

Russia \(\Display \) Finland

Reflection on Neighbours Next Door

1 June 2018

Alexey Igonen, Arturs Polis Ilona Repponen, Miika Lampi Mila Oiva, Victoria Tkachenko

Group leaders:

Andrey Indukaev, Daria Gritsenko

Research question





What are the images of Finland in the Russian media and Russia in the Finnish media?

FINLAND

120.000+ articles





Newspapers:

Russian Federal

Russian Regional

YLE Finland



Case studies





- 1. News agenda
- 2. Dynamic geography
- 3. Understanding of the 'neighbour'

What's on the agenda?



SPORTS and POLITICS

Finnish media (YLE)









sopimus

with kausipoliisi oyhdysyallat

palipulati oyhdysyallat

palipulation

palipulation

palipulation

palipulation

palipulation

palipulation

palipulation

palipul





Russian media (federal)









DIGITAL HUMANITIES HACKATHON

MAY 2018 HELSINKI

Politics EU child war Ukraine NATO

призовой ких ТУ

Integrum Federal - 2012

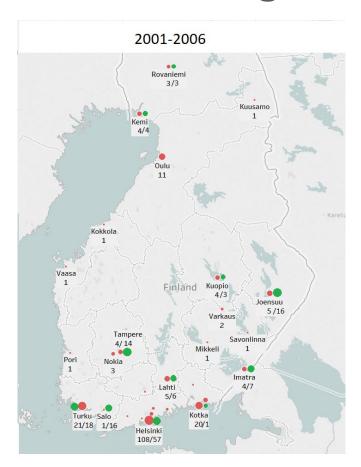


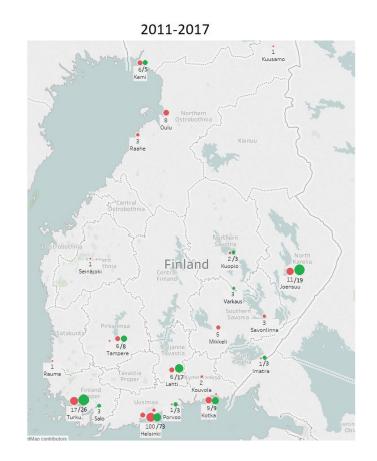




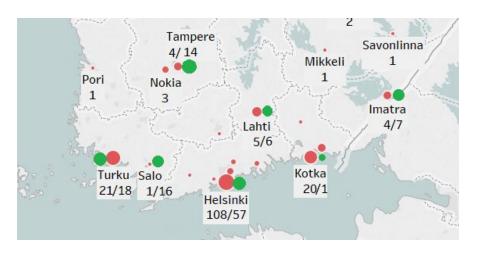


Where things happen?



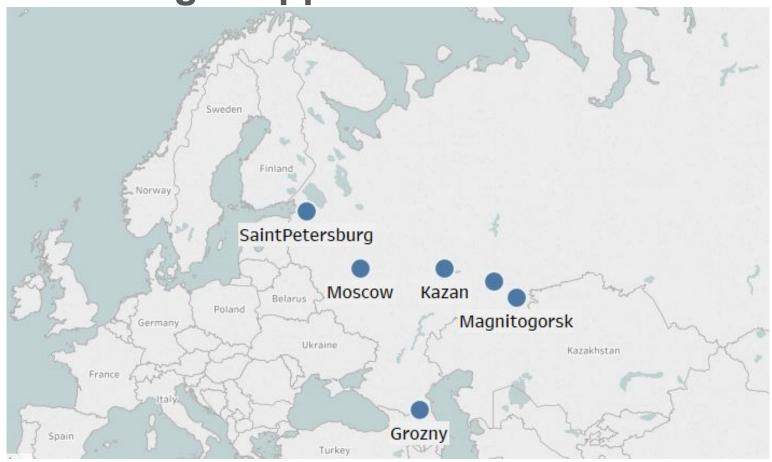


Where things happen?





Where things happen?



What is neighbourhood?







H 18
H 18
DIGITAL HUMANITIES HACKATHON
MAY 2018 HELSINKI

That's where 'neighbour' comes all political

Example period 2004-2009								
RuFed	Fellow citizen	European	West	Ally	Slogan	America	Alarm	

HH18
H18
DIGITAL HUMANITIES HACKATHON
MAY 2018 HELSINKI

That's where 'neighbour' comes all political

Example period 2004-2009								
RuFed	Fellow citizen	European West		Ally	Slogan	America	Alarm	
RuReg	Fellow countryman	Width	Fellow citizen	All over	Palmyra	Finn	European	

HH18

DIGITAL HUMANITIES HACKATHON
MAY 2018 HELSINKI

That's where neighbourhood comes all political

Example period 2004-2009								
RuFed	Fellow citizen	European	West	Ally	Slogan	America	Alarm	
RuReg	Fellow countryman	Width	Fellow citizen	All over	Palmyra	Finn	European	
YLE	Interest	Tension	Stability	Foe	Military	Folkhemmet	Threat	

Patterns in color

Western
countries
Aggression
Community
Finns



Example period 2004-2009								
RuFed	Fellow citizen	European	West	Ally	Slogan	America	Alarm	
RuReg	Fellow countryman	Width	Fellow citizen	All over	Palmyra	Finn	European	
YLE	Interest	Tension	Stability	Foe	Military	Folkhemmet	Threat	



Computational techniques

Data Cleaning

json to csv

lemmatization -> returning the words to their basic form

removing the stop words -> the not meaningful "and", "or"...



Cincinnati Bell Historical Archives



Sports and Politics - techniques

Dominant annual agendas

TF-IDF

The most significant word from each article row 41 Headline:

Министр Рясянен прине

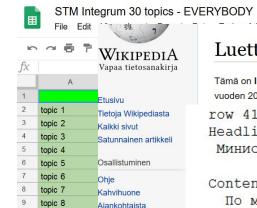
Content:

По мнению отвечающей в «черный список» поли екла информация об оши ладимира Путина в поли и в Македонию и Хорвал основе занесли в регис анный регистр является мации представляет соб утри Полицейского упра е знает, как имя Путин

т, что имя президента РФ было занесено в полицейский регистр совершенно нео-

лохой шуткой. - Мы сегодня выяснили дело, и пришли к выводу, что никаких за

Where things happen? - techniques



Tuoreet muutokset

Tänne viittaavat sivut

Linkitettyjen sivujen muutokset

Toimintosivut

Sivun tiedot

Viitetiedot

Wikidata-kohde

Tulosta tai vie

lkilinkki

Työkalut

topic 9

topic 10

topic 11

topic 12

topic 13

topic 14

topic 15

topic 16

topic 17

topic 18

Luettelo Venäjän suurimmista ka

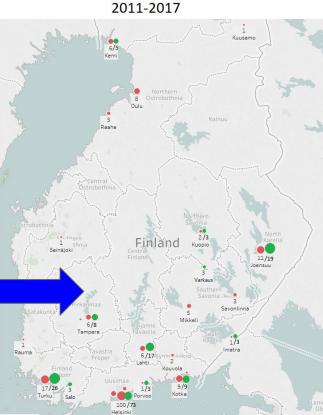
Tämä on luettelo Venäjän federaation suurimmista kaupungeis vuoden 2010^[1] virallisen väestönlaskennan mukaan oli vli 50 000 h

row 41 Headline:

Министр Рясянен принесла извинени

Content:

По мнению отвечающей в правитель в «черный список» полиции связано екла информация об ошибке в регист ладимира Путина в полицейский реги и в Македонию и Хорватию. - У основе занесли в регистр лиц ОДС анный регистр является секре мации представляет собой вто утри Полицейского управления. поли е знает, как имя Путина могло попа т, что имя президента РФ было зане лохой шуткой. - Мы сегодня выяснил



IUMANITIES HACKATHO

AY 2018 HELSINKI

ΤП

OIL

1eH

QNI

I II

1.

CTY

DKI

DCY

DTB

1eo

зa

Understanding the 'neighbour' - techniques

('латвия', 0.6271222829818726),

W2V library

nearby words "use-synonyms"

5-year window

clustering of nearby words



Challenges and Limitations

Wordclouds, Method: TF-IDF

- Lemmatization
- Timestamps
- Running TF-IDF on individual articles VS combined yearly data
- 1 pass VS 2 pass TF-IDF
- Can miss short-lived keywords!



Challenges and Limitations

Geo Mapping and Topic Modelling, Method: POI mapping, STM

- Lemmatization
- Place name transliteration and disambiguation
- Selecting the number of topics and topic clusters
- Ambiguous articles and topics



Challenges and Limitations

Word neighbourhoods, Method: Word2Vec

- Timestamps and Lemmatization
- Quantity of data for shorter periods (5 years at a time)
- Picking the right dimensionality and threshold
- Reading too much into it!



Ideas for future research

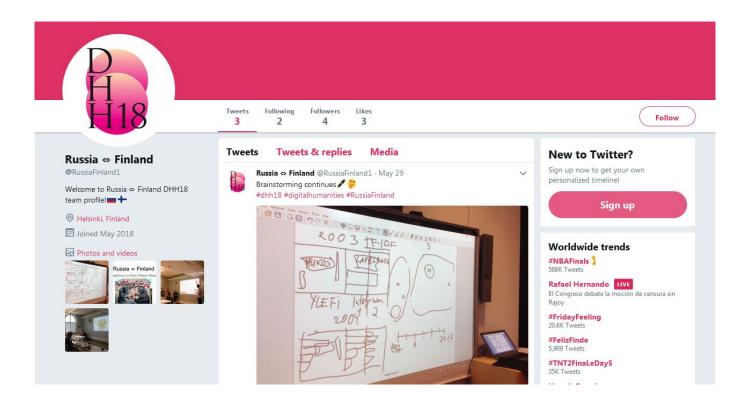


Yearly words per topic - sports, culture, economics

Compare seemingly similar concepts across languages - same or not?

Causality and sentiment analysis - connect events to sentiment

Public outreach during the hackathon





Спасибо!

Kiitos!

Thank you!

HH18

DIGITAL HUMANITIES HACKATHON
MAY 2018 HELSINKI

Questions?