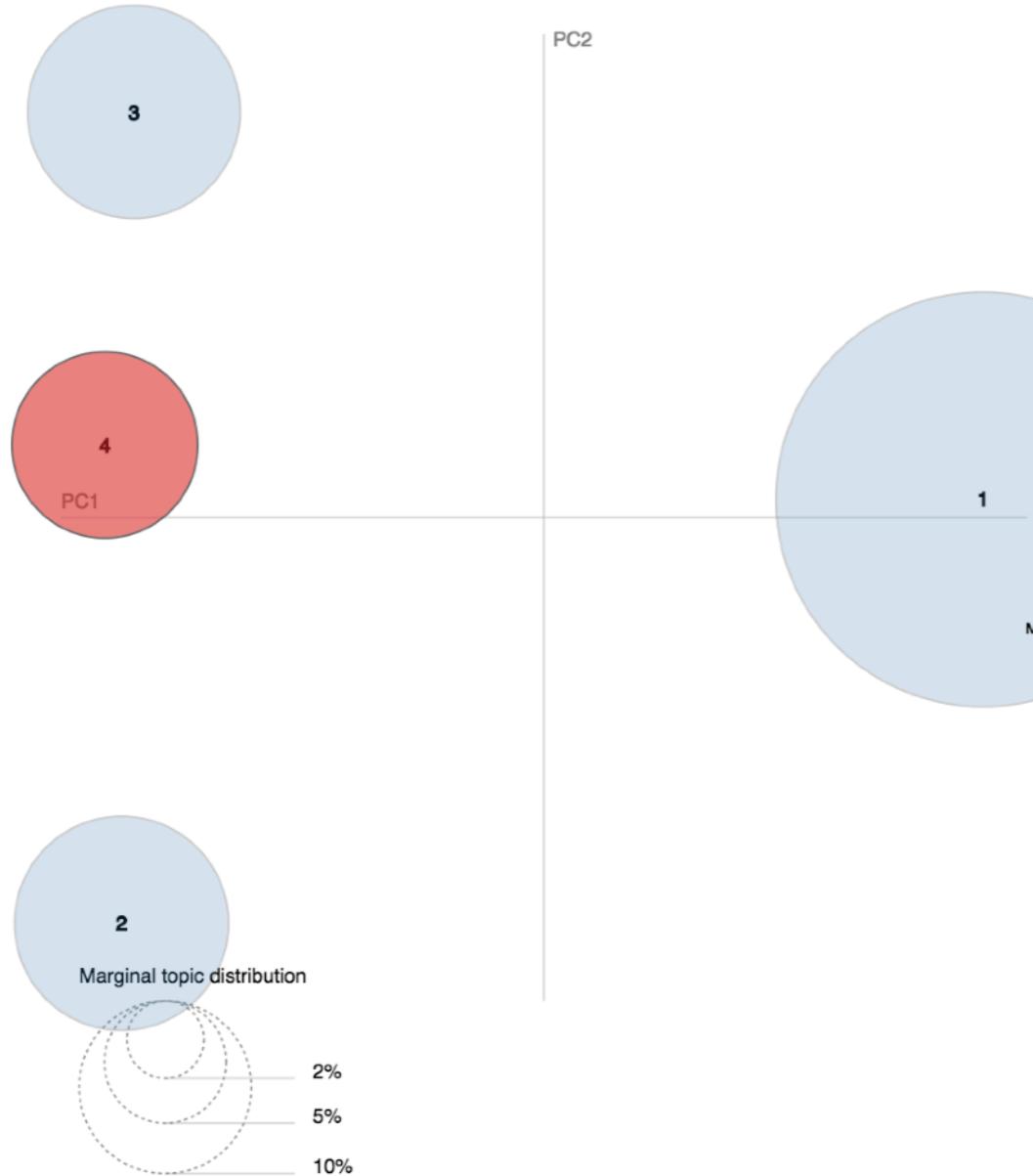
Selected Topic: 4

Slide to adjust relevance metric:(2)

$\lambda = 1$

0.0 0.2 0.4 0.6 0.8 1.0

Intertopic Distance Map (via multidimensional scaling)



Top-30 Most Relevant Terms for Topic 4 (11.7% of tokens)



*Russian
Topic IV
New Negotiations*

*New, Year, Friend, Begin,
Discuss, Russian*

Overall term frequency
Estimated term frequency within the selected topic
1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

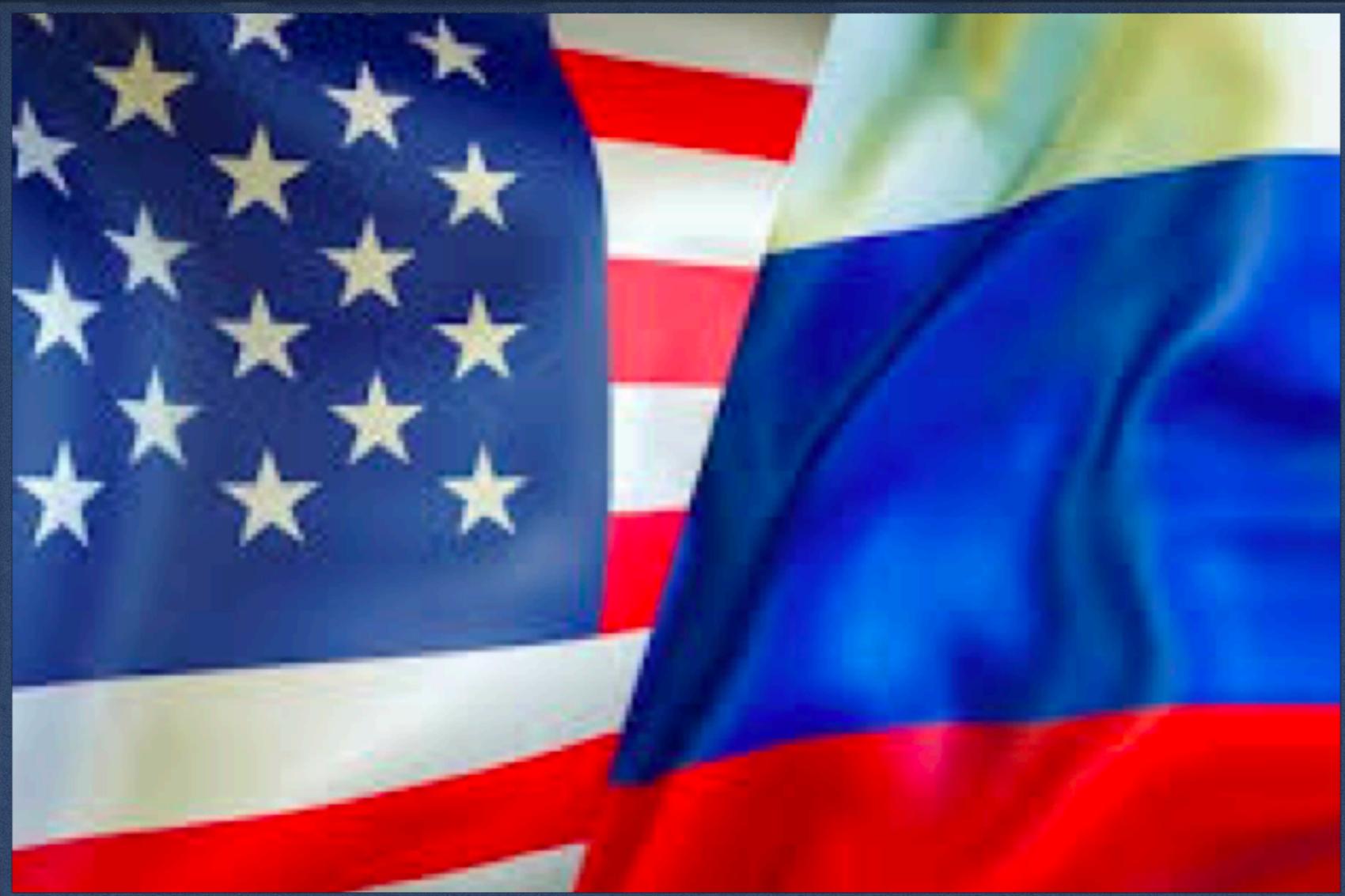
Summary

- ✿ *Predictably, the concerns of Russians and Americans differ, as does their twitter post content.*

<https://www.youtube.com/watch?v=3Ar8osFzViw>

SNL opening skit, Time: 1:39 to 3:51

- ✿ *Individuals seem interested in discussing topics that more directly relate to their everyday lives.*



<https://image.shutterstock.com/usa-russia-flags-3d-waving>

Let's hope for improved relations and lasting peace!

After all, both countries have loads of nukes!