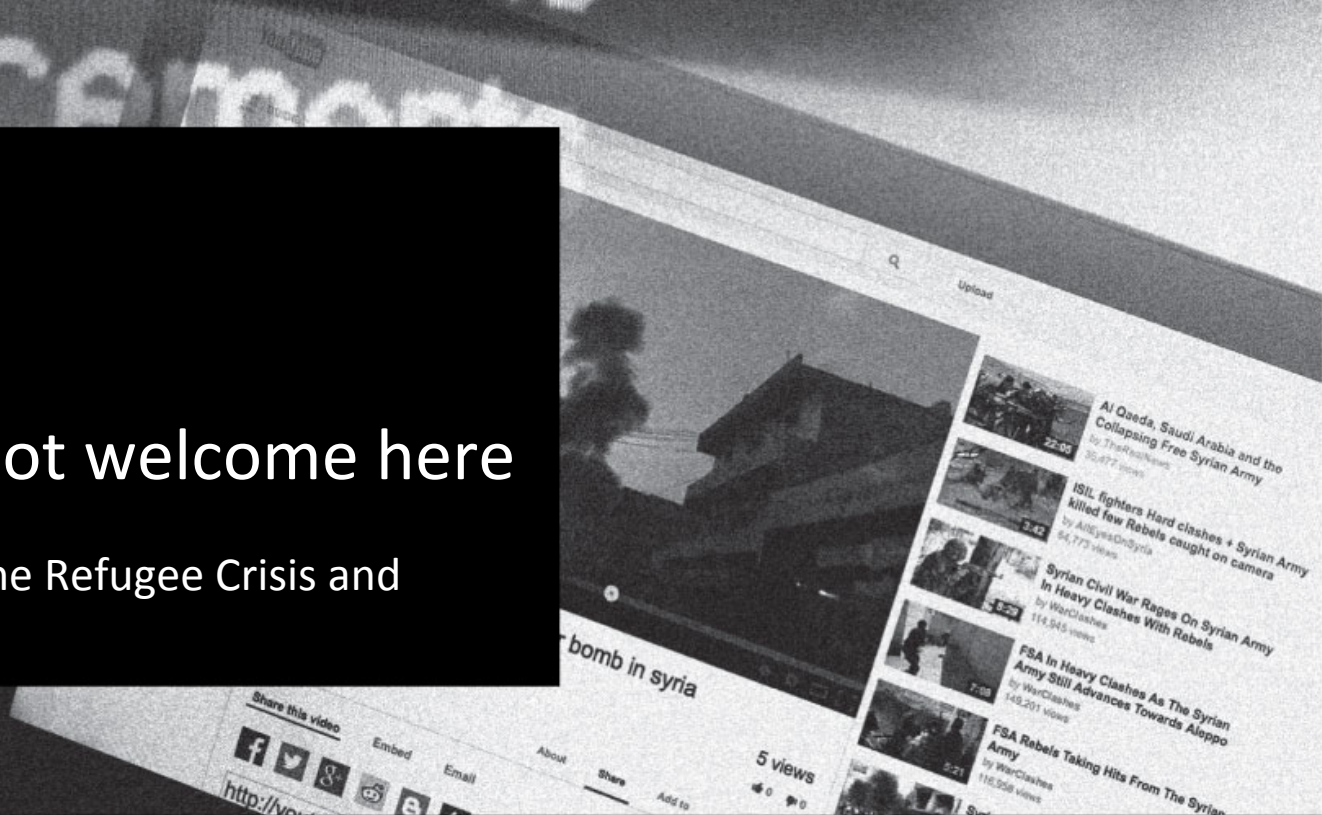


**NETWORK OF EXCELLENCE FOR RESEARCH
IN VIOLENT ONLINE POLITICAL EXTREMISM**



“Rapefugees” not welcome here

**Online Hate Speech, the Refugee Crisis and
Far Right Extremism**



This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 312827

www.voxpol.eu



STRUCTURE OF PRESENTATION

- + **Background** to VOX-Pol research
- + **Research design**, method and case study
- + **Preliminary findings** and results
- + **Conclusions** (and conjectures)



BACKGROUND TO RESEARCH

How a comparative approach can help in CVE research...



mechachal

Online Speech and Elections in Ethiopia



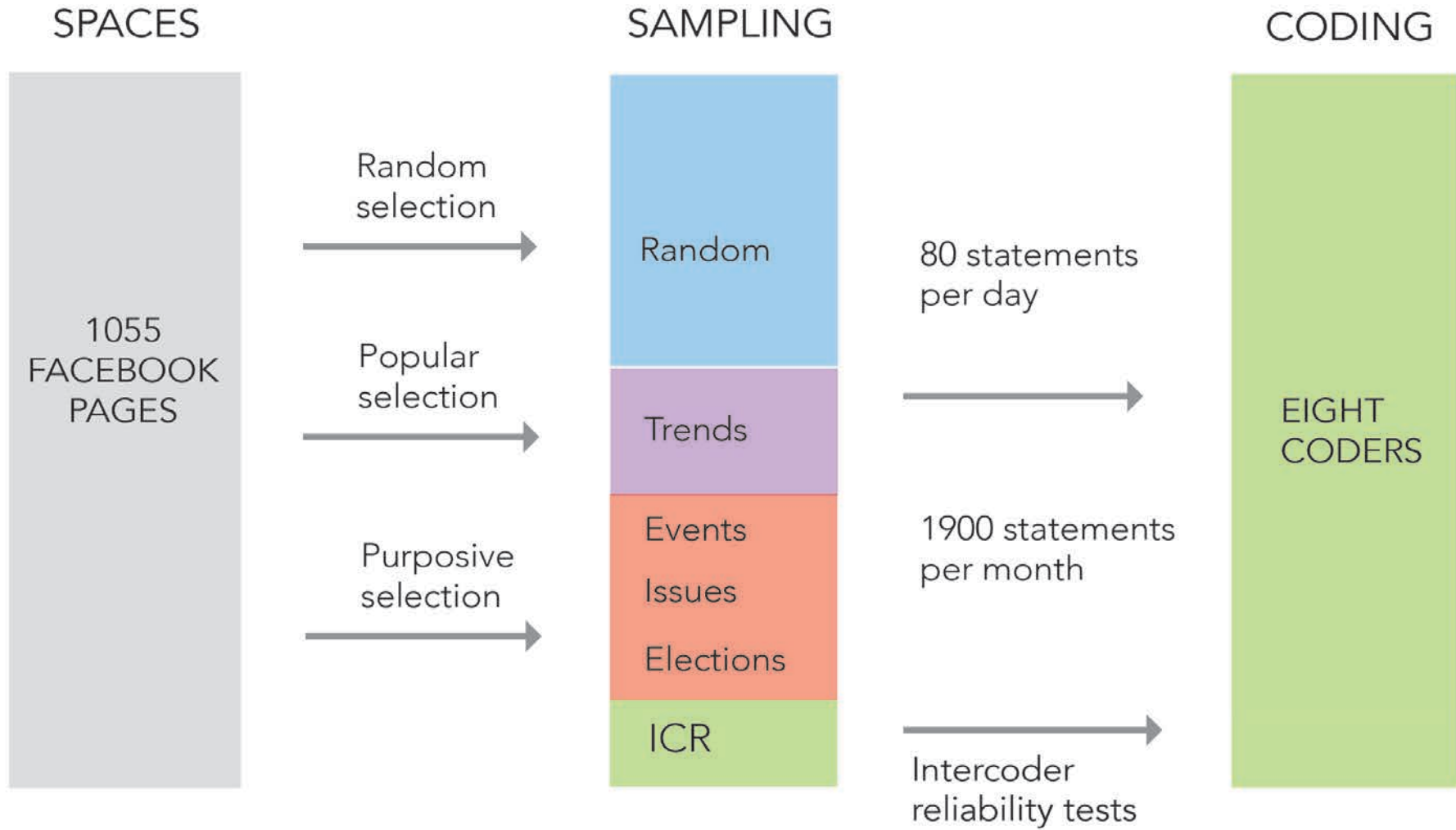


... the government of Ethiopia has on many occasions imprisoned members of the independent press and cracked down on the opposition. At the same time, the government itself has extensively engaged in hate speech against the opposition ... **in a country of ethnicised politics, talking about ethnic injustice is both inescapable and more likely to be punished as hate speech. Hence, what one observes under such circumstances is debate by proxy.**

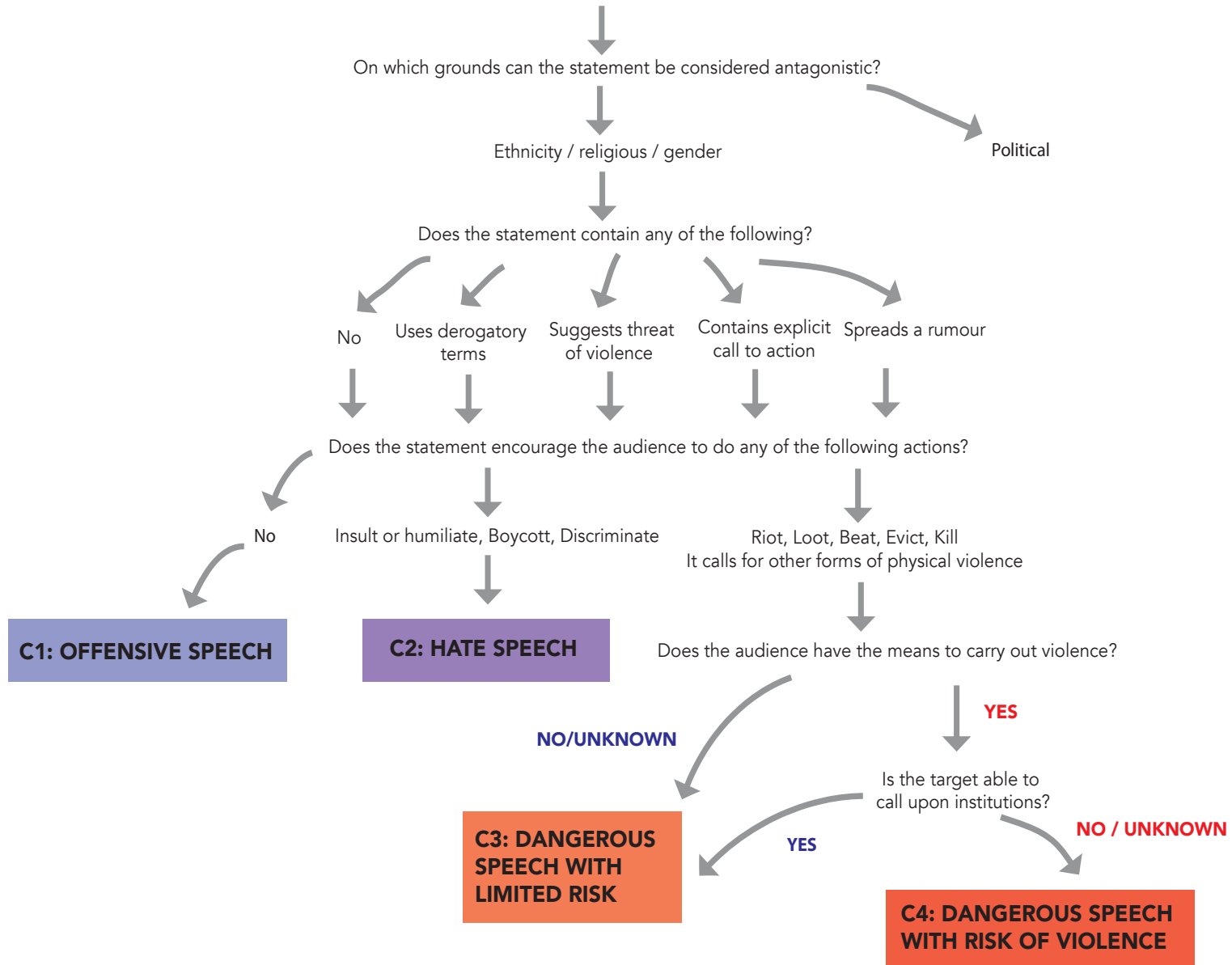
(Legesse 2012: 317)



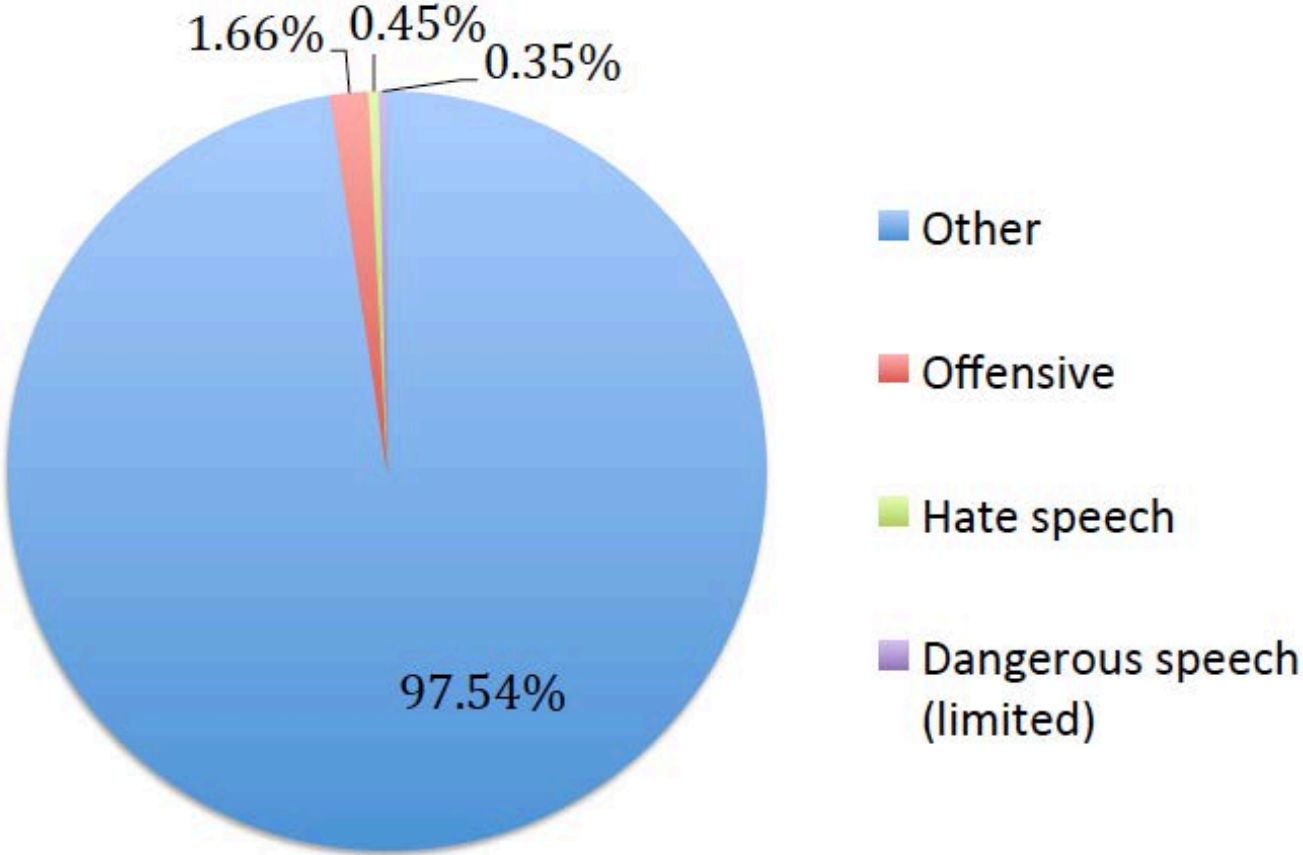
MECHACHAL WORKFLOW



GOING AGAINST EQUATION



**ONLY 0.3% OF SOCIAL MEDIA DISCUSSIONS
WERE POTENTIALLY DANGEROUS (limited risk)**



FREE SPEECH

Posting on Facebook is now a crime under Ethiopia's state of emergency



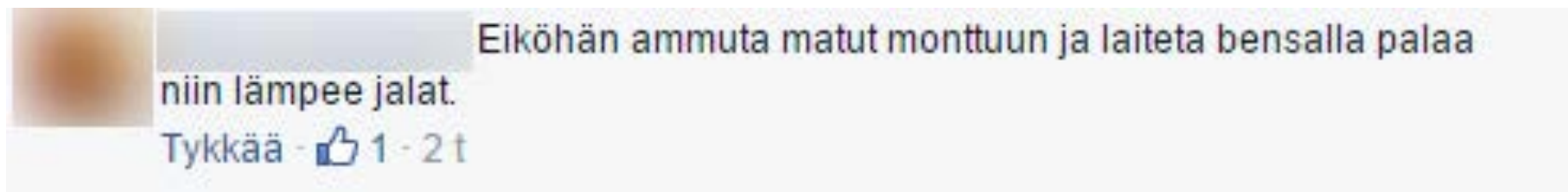
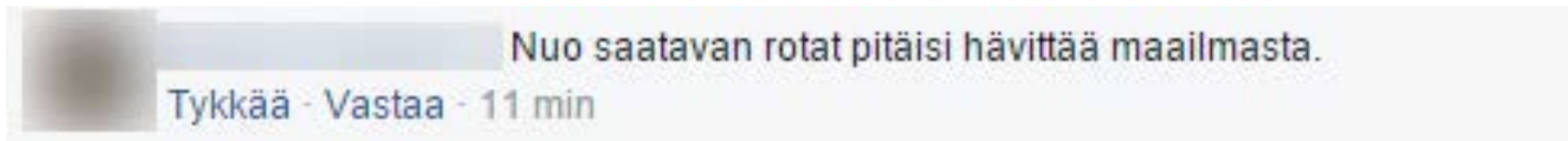


LESSONS LEARNED:

- + **Methodological:** The importance of *not* focusing on "hate speech" *only* in sensitive political climates (i.e. avoiding sample bias and moral panics)
- + **Comparative:** The importance of *comparative research* (i.e. "hate speech" needs to be understood in cultural and political context and not as a universal category)
- + **Communicative relations:** The importance of understanding the *knowledge networks* behind online speech (i.e. when is "hate speech" actually significant/dangerous and its socio-political context?)



*"Those fu***ng rats should be exterminated from the world!"*



"Why don't we shoot the immigrants (invaders) into holes, add some fuel and burn them - this will warm our feet."

Headline (biggest online anti-immigration site in Finland): " A German man rants:
"Open the ovens. They must be destroyed (video)."

i Viff Mitä Vittua?
Leh Eilen 9:00 · 🌐

Saksalaisherran tulivuorenpurkaus
<http://mvlehti.net/.../.../uunit-auki-heidat-on-tuhottava-video/>

"UUNIT AUKI!!! HEIDÄT ON TUHOTTAVA!!! (video)
Saksalaismies saa eppiset raivarit Kölnin raiskausmellakoista.
MVLEHTI.NET

Täyttä asiaa, muslimit vittuun Euroopasta välittömästi ja neekerit perään ja jos meno ei maita niistä on tehtävä lannoitetta. 😡
Tykkää · Vastaa · Eilen 18:36

mutaiset uuniin 😊

Comments: "This is correct! Muslims the hell out of Europe and niggers too. If they do not agree we can make fertiliser out of them | Niggers to the ovens :)"



OBSERVATIONS ON THE "REFUGEE CRISIS"

- + "Hate speech" around the refugee crisis has become more vitriolic and violent than what we found in Ethiopia and other East African countries (countries with a history of ethnic violence, mass killings and genocide)
- + Widespread mistrust of "mainstream" news media and the use of "alternative" online sources for knowledge
- + Purposeful dissemination of "propaganda" by the extreme far right



PROBLEM: HOW TO RESEARCH THIS?

- + Does "hate speech" propagate on social media networks? How does it compare to other kinds of speech (engagement)?
- + What role do different social media groups/users have in promoting or countering/debating "hate speech?" How do communicative dynamics between "ideological" groups differ?
- + How and when do alternative news sources and far right extremists spread hate propaganda and to what effect?



RESEARCH DESIGN:

developing new methods for understanding "hate speech" ...



VOX-POL RESEARCH:


+ **PHASE 1:** A critical analysis of comparative debates on hate speech, violent online political extremism and methods used to counter it (such as counter-speech)

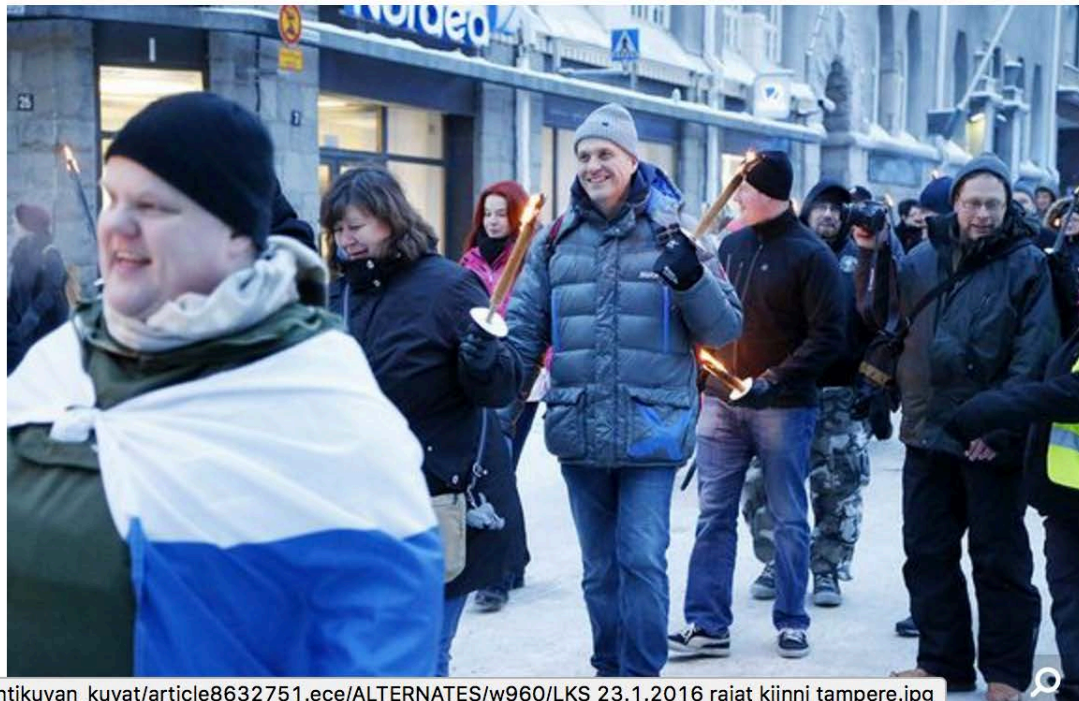
+ **PHASE 2:** Develop and scale up data analysis methods to analyse social media hate speech during the "refugee crisis"

News 29.1.2016 7:03 | updated 29.1.2016 7:52

Helsinki police brace for anti-immigration and anti-racism demos

The anti-immigration group "Close the Borders" (Rajat kiinni) has planned a march in Helsinki to promote peace for women on Saturday. At the same time the anti-racism movement "No racism in my name" (Ei rasismia minun nimissäni) will take to the streets of the capital to protest racist speech and actions.

 Recommend 217 people recommend this. Be the first of your friends.



utiset/lehtikuvan_kuvat/article8632751.ece/ALTERNATES/w960/LKS 23.1.2016 rajat kiinni tampere.jpg

FACEBOOK DATASET



RAJAT KIINNI (anti-refugee)

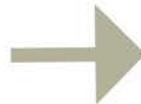
Members = 12443
Posts = 81457
Comments = 654666
Unique posting = 3416
Unique comments = 7734
Unique URLs = 2700
mean comments / post = 8.4
mean likes / post = 27.0
mean shares / post = 1.81

ASIALLISTA KESKUTELUA (group in-between)

Members = 1259
Posts = 1068
Comments = 15763
Unique posters = 262
Unique commenters = 642
Unique URLs = 183
mean comments / post = 15.05
mean likes / post = 5.5
mean shares / post = 0.08

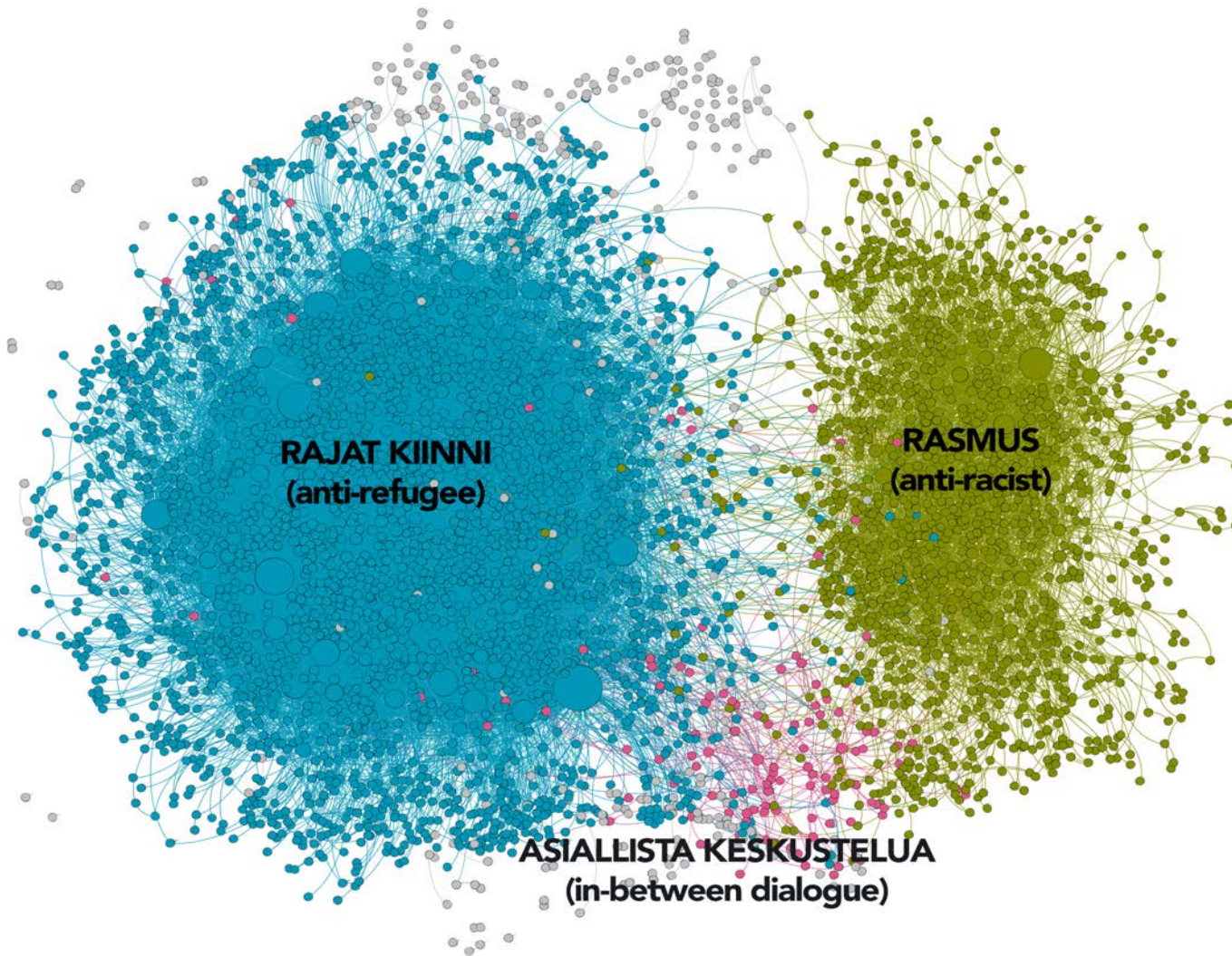
RASMUS (anti-racist)

Members = 13787
Posts = 22836
Comments = 174062
Unique posters = 2263
Unique commenters = 5710
Unique URLs = 2273
mean comments / post = 8.4
mean likes / post = 31.5
mean shares / post = 0.36



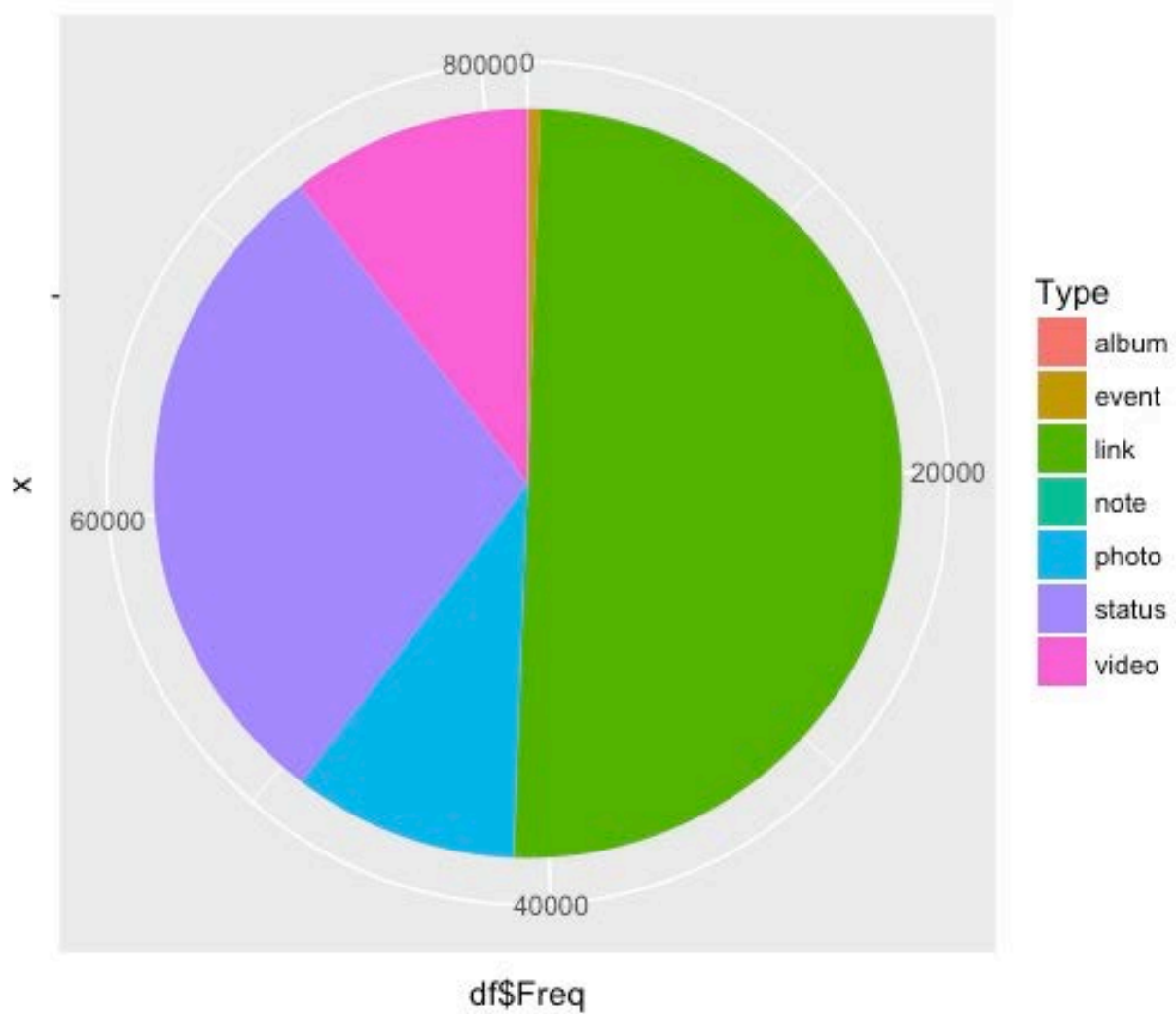
FB RELATIONSHIP BETWEEN GROUPS

(n = random sample of 20,000 out of 840000 - comments x post)

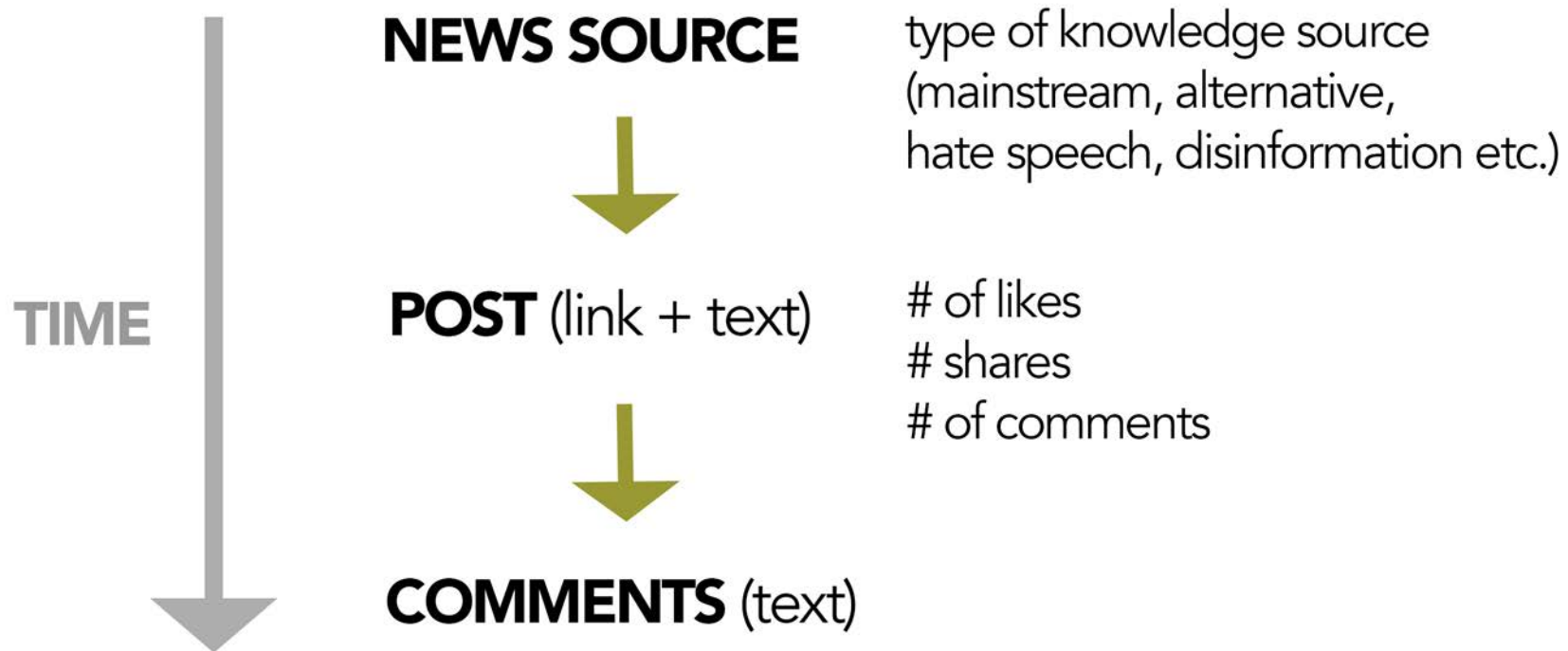


TYPES OF COMMUNICATION

(n = 81457 posts; Rajat Kiinni)

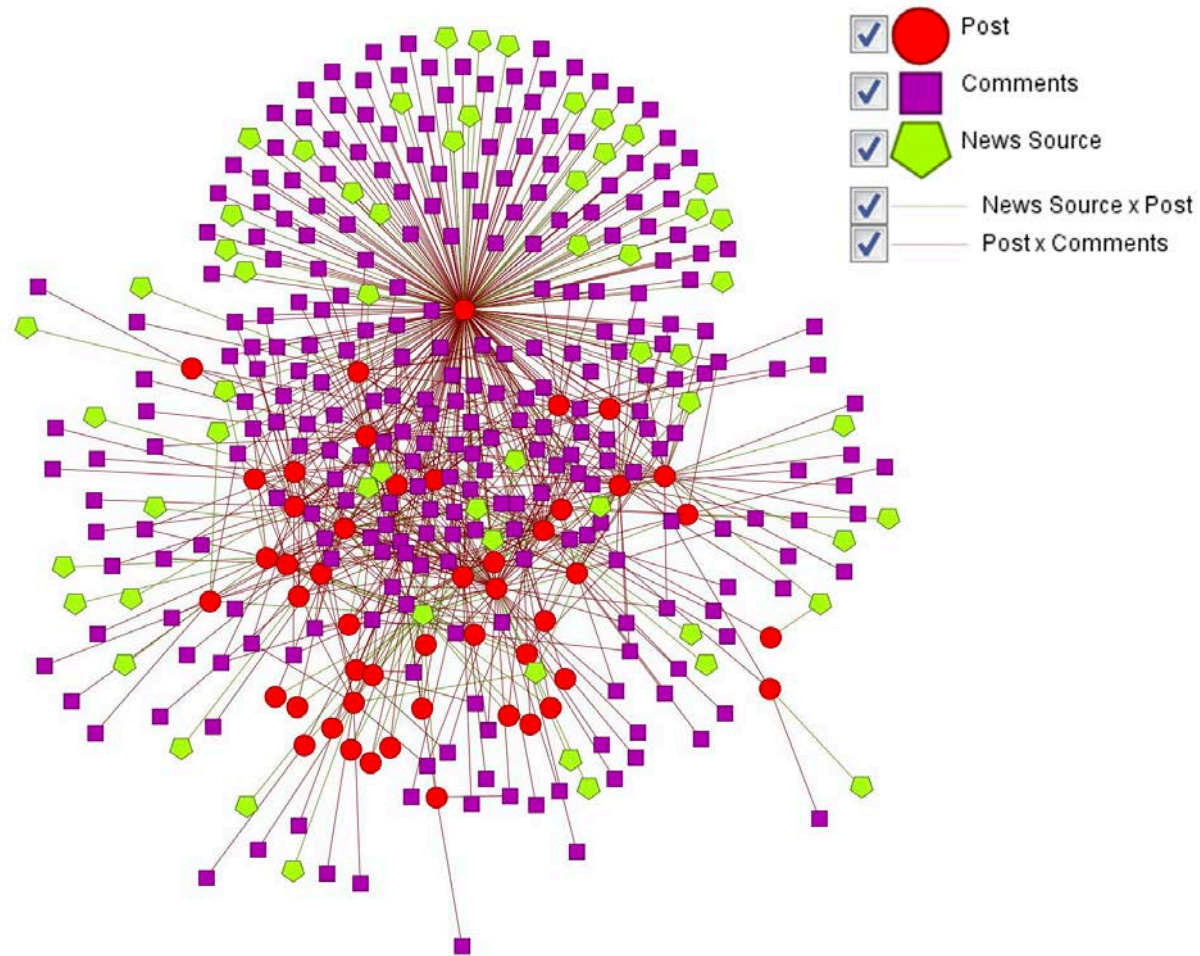


FLOW OF COMMUNICATION



MULTI-LEVEL NETWORK STRUCTURE

(Rajat Kiinni - one day view - by user ID)





SAMPLING STRATEGIES:

- + **Random sample:** baseline
- + **Event-based sample:** Brussels and Nice terrorist attacks.
- + **Purposeful sample:** Different types of information sources - alternative news sites vs. mainstream news.
- + **"Thick" data:** qualitative descriptions selected from the data analysis



DYNAMIC SOCIAL NETWORK ANALYSIS:

+ Where does different kinds of speech originate from? Do these sources change over time?

+ How does information from different knowledge sources (URL) propagate over time in different social networks (posts and comments)?

+ Who are the key entities and their spheres of influence over time?

METHODS: *Static and dynamic social network analysis, near term impact analysis (Gephi & ORA)*



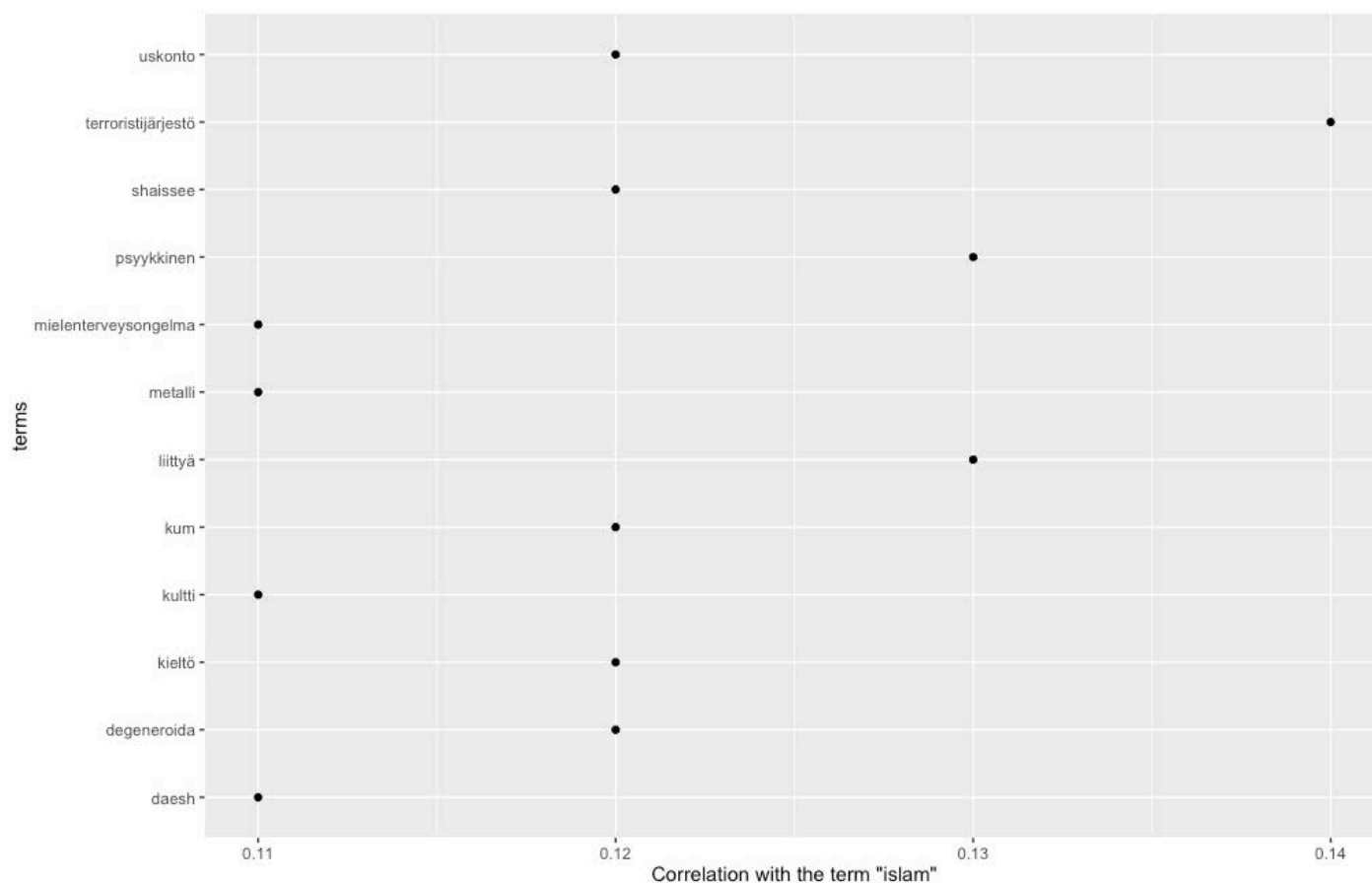
TOPICAL NETWORKS & TEXT MINING

- + What is the relationship between the semantic content in **sources** --> **posts** --> **comments**? How does this change over time?
- + How is this relationship different during events, different types of speech, and in different ideological groups?

METHODS: *Bag of words (BoW) text mining of content for word categorisation (TF-IDF word associations, topic modelling/LDA, supervised classification) I have also lemmatized the text from the Finnish Internet Parsebank to get rid of morphological redundancies in Finnish)*

WORD ASSOCIATIONS (TF-IDF)*

(Rajat Kiinni: n = random sample of 10,000 of 81,457 posts, seed term: "islam")



SOME CORRELATED TERMS:

RELIGION
TERRORIST GROUP
S**T
PSYCHOLOGICAL
MENTAL ILLNESS
METAL
JOIN
DEGENERATE
DAESH

***TF-IDF**, frequency-inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus

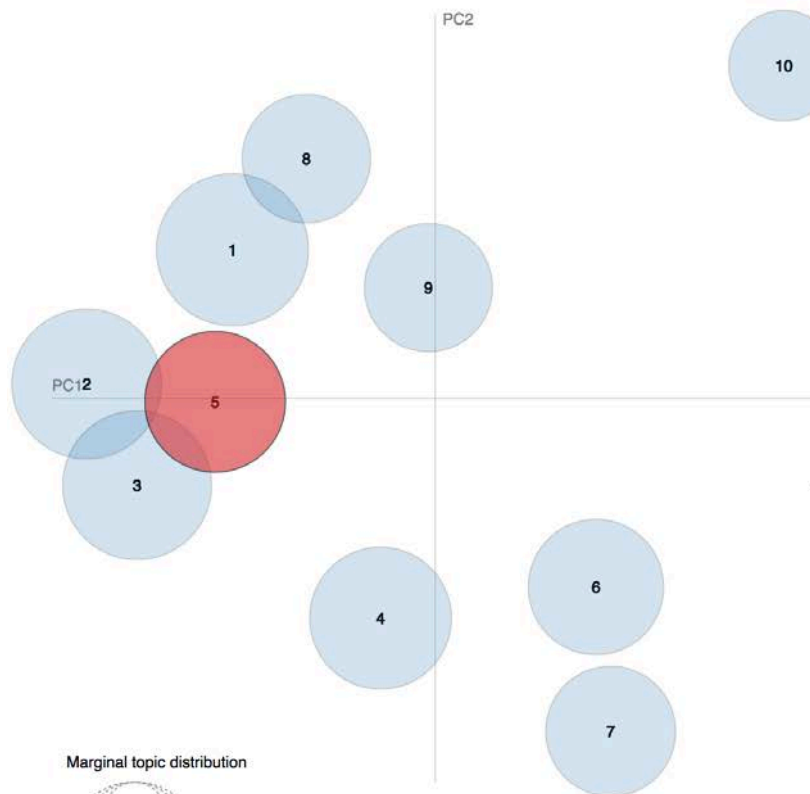
TOPIC MODEL (LDA) RAJAT KIINNI

(n = random sample of 50,000 out of 654,666 comments, 1000 iterations, 10 topics)

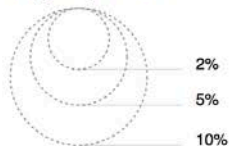
Selected Topic: Previous Topic Next Topic Clear Topic

Slide to adjust relevance metric:⁽²⁾ $\lambda = 0$ 0.0 0.2 0.4 0.6 0.8 1.0

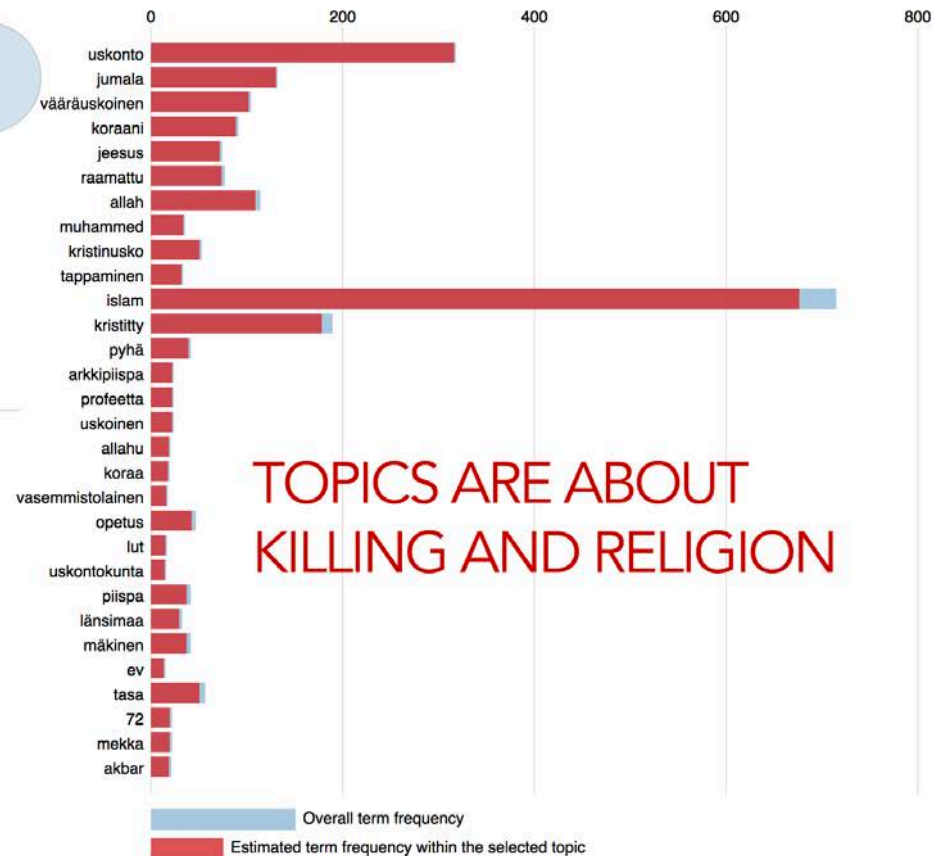
Intertopic Distance Map (via multidimensional scaling)



Marginal topic distribution



Top-30 Most Relevant Terms for Topic 5 (10.6% of tokens)



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)

SUPERVISED LEARNING ("CLASSIFICATION")

TEXT	LABEL
barbaari ottaa äiti ja lähteä koti päin kun täällä olla niin vaikekea . . niin mä tehdä . ."	OTHER
nyt parpaari munata ittes lopullisesti . vituttaa vituttaa kuunnella ? voi resukka ruikku , ruik ."	HATE
kuten tuosta tulla ilmi nuo ei olla koskaan luottaminen , minä olettaa aina että ne valehdella ja tuo juttu vaan vahvistaa asia"	OTHER
jos tommonen sanoi päivä kyllä kattosin kaksi kerta missä kohdata se aurinko olla"	OTHER
" neekeri paskat saada oikein ansio mukaan . aina valehdella joka asia . onnek olla	HATE
" mikä vittu barbaari , idiootti neekeri , lelliä paskahousu ."	HATE
"nolo ."	OTHER
"niin , oisko vaan kannattaa näyttää ne henkkari ja jatkaa matka . ."	OTHER
"hupsis . miten se nyt silleen ."	OTHER

METHODS: Manual labelling of 2000 comments based on terms used in hate speech. Dividing into training and test set (60/40) and testing 8 different machine learning classifiers. Selected **classification and regressions trees** with highest predictive accuracy (F-Score = 0.87)

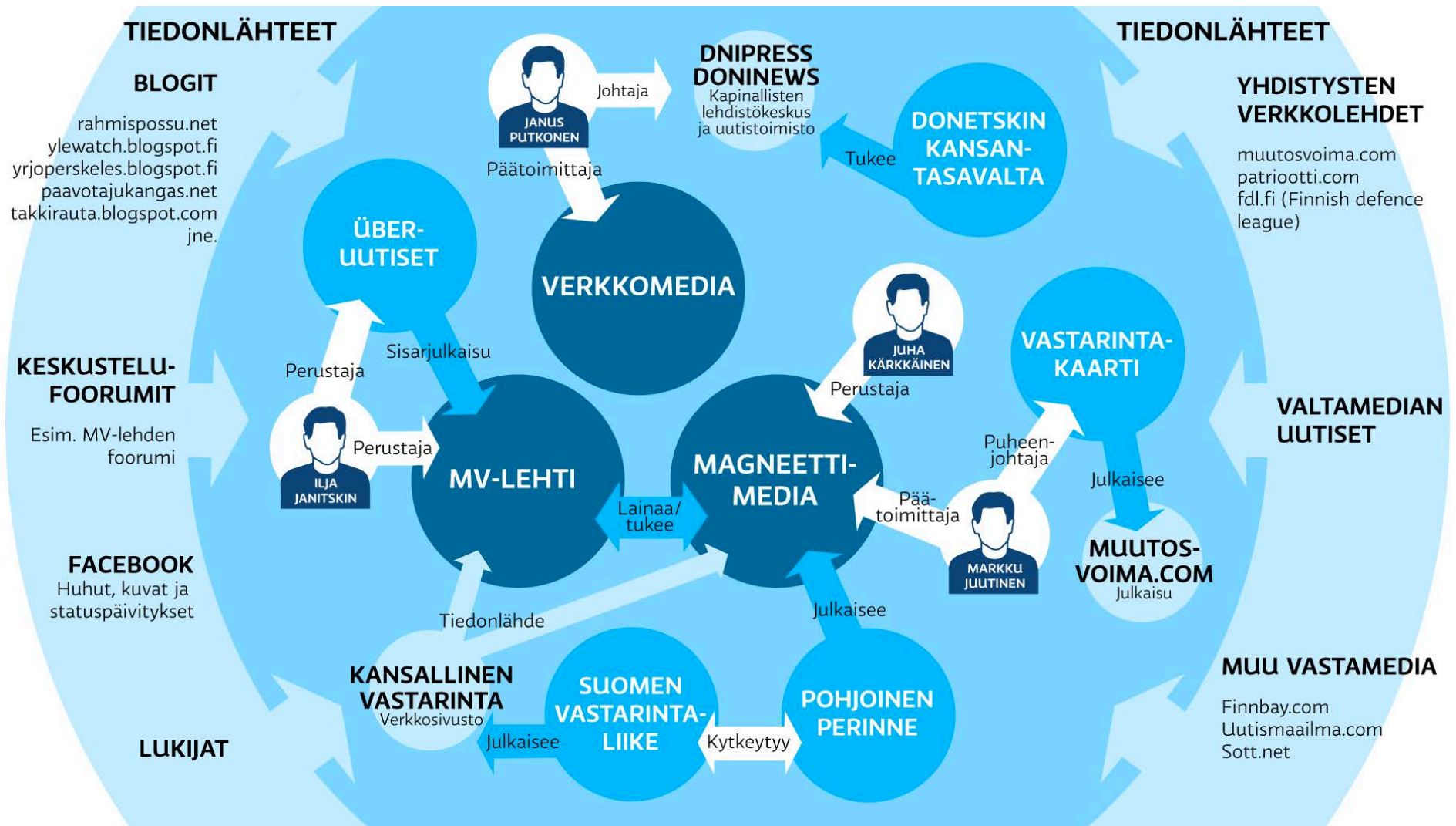


PRELIMINARY FINDINGS:

discoveries from a complex dataset ...

THE INTRICATE XFR ONLINE NETWORKS

(Alternative online news sites and their links to the extreme far right)




THE CASE OF MV-LEHTI

(MV-lehti = most popular online news site promoting anti-immigration sentiment)

Police to seek European arrest warrant for right-wing site founder

Helsinki police want to question Ilja Janitskin, a Finn living in Spain who is suspected of a range of crimes connected to his anti-immigrant website.

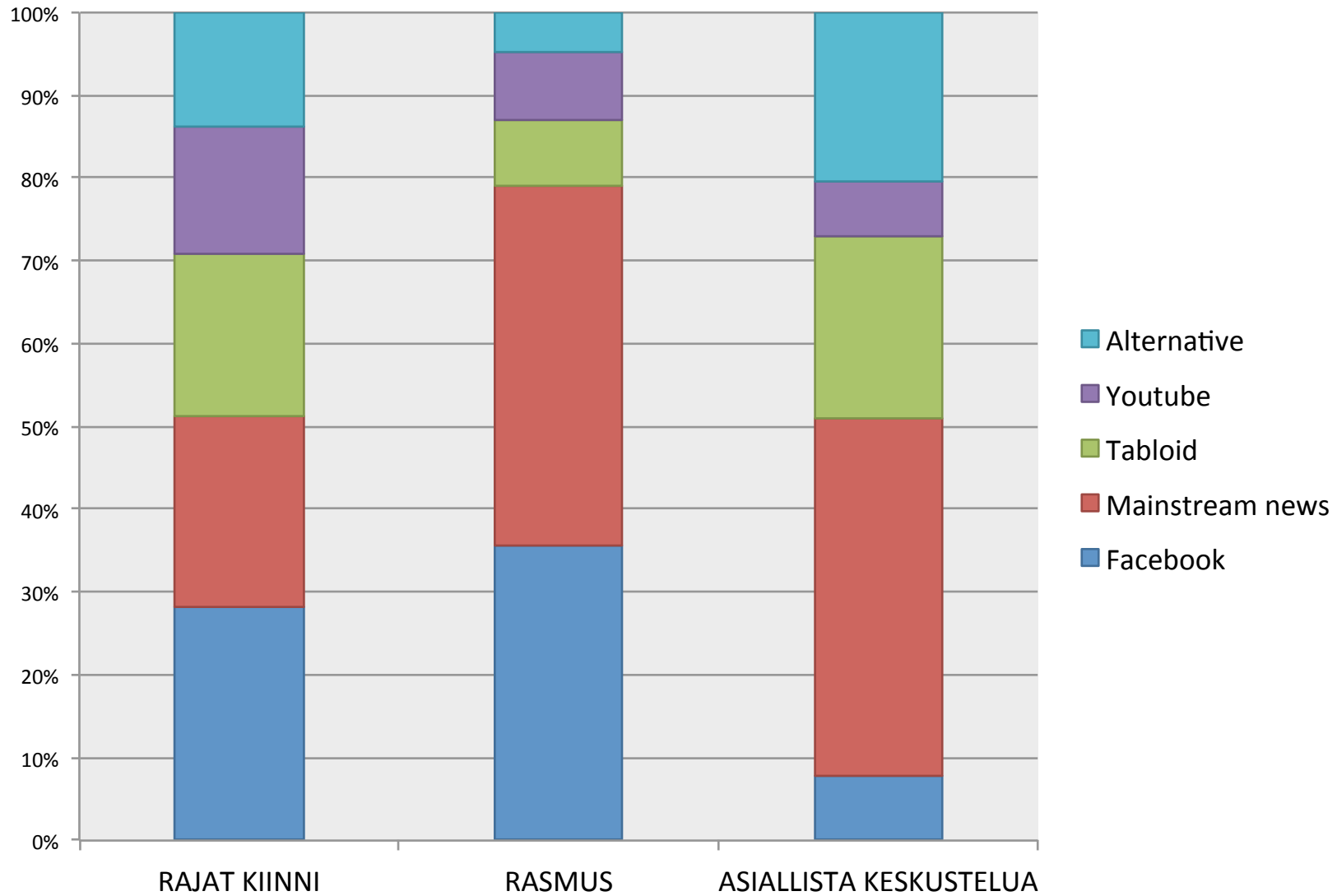
 Recommend 104 people recommend this. Be the first of your friends.



The site features an array of purported news stories with an extreme right-wing slant. Image: Yle

WHAT ARE THE NEWS SOURCES?

(Types of online sources shared by the three groups)

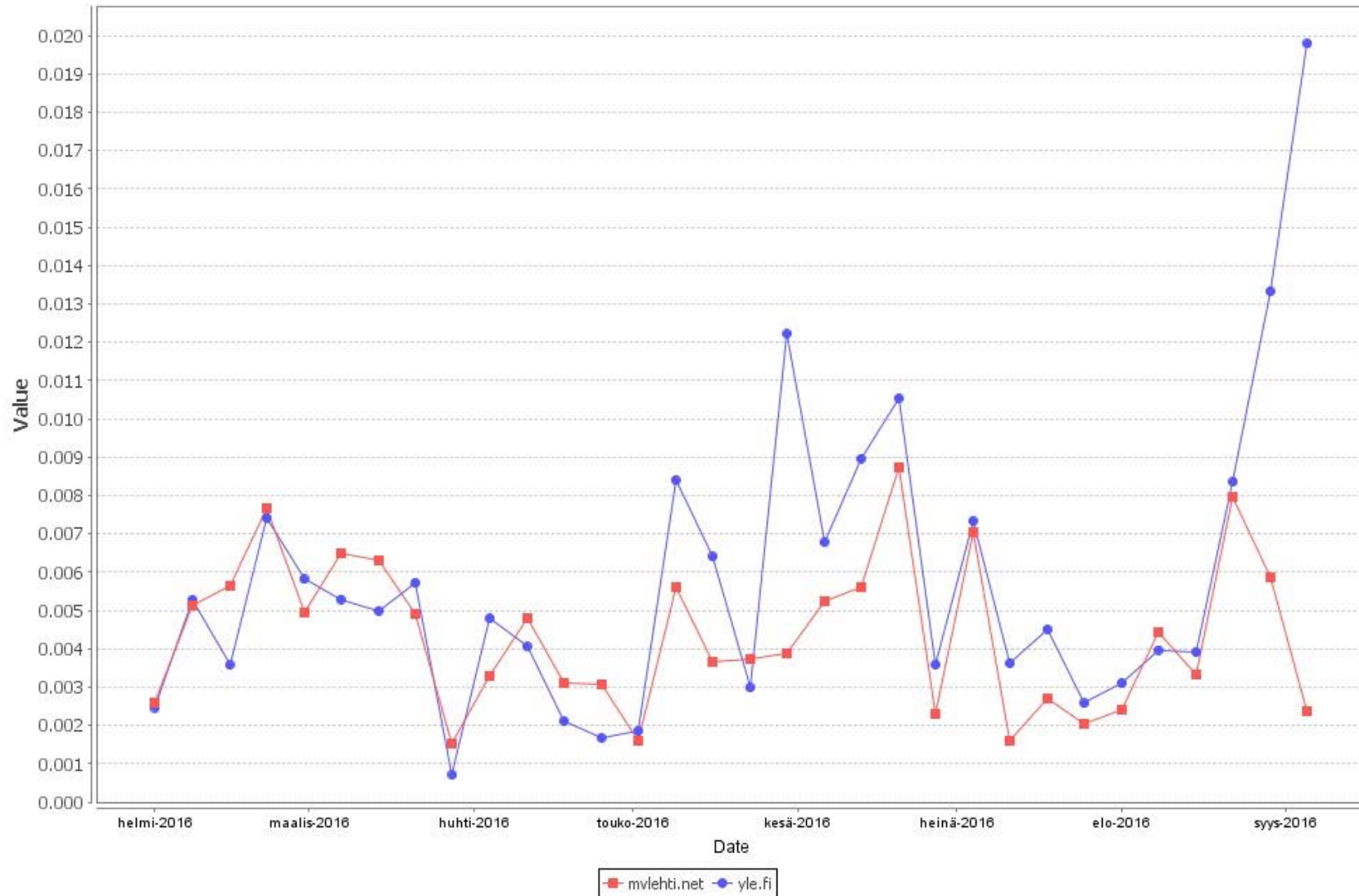


MOST SHARED URL (RAJAT KIINNI)

www.facebook.com	11455	Social media
www.youtube.com	6181	video
www.iltalehti.fi	4092	tabloid
mvlehti.net	3513	alternative online
yle.fi	2916	mainstream online news
www.iltasanomat.fi	2656	tabloid
www.mtv.fi	1871	mainstream television channel
riippumatonmedia.com	1065	alternative online
www.hs.fi	995	mainstream newspaper
www.verkkouutiset.fi	961	alternative online
paavotajukangas.com	839	alternative far right blogger
uberuutiset.fi	621	alternative online
www.finnleaks.net	621	alternative online
www.uusisuomi.fi	620	mainstream newspaper
www.helsinginuutiset.fi	563	local newspaper
www.express.co.uk	532	UK tabloid
www.suomenuutiset.fi	399	True Finns newspapers
www.dailymail.co.uk	333	UK tabloid

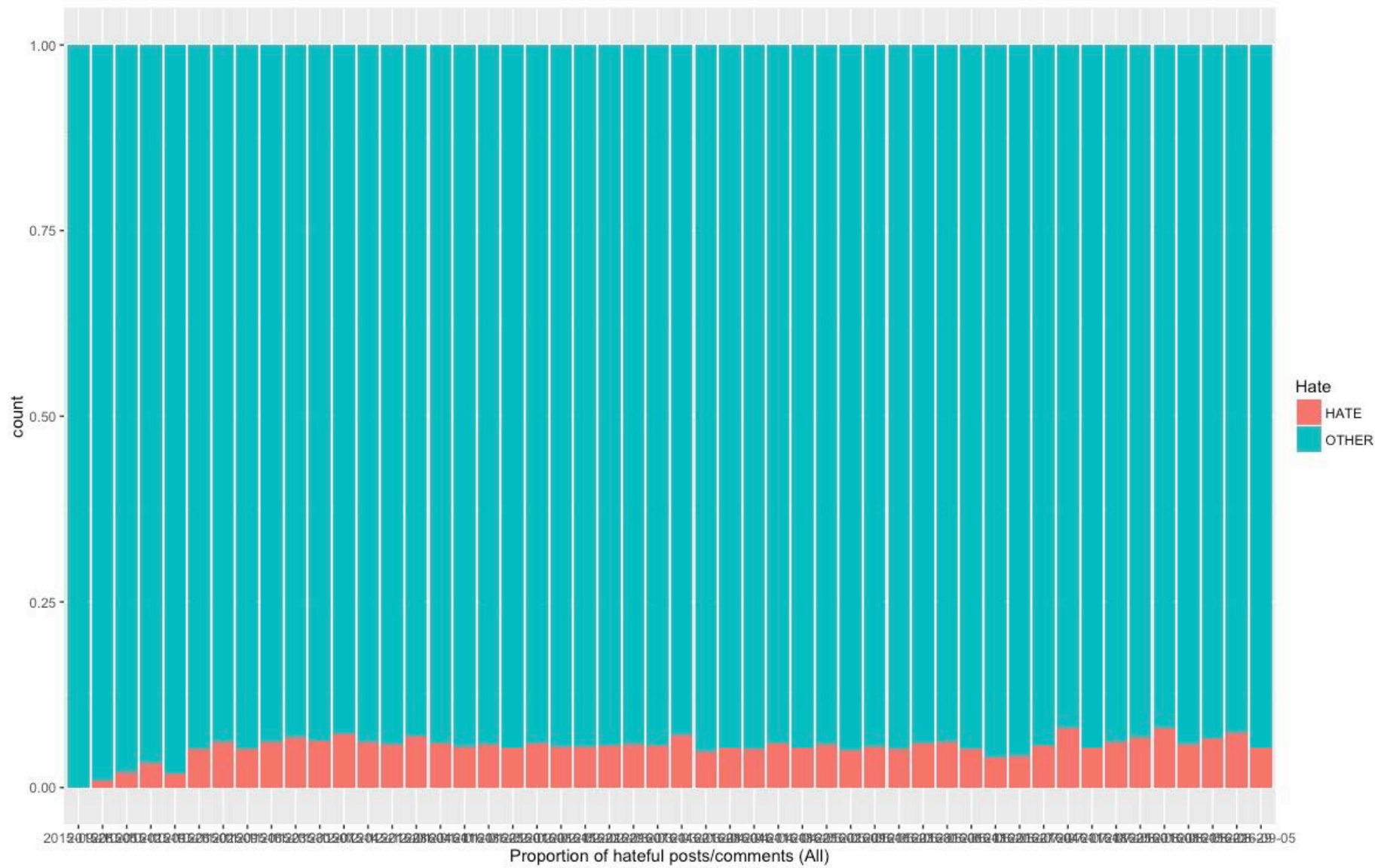
HOW IMPORTANT IS XFR NEWS?

(comparative centrality of **MV-Lehti** and **YLE** in the network over time)



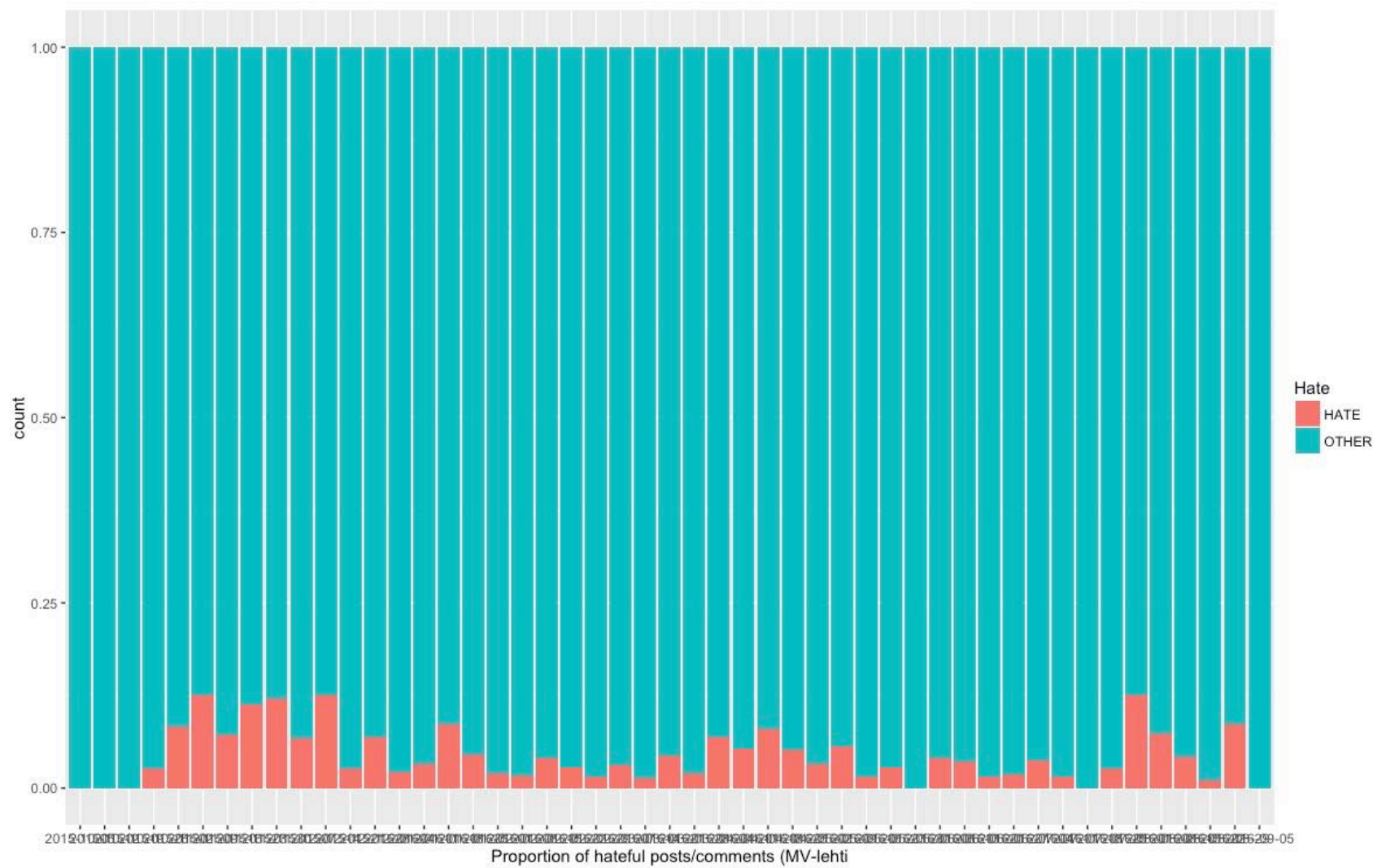
HOW MUCH HATE IS THERE?

(n = 654,666 comments and 81,567 posts - Rajat Kiinni all)



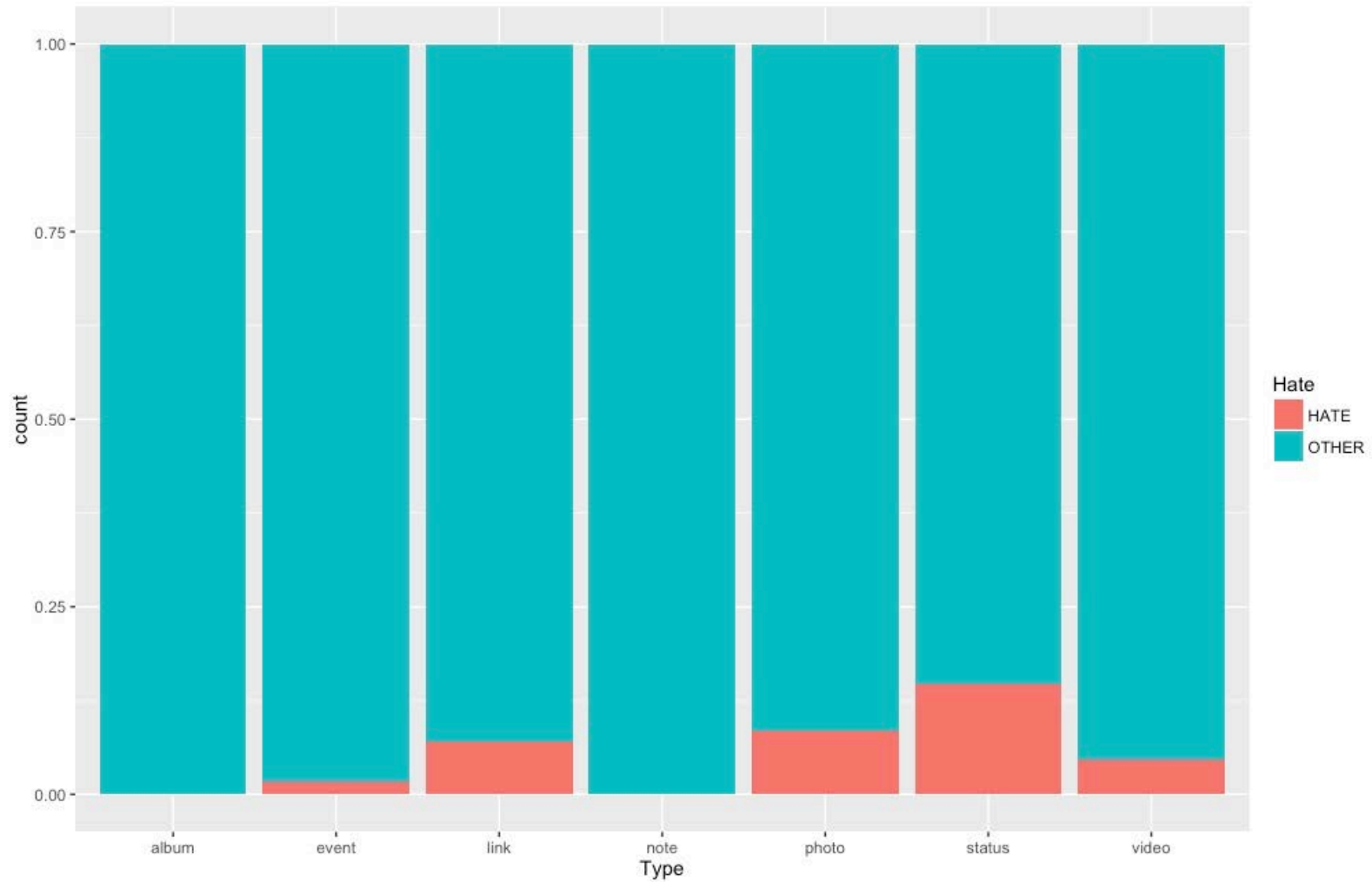
HOW MUCH HATE IS THERE?

(Rajat Kiinni MV-lehti shares (main anti-immigration news))

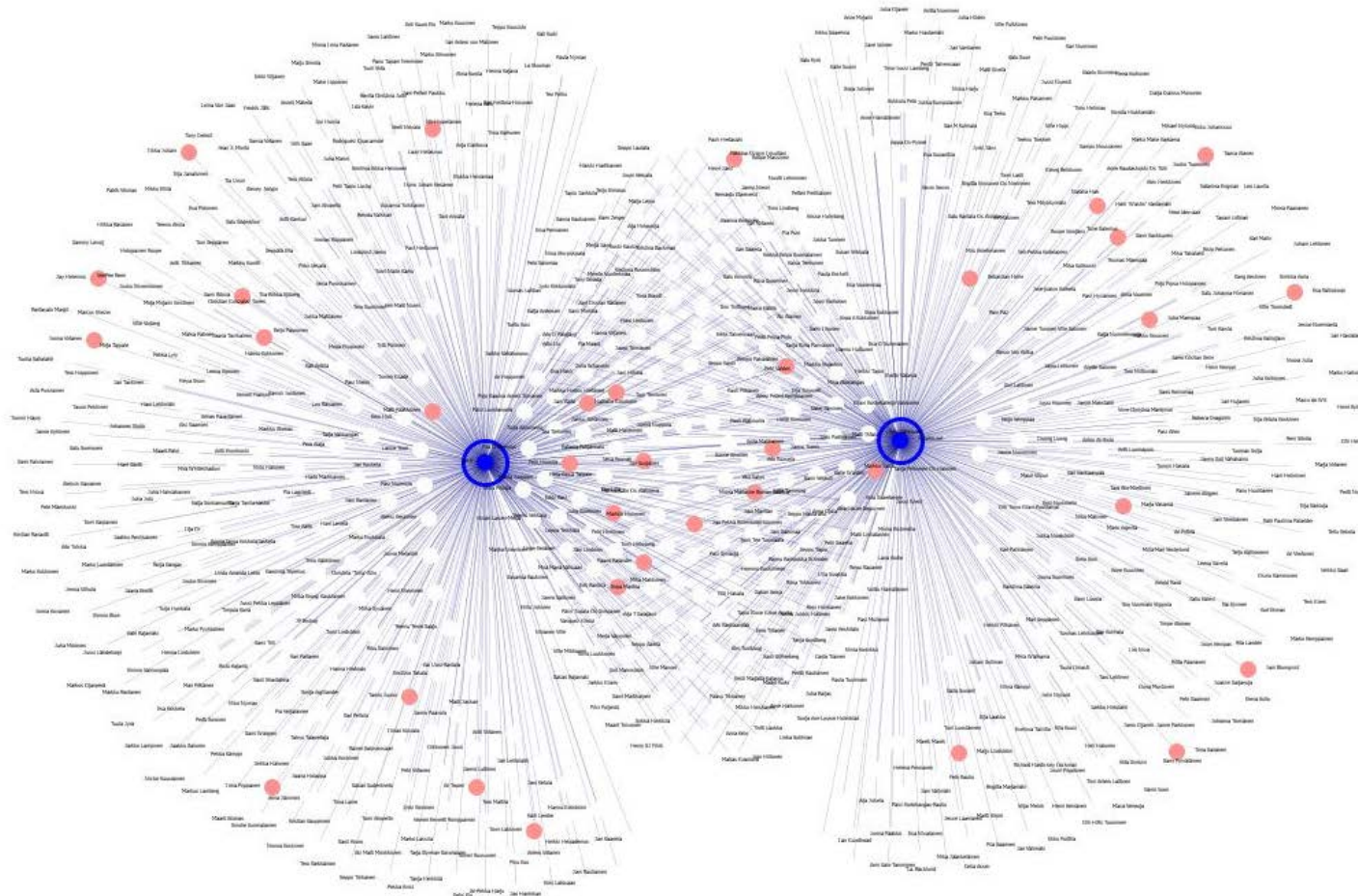


WHERE IS HATE FOUND?

(n = 81,567 posts - Rajat Kiinni)



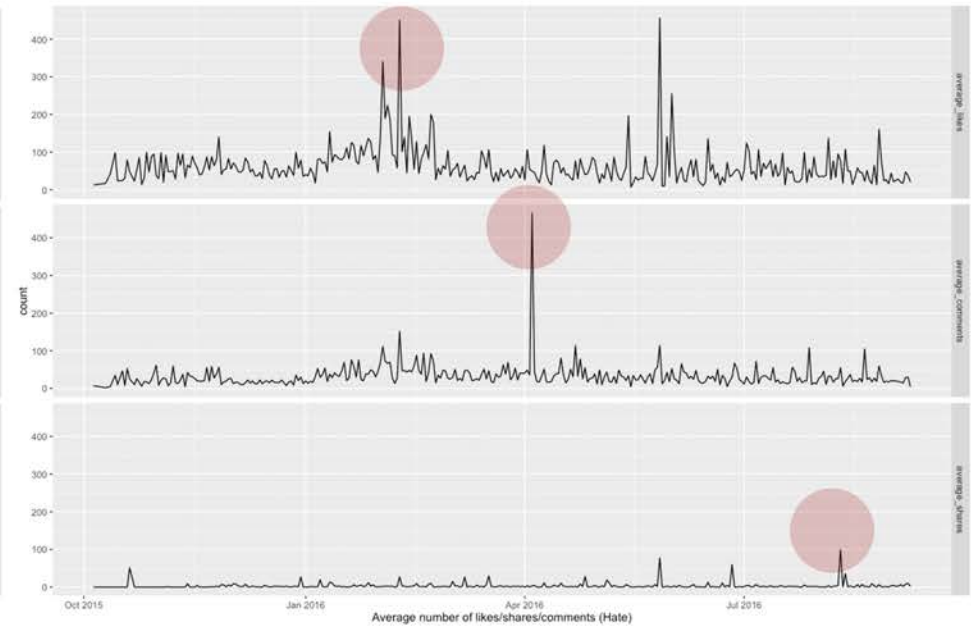
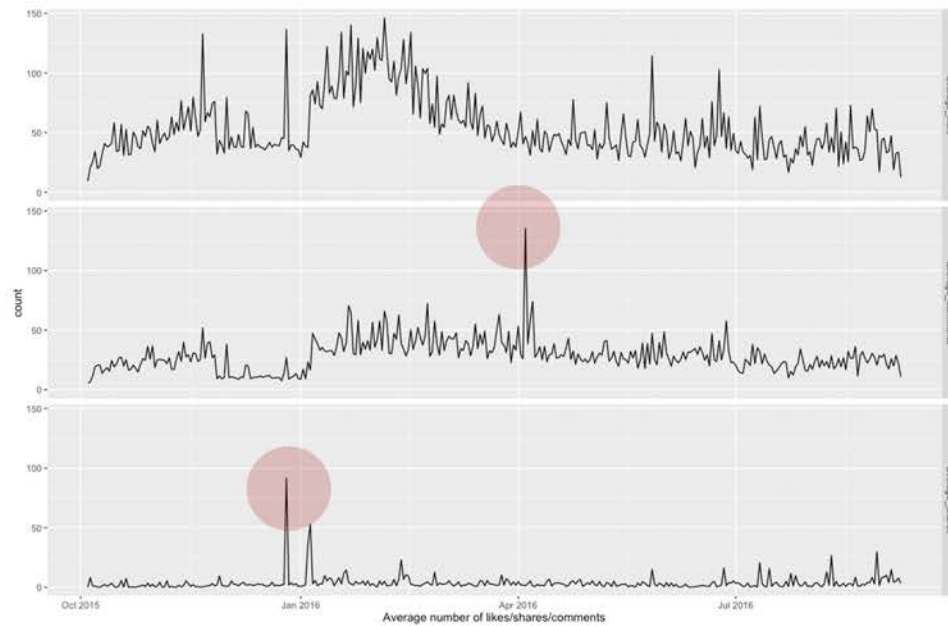
PRESENCE OF HATE IN THE NETWORK? (spheres of influence - MV-Lehti and YLE)



HOW IS DIFFERENT CONTENT REACTED TO?

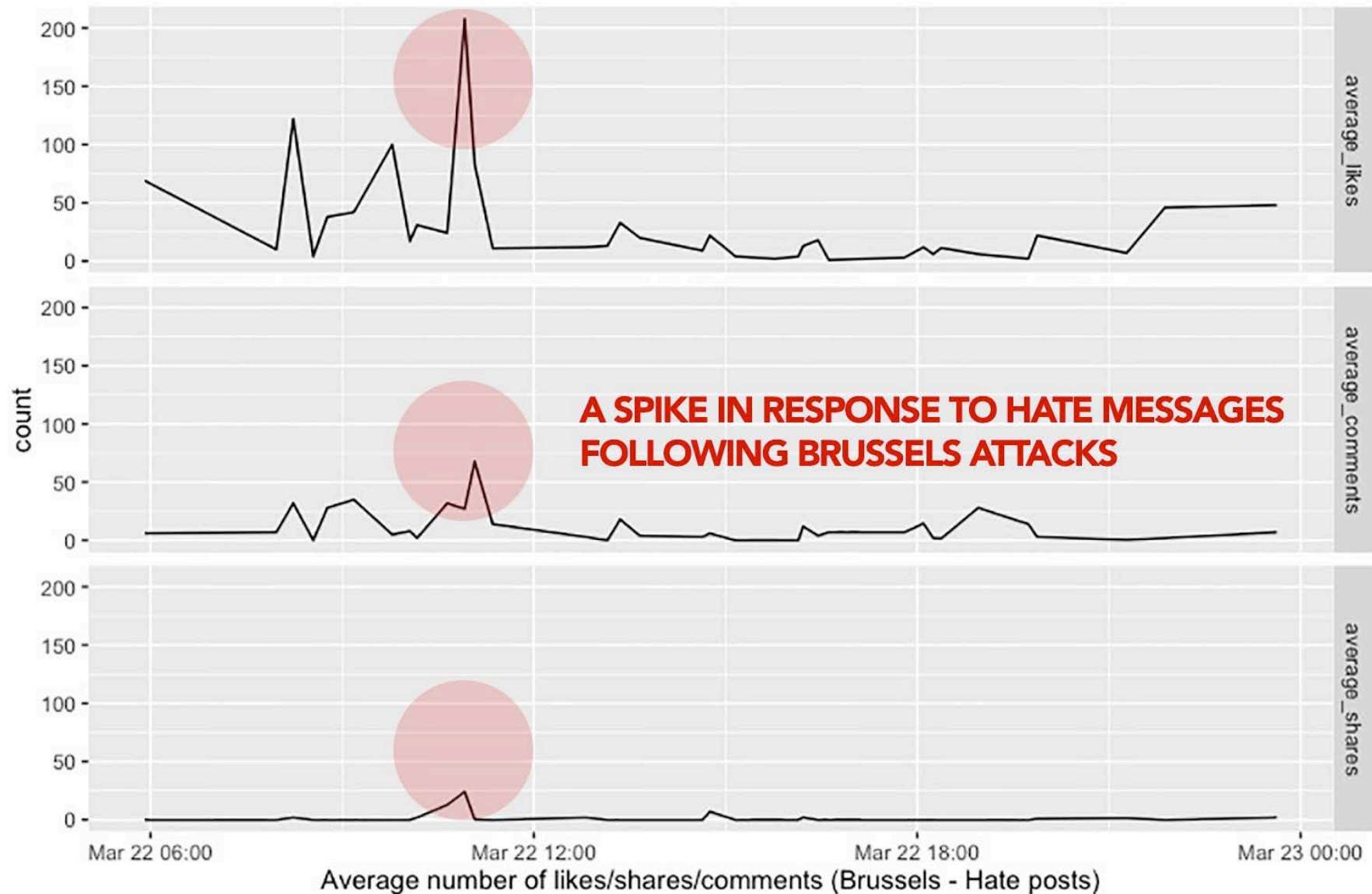
(Average number of daily likes/comments/shares - Rajat Kiinni - All vs. hate)

SUDDEN SPIKES OF ACTIVITY AROUND POPULAR CONTENT



TERRORIST ATTACKS GET AN ANGRY RESPONSE

(Likes/comments/shares - Rajat Kiinni - Brussels - Hate)



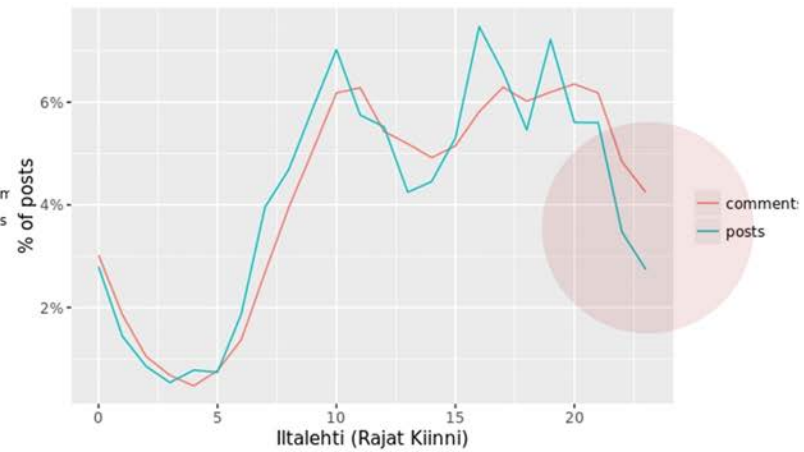
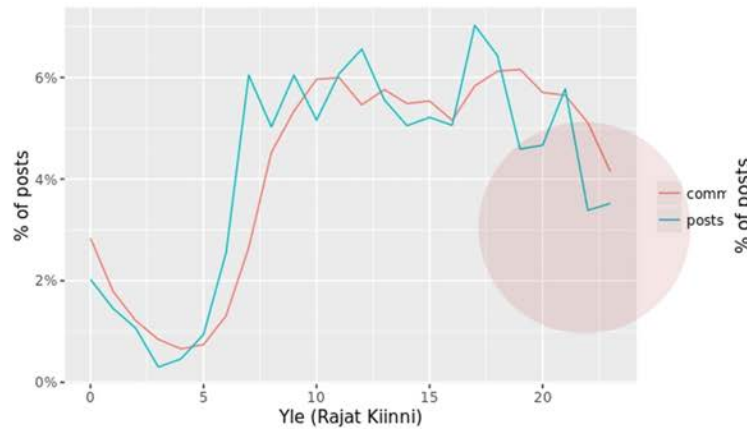
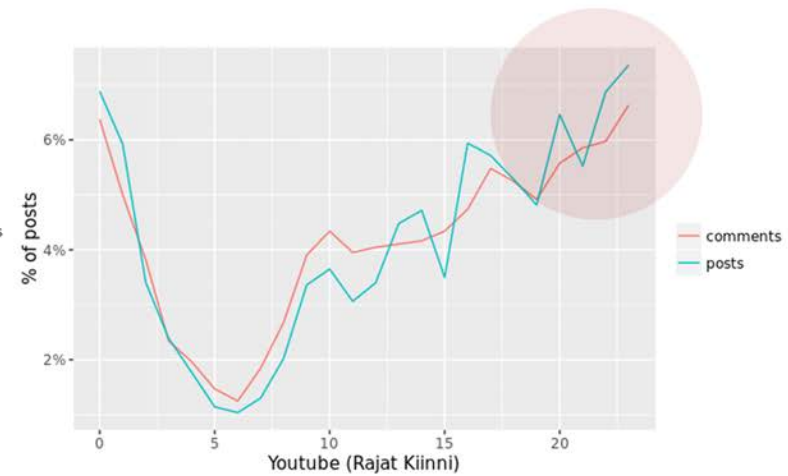
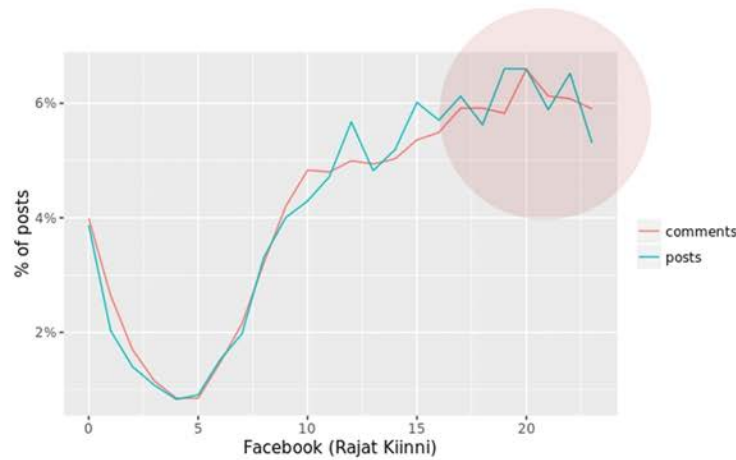
HATE MESSAGES DO SEEM **GET MORE RESPONSE?**

(Likes/comments/shares - Rajat Kiinni)

	RAJAT KIINNI (ALL)	RAJAT KIINNI (HATE)
Likes (mean)	26.98747806	38.28921569
Likes (median)	10	19
Comments (mean)	8.381759701	14.26364024
Comments (median)	3	7
Shares (mean)	1.812490025	1.963768116
Shares (median)	0	0

WHAT TIME ARE THEY POSTED?

(time of day - Rajat Kiinni)





CONCLUSIONS (AND CONJECTURES)

So what should we make of all this for hate speech research ...



SOME TENTATIVE FINDINGS:

+ **ALTERNATIVE MEDIA BUBBLES:** While the communicative dynamics of different ideological groups are similar, there are significant differences in the sources of knowledge they access. The anti-immigration / far activist have set up a parallel news network online where anti-immigrant news are commonly shared from. The anti-racists are more closer to mainstream news sources.

+ **ECHO CHAMBERS & COUNTER-SPEECH.** The different ideological groups are separated into opposing clusters, which often talk about each other, but seldom *to each other*. The *Asiallista Keskustelua* Facebook group - set up to bridge this gap - is an interesting effort to have a dialogue across these echo chambers. In practice, however, it has gradually become a space for the anti-immigration side to express their viewpoints and frustration that they are not taken seriously.

+ OPPOSITIONAL READINGS: There is no relationship between a news source and the amount of hateful comments on it. Hateful comments are more related to individual rather than the source of news. The mainstream news is routinely used to express a point even if the reading is oppositional.

+ XFR PROPAGANDA IS NOT HATE SPEECH: "Hate speech" is also not closely linked to extreme far right propaganda. Hate speech seems to be more of an *emotional response* by people venting or hoping to get a "high five" response from in-group or, alternatively, "trigger" a response from the other group;

+ POLARISATION VS. CONTENT: Far right propaganda (such as the case of *MV-Lehti*), on the contrary, is a more concerted (and dangerous) effort to provide an alternative political framing and to capitalise on people's angry response to events such as the terrorist attacks in Europe. Focusing only on the *vitriolic content* of hate speech alone thus perhaps obfuscates a more complex political polarisation, social conflict and extreme far right radicalisation taking place in EU that we should be worried about?



THANK YOU!

matti.pohjonen@voxpath.eu

[@objetpetitm](#)