



3rd Biennial Conference
Amsterdam August 2018

Decoding the Alt-Right

A Machine Learning Project

Simon Roth¹ & Fabio Votta²

2018-20-08

¹ E-Mail: nomis.roth@gmx.com

² E-Mail: fabio.votta@gmail.com

Slides: <https://decoding-the-altright.netlify.com/>

Goal:

Identify Alt-Right commentary on social media and its spread into mainstream media

Get Started

What is the Alt-Right?

Milo Yiannopoulos' characterization of the Alt-Right:

*The alt-right is a movement born out of the youthful, subversive, **underground edges of the internet**. 4chan and 8chan are hubs of alt-right activity. For years, members of these forums – political and non-political – have delighted in **attention-grabbing, juvenile pranks**. Long before the alt-right, 4channers turned **trolling** the national media into an in-house sport.*

Marwick and Lewis 2017

*Term coined by Richard Spencer to describe a version of **white nationalism** that positions itself as comprised of **younger, wealthier, and better-educated individuals** than traditional white supremacist groups like the Klu Klux Klan. Deeply entwined with social media and staples of **internet culture, like memes**. Believes that “**political correctness**” threatens individual liberty. Example: Jared Taylor of the American Renaissance, Mike Cernovich of Danger and Play.*

Hawley 2017

*Using the loosest definition, we could say the Alt-Right includes anyone with right-wing sensibilities that **rejects the mainstream conservative movement**. But there are certain common, perhaps universal attitudes within the Alt-Right. The Alt-Right is **fundamentally concerned with race**. At its core, the Alt-Right is a white-nationalist movement, even if many (perhaps most) of the people who identify with the Alt-Right do not care for that term. The most energetic and significant figures of the movement want to see the creation of a **white ethnostate** in North America.*

Allum Bokhari and Milo Yiannopoulos, “An Establishment Conservative’s Guide to the Alt-Right,” Breitbart, March 29, 2016, <http://www.breitbart.com/tech/2016/03/29/an-establishment-conservatives-guide-to-the-alt-right/>.

Non-ironic Nazism masquerading as ironic Nazism.

~Andrew Anglin, Founder of the *The Daily Stormer*

What is the Alt-Right?

A group of loosely aligned individuals and subgroups that predominately consist of white, young & male frequent internet user. Although adherence to these beliefs varies, these are some common themes within the Alt-Right:

What is the Alt-Right?

A group of loosely aligned individuals and subgroups that predominately consist of white, young & male frequent internet user. Although adherence to these beliefs varies, these are some common themes within the Alt-Right:

- Antipathy towards both liberals (“*social justice warriors*”) and conservatives (“*cuckservatives*”)

What is the Alt-Right?

A group of loosely aligned individuals and subgroups that predominately consist of white, young & male frequent internet user. Although adherence to these beliefs varies, these are some common themes within the Alt-Right:

- Antipathy towards both liberals (“*social justice warriors*”) and conservatives (“*cuckservatives*”)
- Fighting a “*culture war*” against perceived liberal consensus within the establishment

What is the Alt-Right?

A group of loosely aligned individuals and subgroups that predominately consist of white, young & male frequent internet user. Although adherence to these beliefs varies, these are some common themes within the Alt-Right:

- Antipathy towards both liberals (“*social justice warriors*”) and conservatives (“*cuckservatives*”)
- Fighting a “*culture war*” against perceived liberal consensus within the establishment
- Belief that freedom of speech is being threatened by “*political correctness*”

What is the Alt-Right?

A group of loosely aligned individuals and subgroups that predominately consist of white, young & male frequent internet user. Although adherence to these beliefs varies, these are some common themes within the Alt-Right:

- Antipathy towards both liberals (“*social justice warriors*”) and conservatives (“*cuckservatives*”)
- Fighting a “*culture war*” against perceived liberal consensus within the establishment
- Belief that freedom of speech is being threatened by “*political correctness*”
- Strong animosity towards feminism, multiculturalism and immigration (“*cultural marxism*”)

What is the Alt-Right?

A group of loosely aligned individuals and subgroups that predominately consist of white, young & male frequent internet user. Although adherence to these beliefs varies, these are some common themes within the Alt-Right:

- Antipathy towards both liberals (“*social justice warriors*”) and conservatives (“*cuckservatives*”)
- Fighting a “*culture war*” against perceived liberal consensus within the establishment
- Belief that freedom of speech is being threatened by “*political correctness*”
- Strong animosity towards feminism, multiculturalism and immigration (“*cultural marxism*”)
- Belief in biological differences in between races and genders that justifies separate treatment

What is the Alt-Right?

A group of loosely aligned individuals and subgroups that predominately consist of white, young & male frequent internet user. Although adherence to these beliefs varies, these are some common themes within the Alt-Right:

- Antipathy towards both liberals (“*social justice warriors*”) and conservatives (“*cuckservatives*”)
- Fighting a “*culture war*” against perceived liberal consensus within the establishment
- Belief that freedom of speech is being threatened by “*political correctness*”
- Strong animosity towards feminism, multiculturalism and immigration (“*cultural marxism*”)
- Belief in biological differences in between races and genders that justifies separate treatment
- Ambiguous and troll-like behavior as to ensure *plausible deniability*

What is the Alt-Right?

The Alt-Right isn't a single cohesive movement. It should rather be understood as a collection of subgroups that share similar talking points as well as tactics and beliefs.

The Alt-Right meanders between sub-groups that sometimes have very different agendas or are more or less committed to their beliefs.



What is the Alt-Right?

Neo-Nazi Movements



Neo-Nazi Movements

White Nationalists & Supremacists

Alt-Right

Conspiracy Theorists

Online Trolls

Menosphere

Anti-Feminist & "Skeptic" Community

What is the Alt-Right?

White Nationalists &
Supremacists



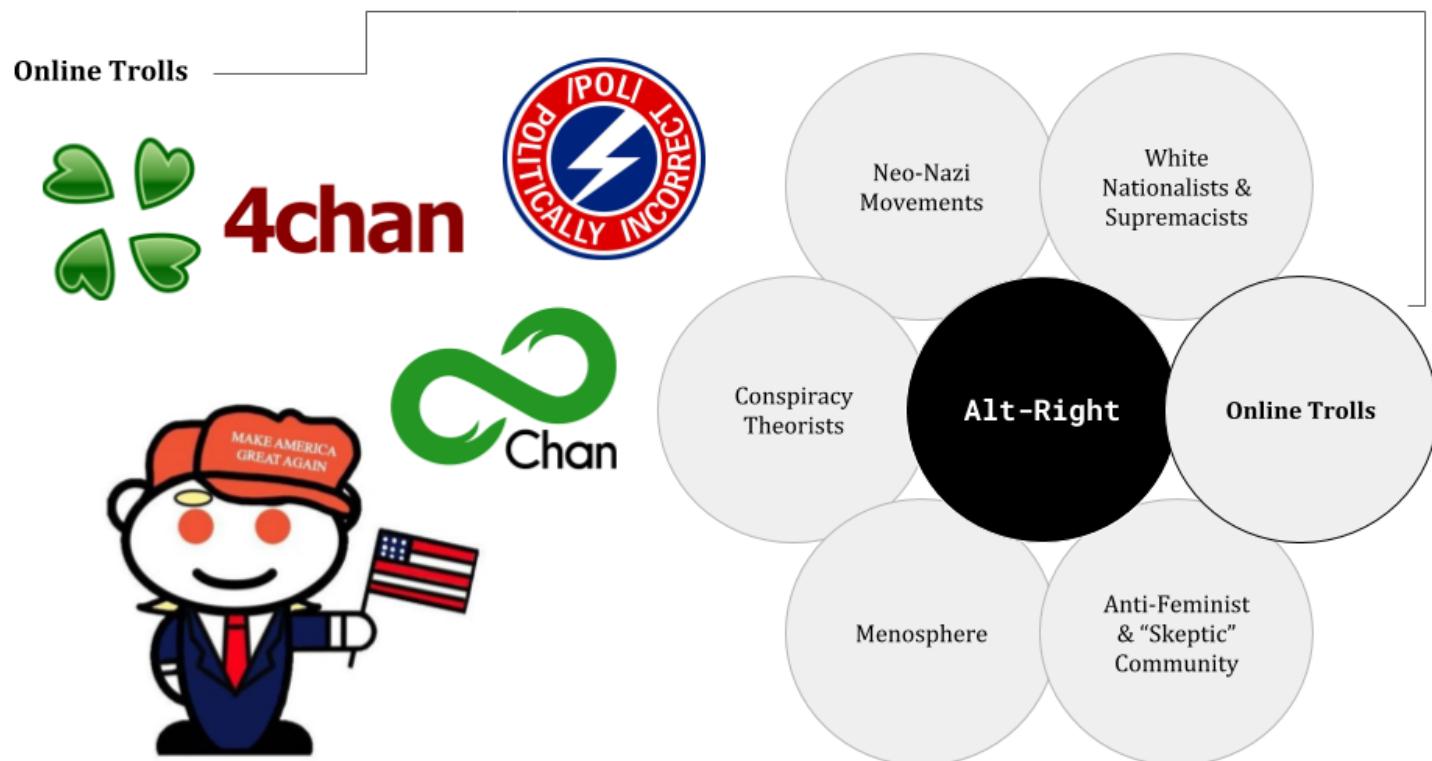
AMERICAN RENAISSANCE



NPI THE NATIONAL POLICY INSTITUTE



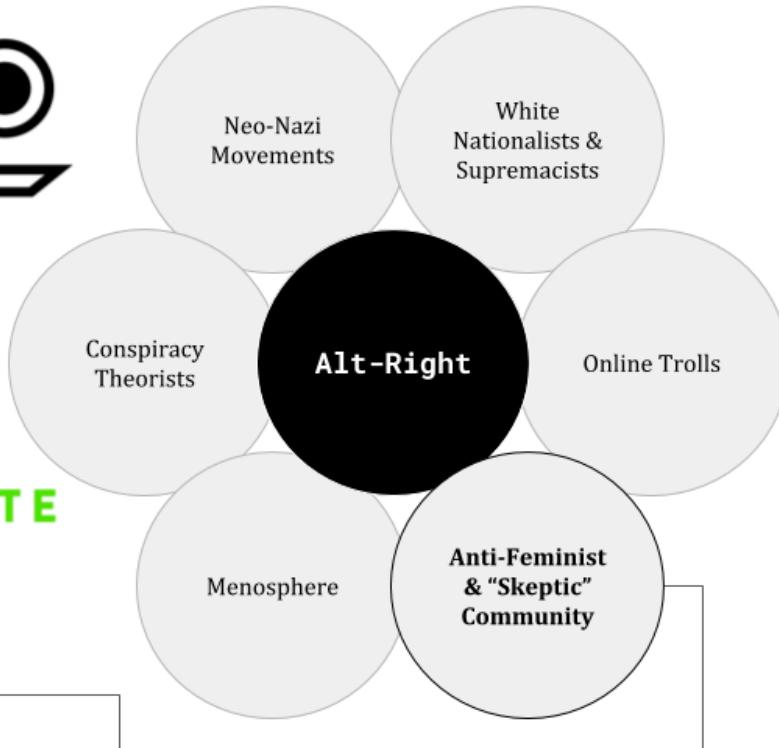
What is the Alt-Right?



What is the Alt-Right?



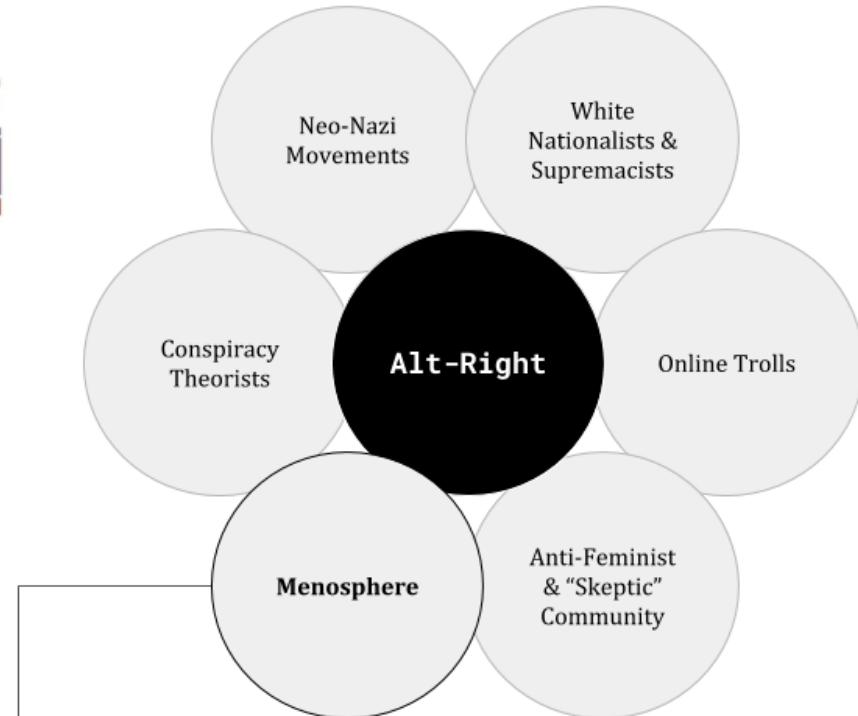
Anti-Feminist & "Skeptic" Community



What is the Alt-Right?



Menosphere



What is the Alt-Right?



Conspiracy Theorists



Data Collection

What data did we collect?

What data did we collect?

- Text data from three Platforms: Facebook, Twitter and YouTube



What data did we collect?

- Text data from three Platforms: Facebook, Twitter and YouTube



- Comments from 36 Alt-Right connected profiles/YouTube channels

What data did we collect?

- Text data from three Platforms: Facebook, Twitter and YouTube



- Comments from 36 Alt-Right connected profiles/YouTube channels
- Comments from 6 Media Outlets (CNN, MSNBC, ABC, Fox News, Washington Post, New York Times)



The Washington Post



The New York Times



What data did we collect?

- Text data from three Platforms: Facebook, Twitter and YouTube



- Comments from 36 Alt-Right connected profiles/YouTube channels
- Comments from 6 Media Outlets (CNN, MSNBC, ABC, Fox News, Washington Post, New York Times)



The Washington Post



The New York Times



- Comments from the entire year 2017

What data did we collect?

- Text data from three Platforms: Facebook, Twitter and YouTube

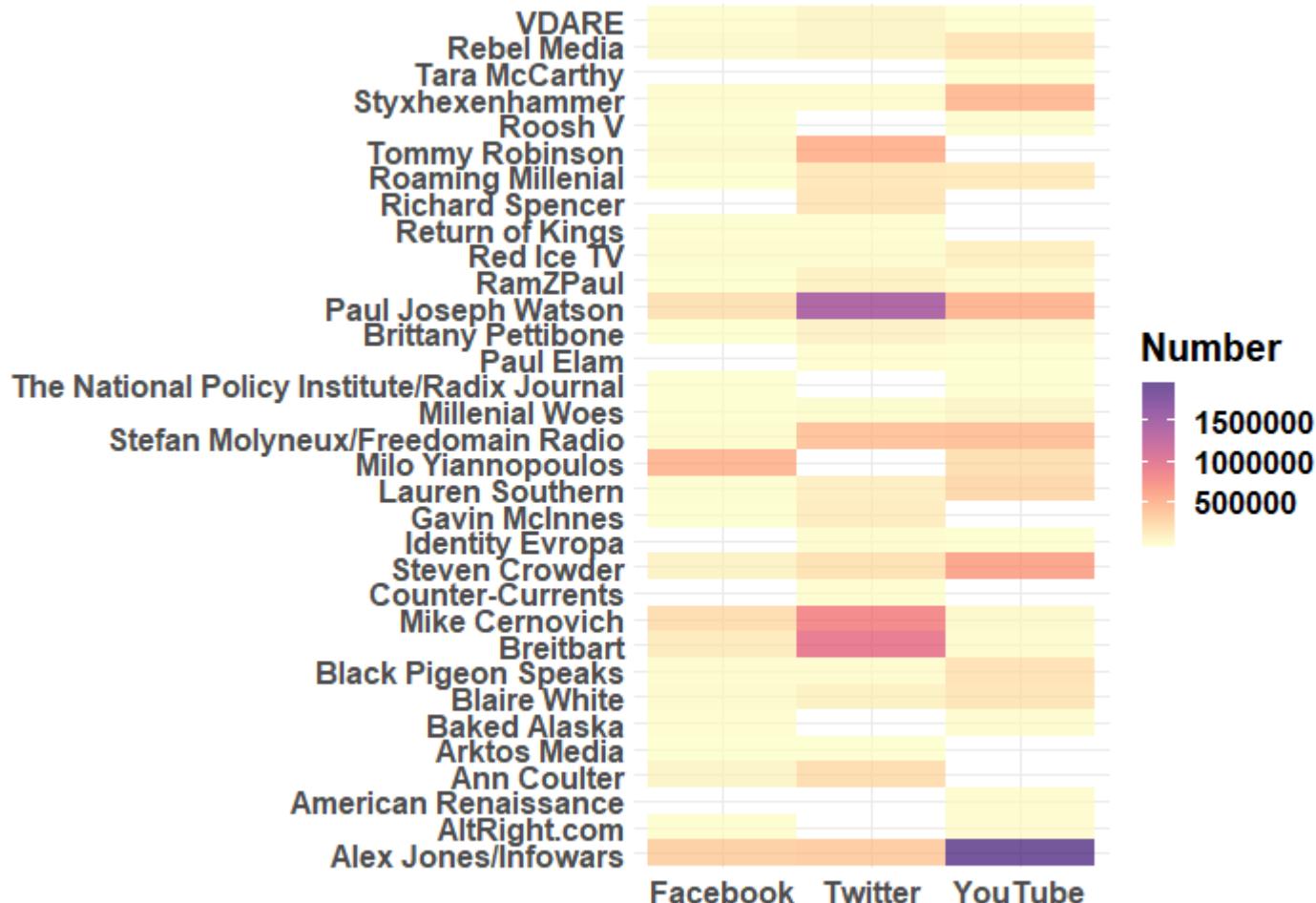


- Comments from 36 Alt-Right connected profiles/YouTube channels
- Comments from 6 Media Outlets (CNN, MSNBC, ABC, Fox News, Washington Post, New York Times)

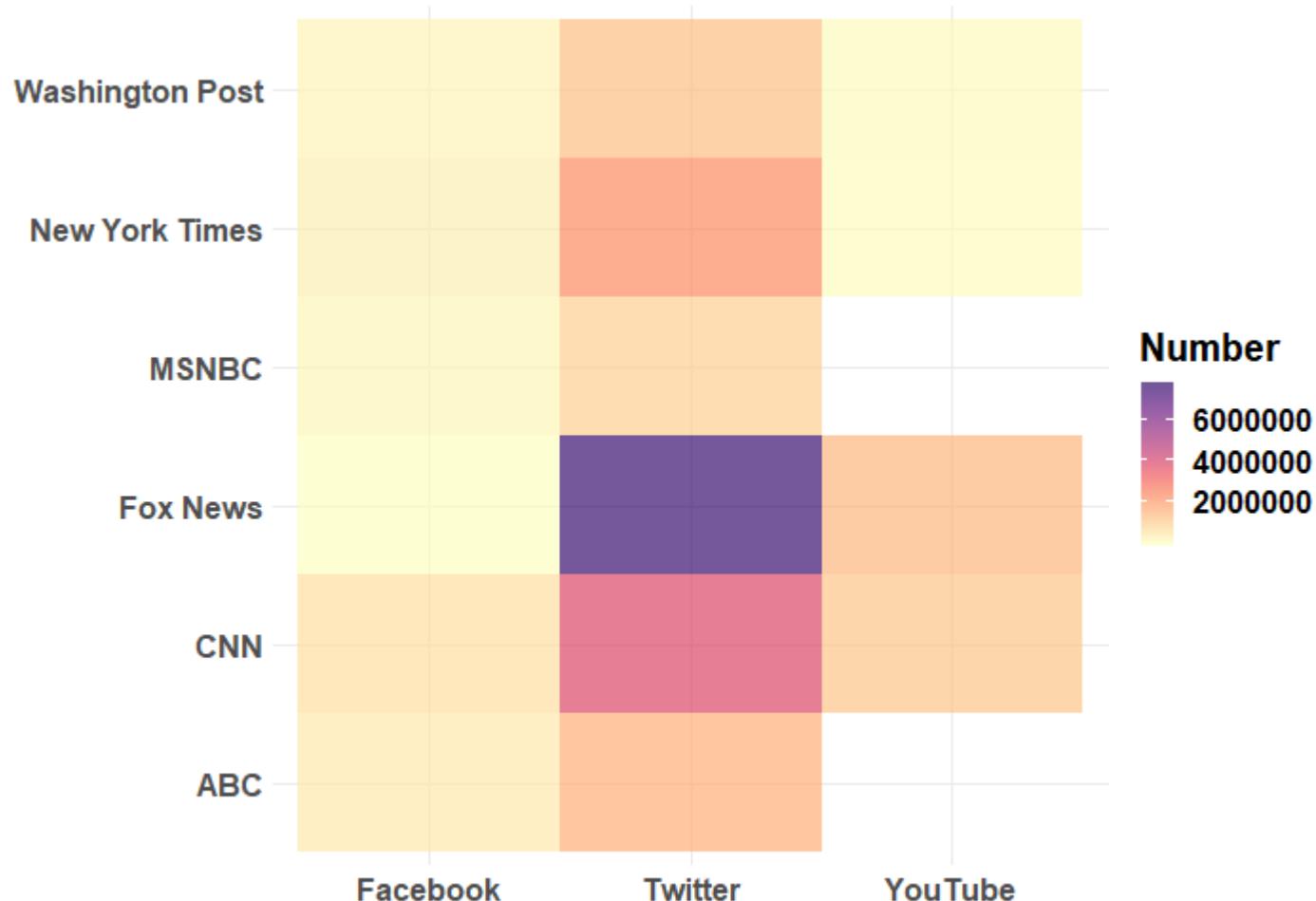


- Comments from the entire year 2017
- In total over 30 million comments/tweets

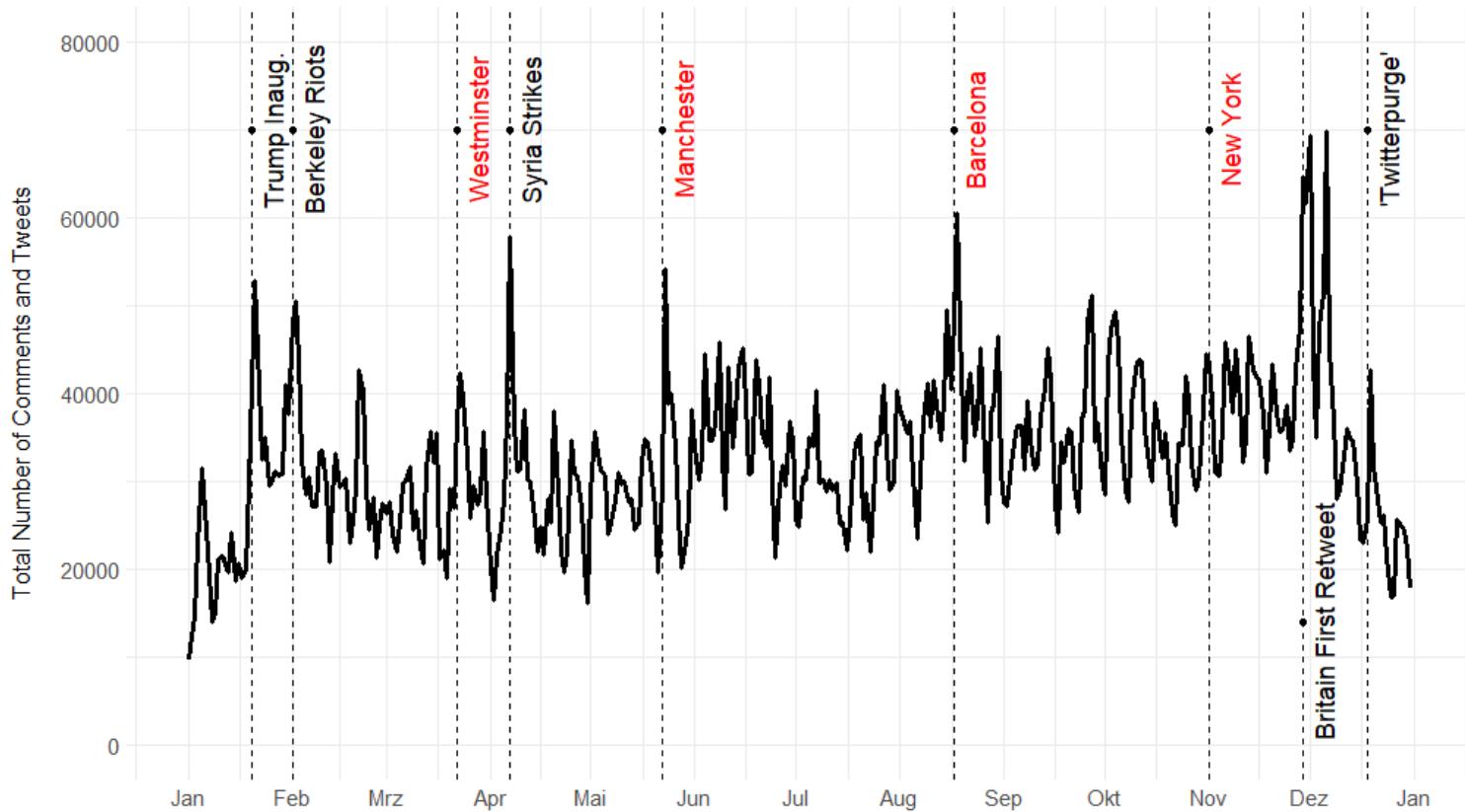
Who did we scrape? Alt-Right



Who did we scrape? - Media



Total Number of Posts and Comments



Language Modeling: Toxicity

Training Data

TODO

- update three slides
- animate toxicity on text

Training Data

TODO

- update three slides
- animate toxicity on text

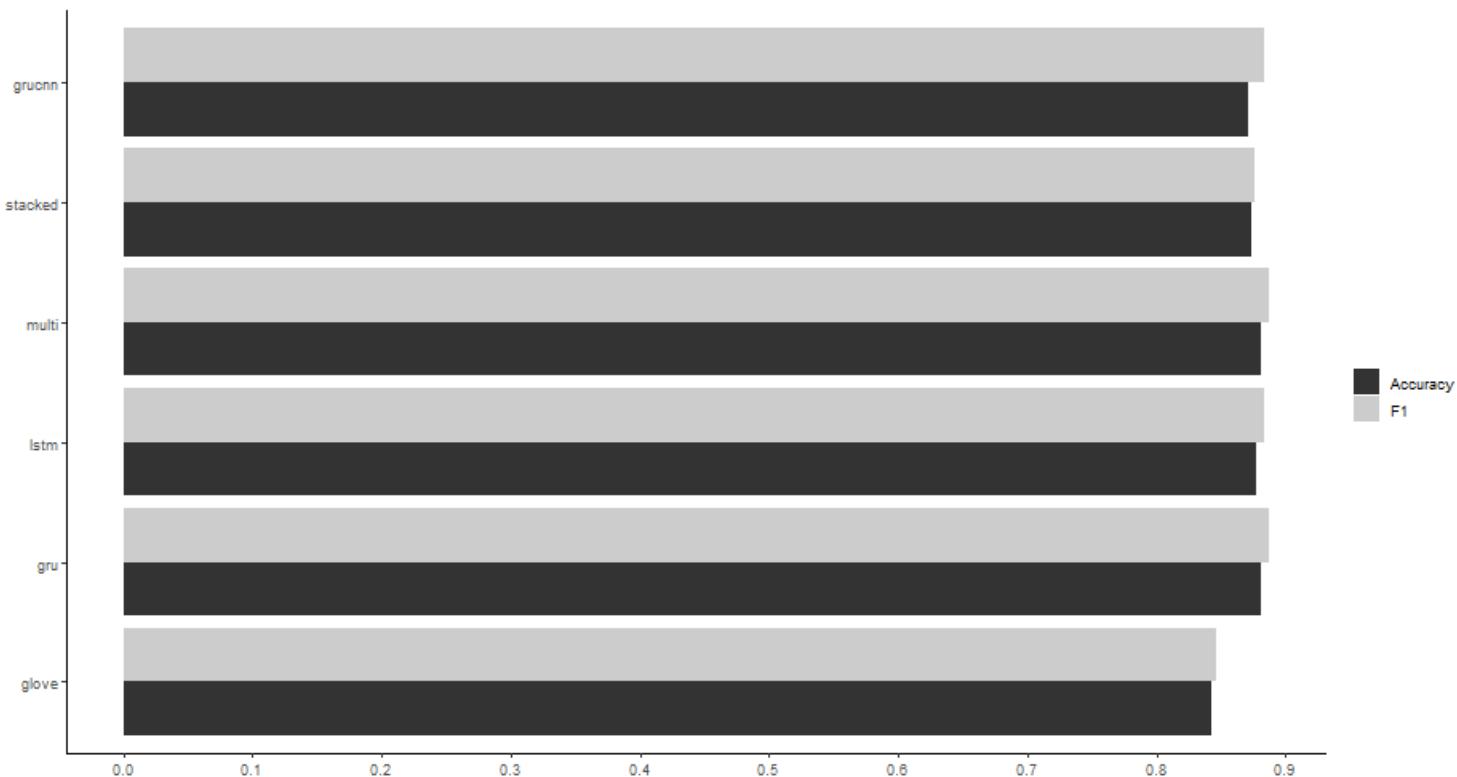
Data Source: Wikipedia comments annotated with **toxicity** scores available in as [Kaggle Competition](#). Five thousand crowd-workers were involved in labeling different toxicity and sub-types like personal attacks for approx. 160k comments. The dataset is documented [here](#).

- balanced data to 30.000
- Text Normalization
- logo toxic/Jigsaw
- ask for other interesting datasets



Predictions

- different keras models
- confusion matrix



Explainable AI

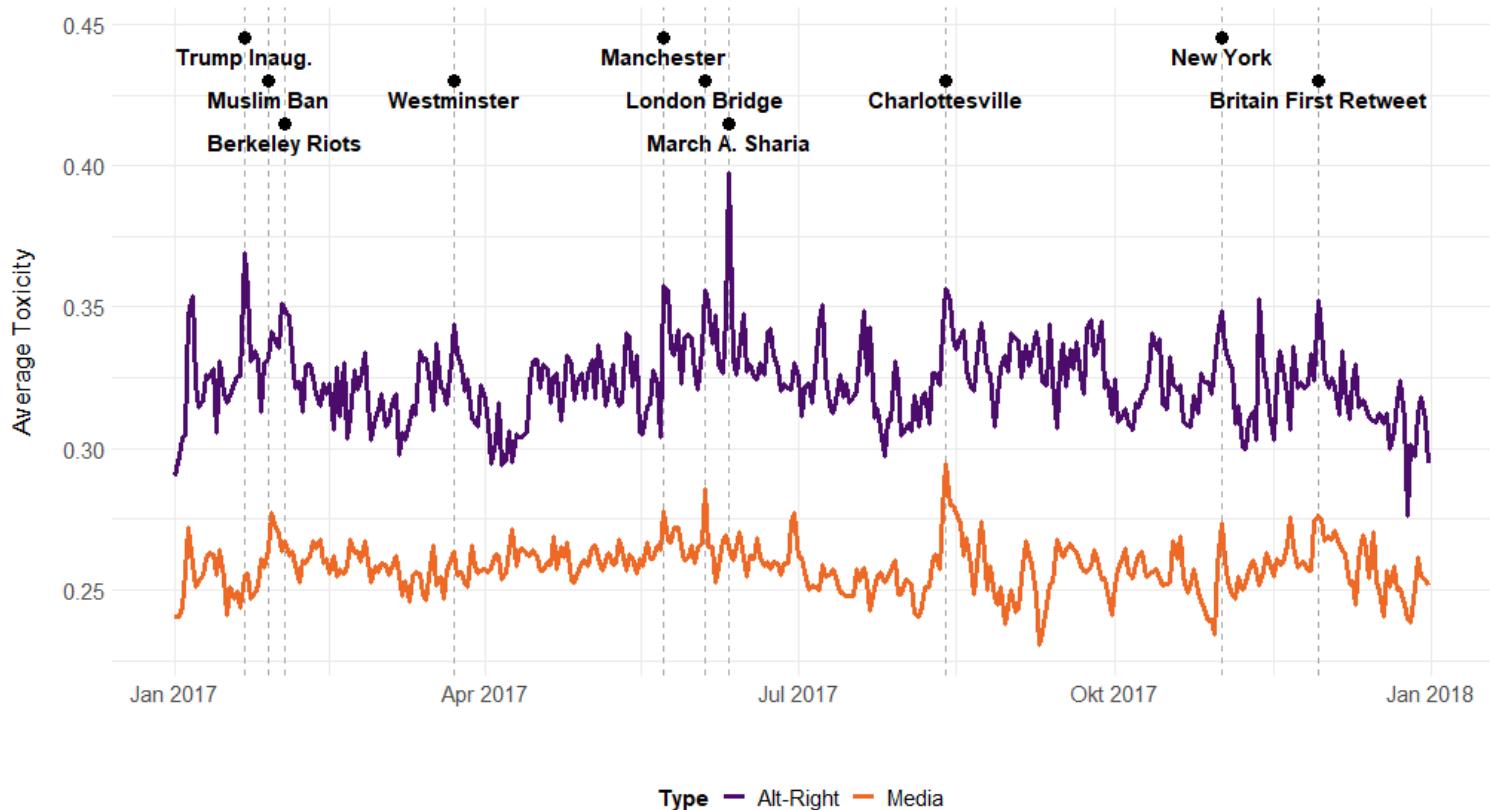


What does a language model learn?

word	glove	lstm	gru	multi	stacked	grucnn	ngram
Fuck	0.753	0.997	0.992	0.996	0.995	0.993	Fuck those cowards
those	0.436	0.687	0.598	0.854	0.922	0.785	those cowards that
cowards	0.458	0.79	0.722	0.71	0.914	0.727	cowards that attacked
that	0.501	0.646	0.352	0.343	0.41	0.443	that attacked you
attacked	0.5	0.563	0.389	0.672	0.414	0.547	attacked you They
you	0.516	0.534	0.414	0.625	0.484	0.432	you They will
They	0.491	0.438	0.392	0.306	0.428	0.27	They will get
will	0.476	0.426	0.249	0.232	0.438	0.191	will get what's
get	0.463	0.242	0.222	0.128	0.341	0.064	get what's coming
what's	0.371	0.135	0.12	0.079	0.295	0.045	what's coming to
coming	0.379	0.184	0.117	0.094	0.325	0.151	coming to them
to	0.386	0.249	0.124	0.096	0.402	0.153	to
them	0.415	0.304	0.143	0.112	0.421	0.233	them

- concept
- our examples
- compare classifier

Average Toxicity by Day



Statistical Modeling

modeling todos

- Multilevel Logit/Regression
 - fit1: shows difference in toxicity by user type (ar, media, mixed)
 - fit2: shows difference in toxicity by user type and media type (media_type): media and alright and mixed toxic comparison
 - fit3: shows keywords and associated toxicity by user type and media type
 - fit4: shows likes, comments and shares toxicity by user type and media type

Maybe this but less important

- Mean toxicity by page
- Keywords count (parallel chart) on media and alright

User Activity

How do Media and Alt-Right interact online?

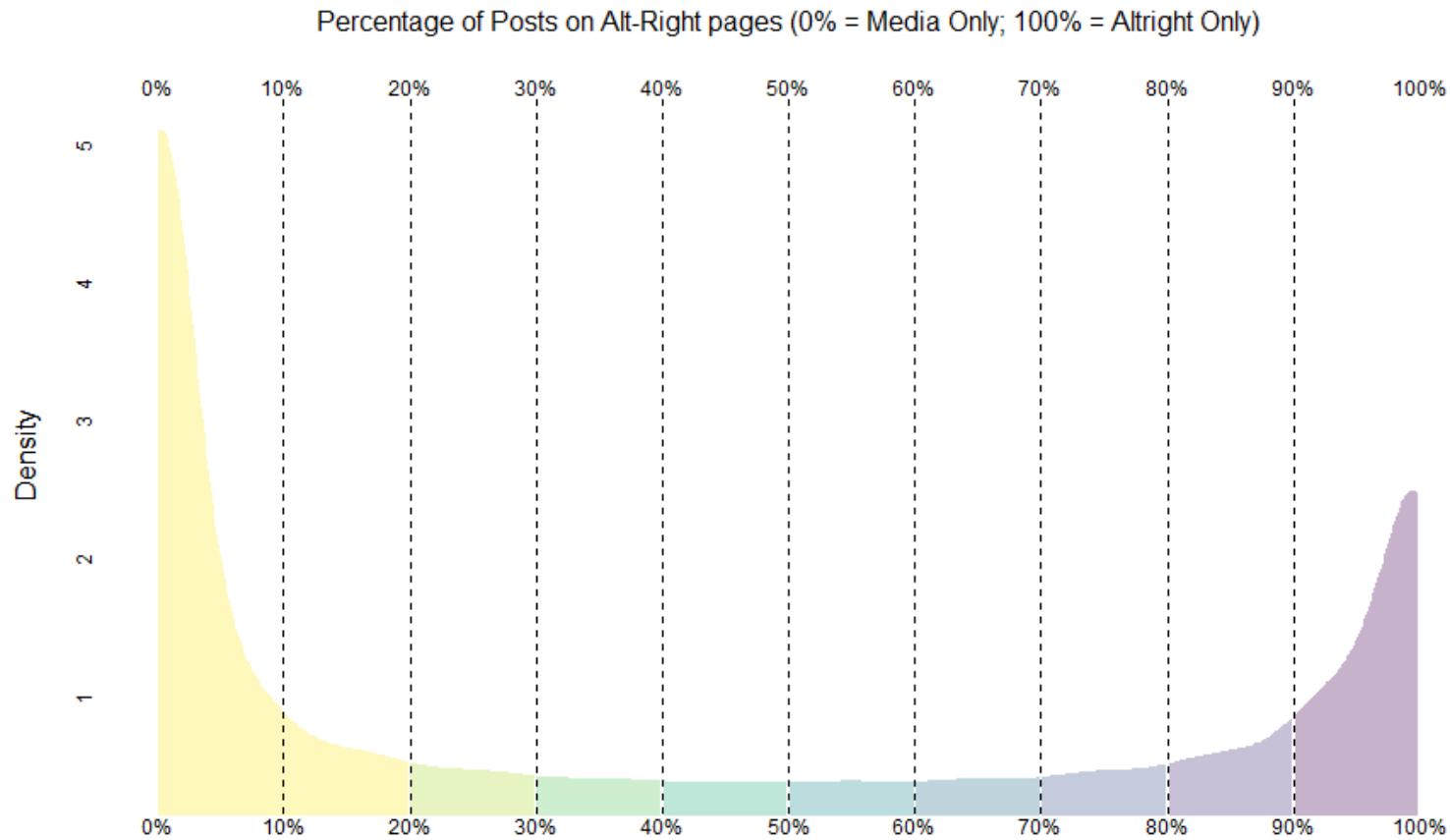
User Activity ToDos

- Flow of Toxicity (event toxicity)
 - implement event titles
 - data in file name
 - get average toxicity over time
 - plot a line chart below

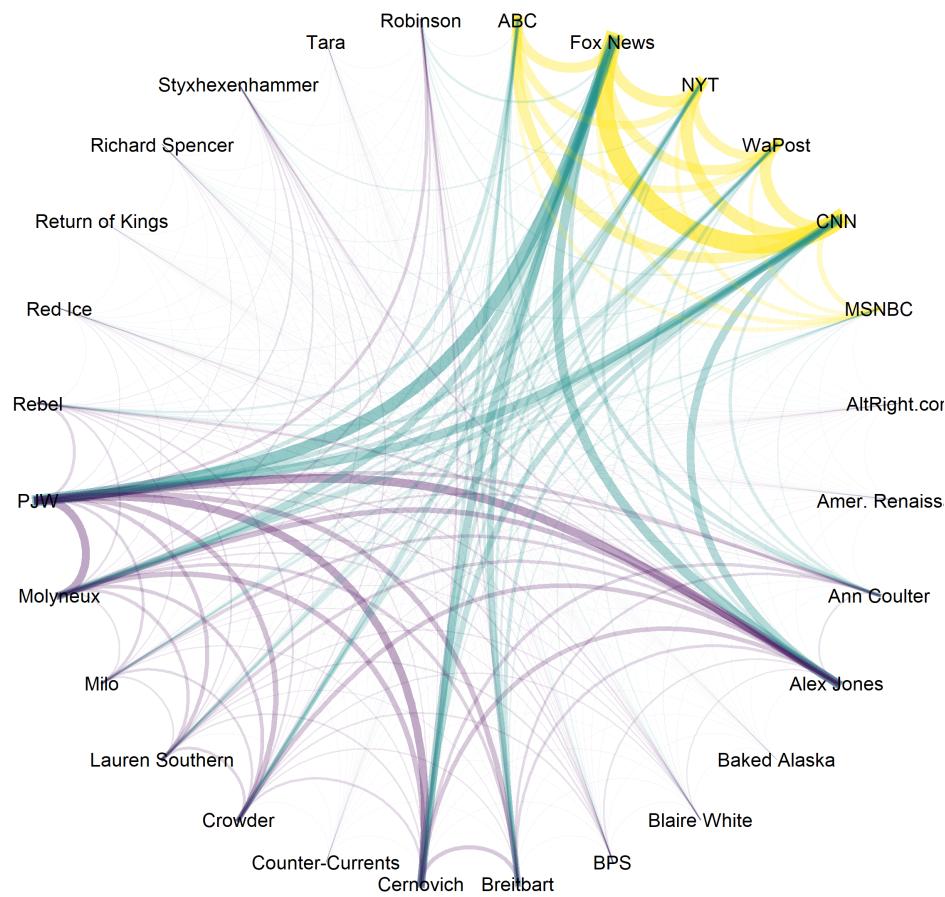
probably not

- How does the Alt-Right intersect with media? Timeline plots
 - What are days where the Alt-Righters came to the media? Why?

Differentiating between User Activity Types



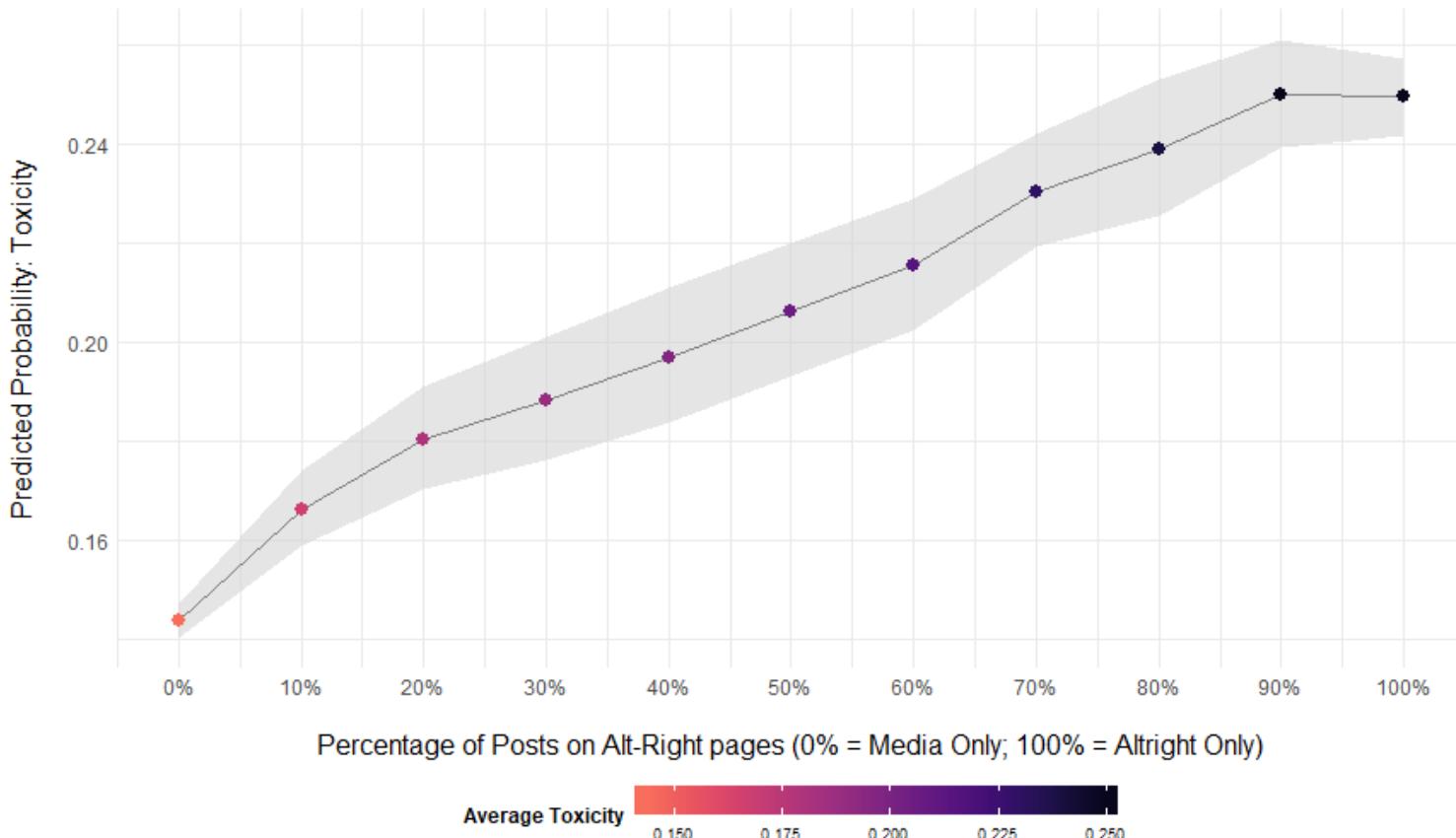
Shared Active Users: Alt-Right Profiles and Media Outlets



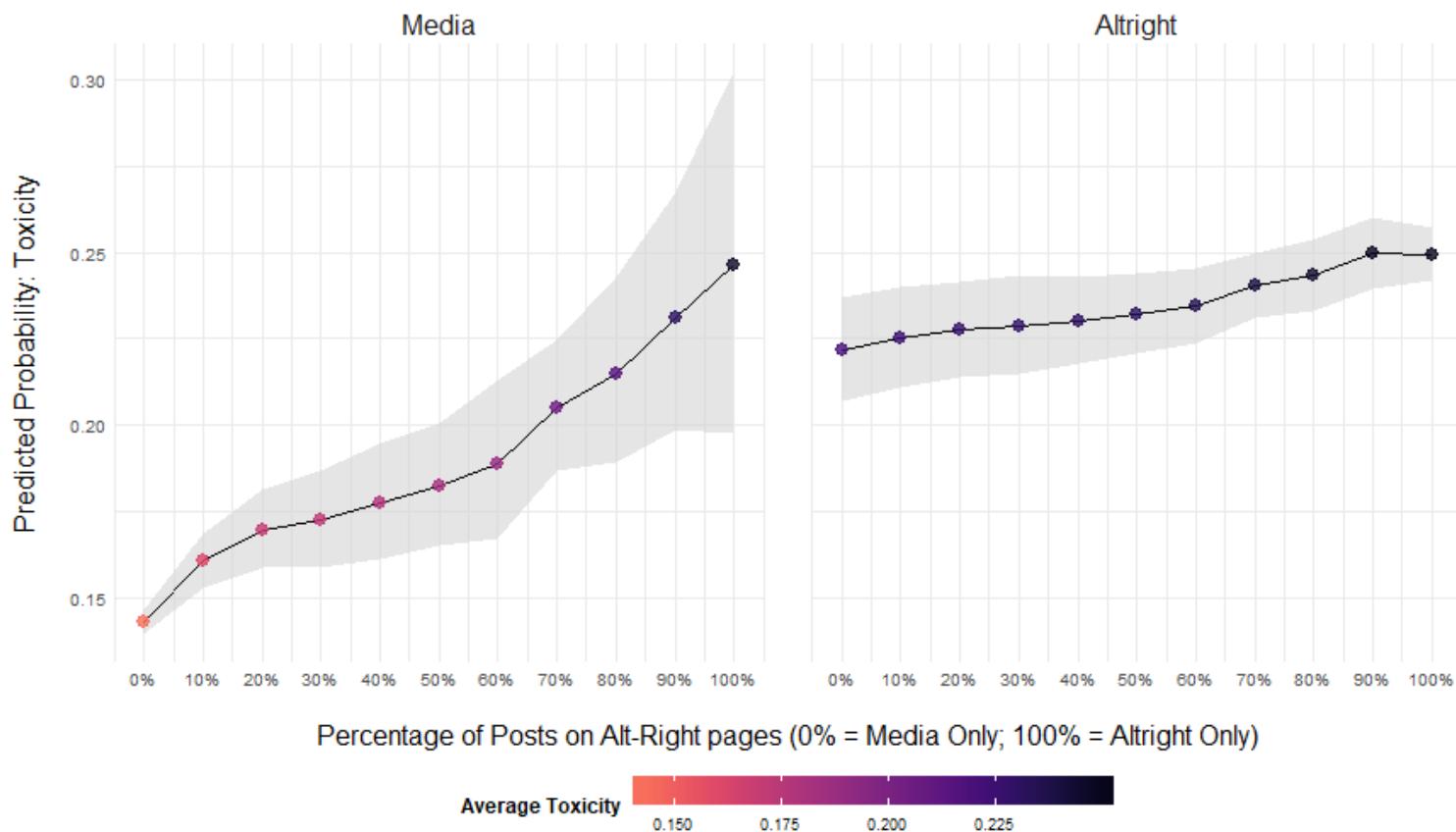
Link Type — Alt-Right — Alt-Right ~ Media — Media

Number of Users — 50000 — 100000 — 150000

Null Model

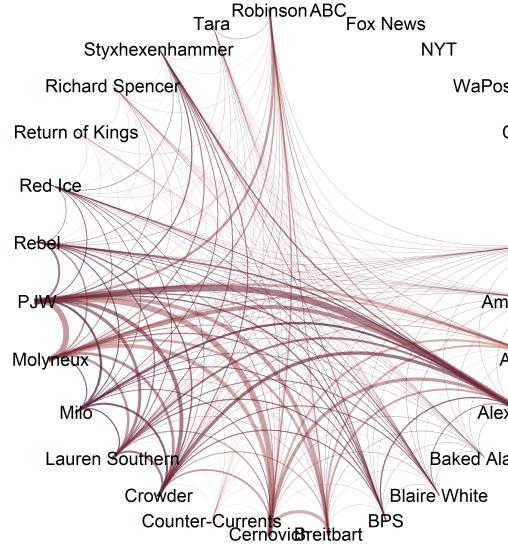


Model 1

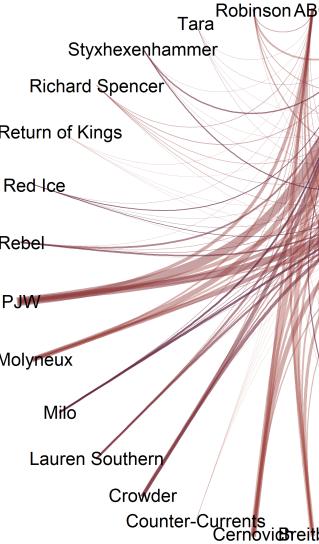


The Flow of Toxicity

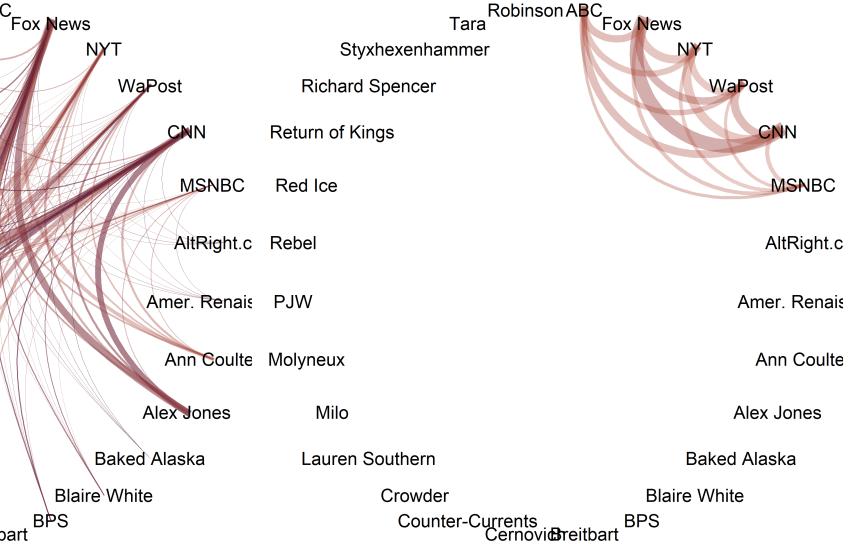
Alt-Right



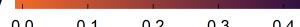
Alt-Right ~ Media



Media



Average Toxicity

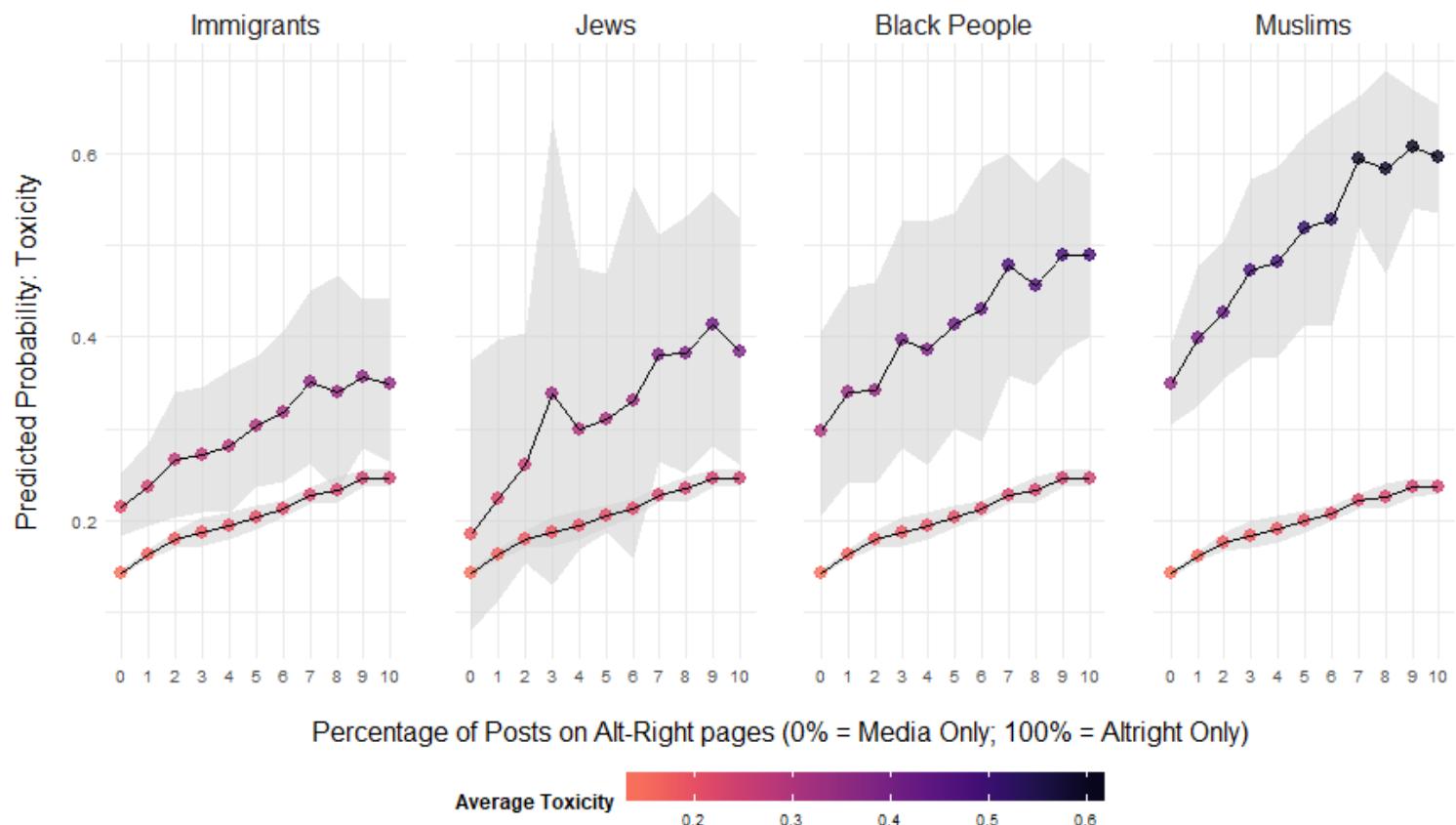


Number of Users

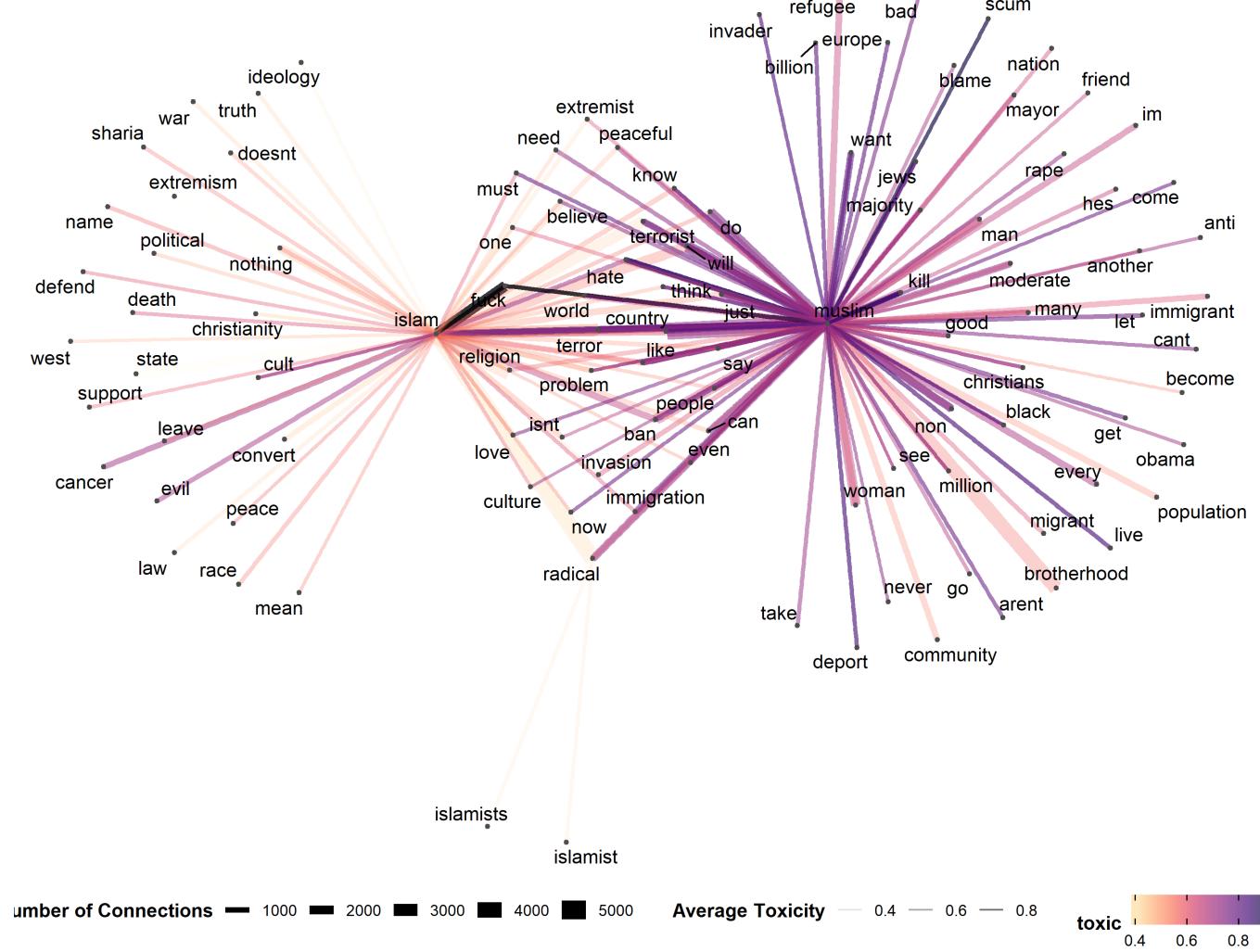


Content Analysis

Model 3



Bigram Network: Muslims



Bigram Network: Jews

Bigram Network: Migrants

Future Research

Possible other applications:

- How does the Alt-Right react to terror attacks
- Tracking stories and misinformation (like the hoax that two Syrian refugees were responsible for the Quebec Mosque Attack)
- Distinguishing between Alt-Light and Alt-Right content

Conclusions

- TODO TODO

Conclusions

- Account for the *multidimensionality* of extremism
 - extremism is not only limited to hate speech, it's also ideological
- Taking complaints about ambiguity of hate speech seriously
- More accountable AI programming:
 - AI deletion: problems with uncertainty and ambiguity of language
 - Need for *human in the loop*
 - Include Feedback: possibility to appeal (restricted times)
 - Algorithms must be responsive and transparent
 - Highlight content in question and reasons
 - Cross-checking algorithms to reduce error rate

Thanks for Listening!

<https://decoding-the-altright.netlify.com/>

GitHub

References

Hawley, G. (2017). Making Sense of the Alt-right. Columbia University Press.

Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online. New York: Data & Society Research Institute.

Appendix

Hyperparameter Tuning

