



Targeting African American Voters: 2020 Ongoing Influence Operations Deciding US Senate Runoff Election in Georgia

Cody Webb, Shireen Mitchell, and Jessalyn Hernandez



Table of Contents

3	Introduction
5	Our Findings
6	Authentic/Inauthentic Accounts
7	Tweet Network Analysis
14	UMAP Analysis of Hashtags
17	Toxicity Analysis
19	Top Spammers — UMAP Account Analysis
28	Top Accounts By Volume, With Tweets Labeled by Toxicity
33	Top Trolls by Tweet Toxicity
36	Wrapping up
38	Conclusion
39	References
40	Appendix



Introduction

The state of Georgia is currently undergoing a run-off election for two hotly contested Senate seats that could shift the power of the Senate to the Democratic party, tremendously altering the current political landscape going into 2021. African Americans make up the second largest voter group in Georgia, and are now under an influence operation geared toward flipping their votes, and even stopping them from voting at all.

With 4 in 5 U.S. adults are on social media and using it as part of their information diet, there is not one online community that isn't being held captive by accounts marionetting personas, spammers, and other threat actors looking to harass and push a narrative of beliefs onto other users. This has resulted in public attention and debate among social platform hosts about what is too much moderation versus what is too little. Many mainstream social media platforms have opted for little to no regulation around the type of content shared, using free speech as their defense. This has left communities to fend for themselves, giving a rise to citizen data scientists committed to using their expertise to map and combat the spread of disinformation.

Project Domino is continuously monitoring the narratives and communities surrounding the Georgia runoff elections. Over the past 30 days we have collected more than 150,000 tweets around the election along with specific astroturfing campaigns. "Overall, Project Domino saw several major campaigns, 3 of 4 of which seem manipulated, and enough so to dominate 60.5% of the overall discourse."



The specific campaigns include:

ADOS (#ADOS) - American Descendants of Slaves. ADOS is a campaign focused on reparations for African Americans but only for individuals who are descendants of slaves on American soil. The campaign makes demands of Democratic candidates to develop a Black agenda.

Tactics employed: influencer-driven Astroturfing and harassment, narrative control, disinformation and conspiracy theories, spamming.

Blexit (#BLEXIT) - Black exit from the Democratic party. BLEXIT is a campaign that claims freedom from key cities in states that hurt black voters due to democratic policies (including Chicago, Philadelphia and Detroit).

Tactics employed: Influence- driven astroturfing and harassment, narrative control, disinformation and conspiracy theories.

Black Voices for Trump - Black Voices for Trump is a Trump Campaign-driven tactic parading as a legitimate movement. Part of the launch falsely highlighted a key initiative that would redirect dollars for climate change to the Black community..

Tactics employed: Spamming, astroturfing, and harassment.

Walk Away (#WalkAway) - Launched around 2018, WalkAway is specifically designed to get black voters to leave the democratic party as an empowerment strategy. Current celebrity spokespersons supporting this campaign include Isiah Washington¹ and Brandon Straka.

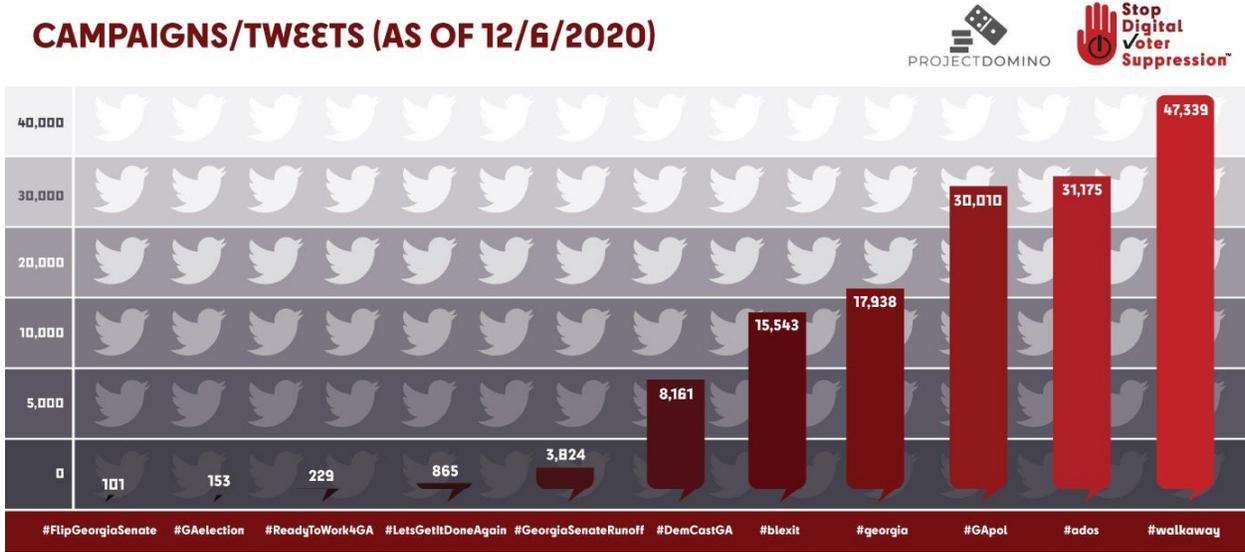
Tactics employed: Influencer-driven astroturfing and harassment, narrative control, disinformation and conspiracy theories, spamming.

These narratives target Black Americans with the intent to sow distrust of the Democratic party and process of the United States. The data collected have been run through multiple machine learning driven analyses to **identify key threat actors, narratives, topics, and active information operations**. These analyses allow us to gather insights into the community at risk for digital manipulation. The information in this report has been garnered from Twitter. However, to remove outside noise of other ongoing election- related topics, we have narrowed our search queries to terms surrounding the Georgia Election.



Our Findings

We have currently identified active influence operations taking place to disenfranchise and dissuade citizens from taking place in the electoral process, with different methods ranging from disinformation and digital voter suppression, to harassment and threats of violence.



(Above: Tweets that were analyzed for this report)



Authentic/Inauthentic Accounts

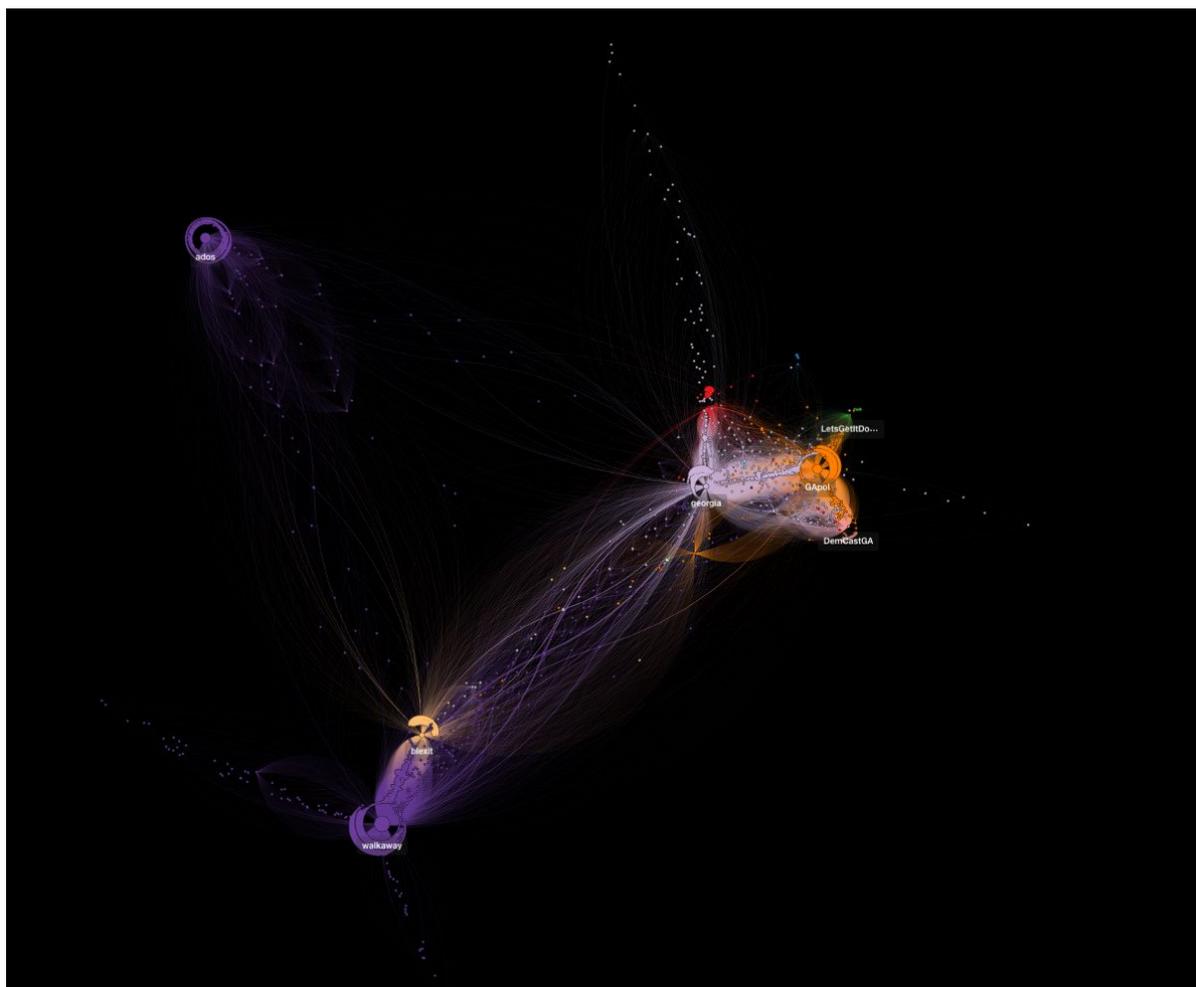


To categorically label accounts as authentic/inauthentic with a high degree of confidence would take access to data, with which we as citizen data scientists do not have access to. So instead, we have shaped our analyses around confidently identifying the key parts of an influence operation:

1. **Influencers**- *Give Legitimacy to the Narrative*
2. **Spammers**- *Help the Narrative reach a wider audience*
3. **Trolls**- *Reinforces the Narrative through harassment and interaction*



Tweet Network Analysis

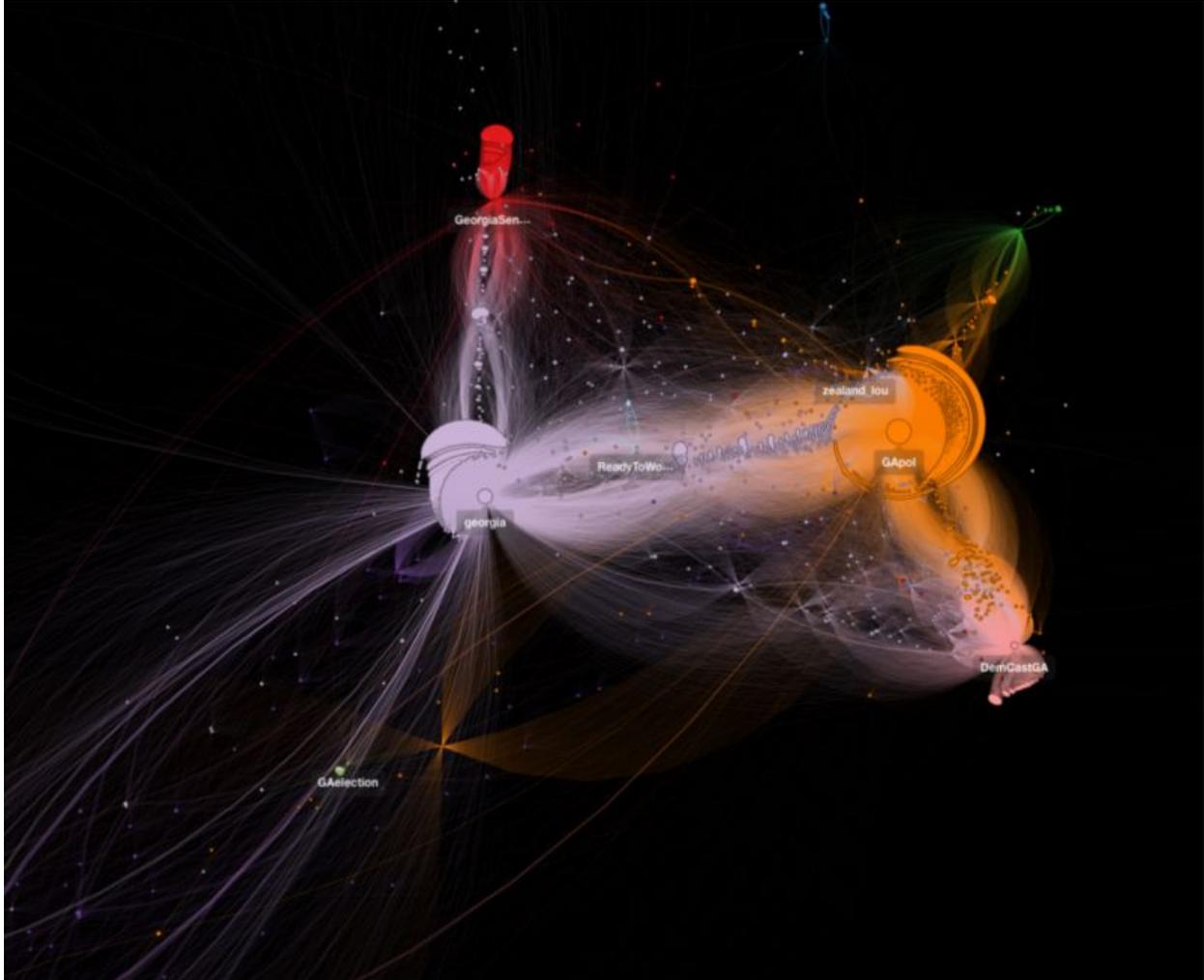


(Above: a network of 150k tweets surrounding the Georgia Senate election and ongoing influence operations.)

The above Network visualization (based on the entire set of tweets), provides a bird's eye view of all of the tweets and actors in the information space. Each node is a hashtag or a username, and each edge is a tweet.

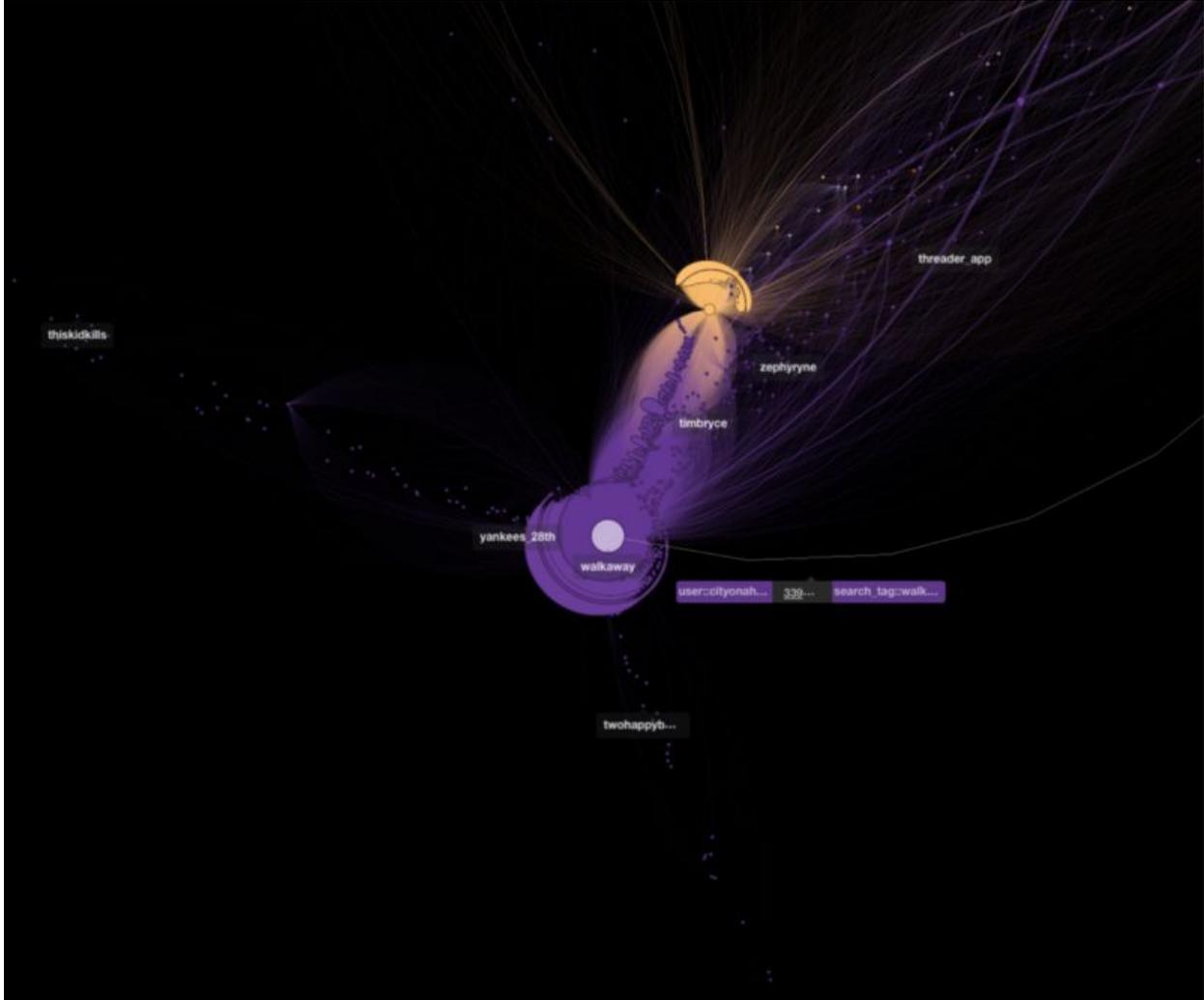
We have identified three distinct sections of our network, two of which are **active influence operations targeting African American Voters** (ADOS & Walkaway);the third is the Georgia runoff election topics. The Georgia election topics portion of the network is closer to and shares more connectedness with the #walkaway portion of the network compared to the other active information operation. We will focus our initial analysis here as it is a clear indication of which information operation is actively influencing the conversation around the Georgia Runoff Election.





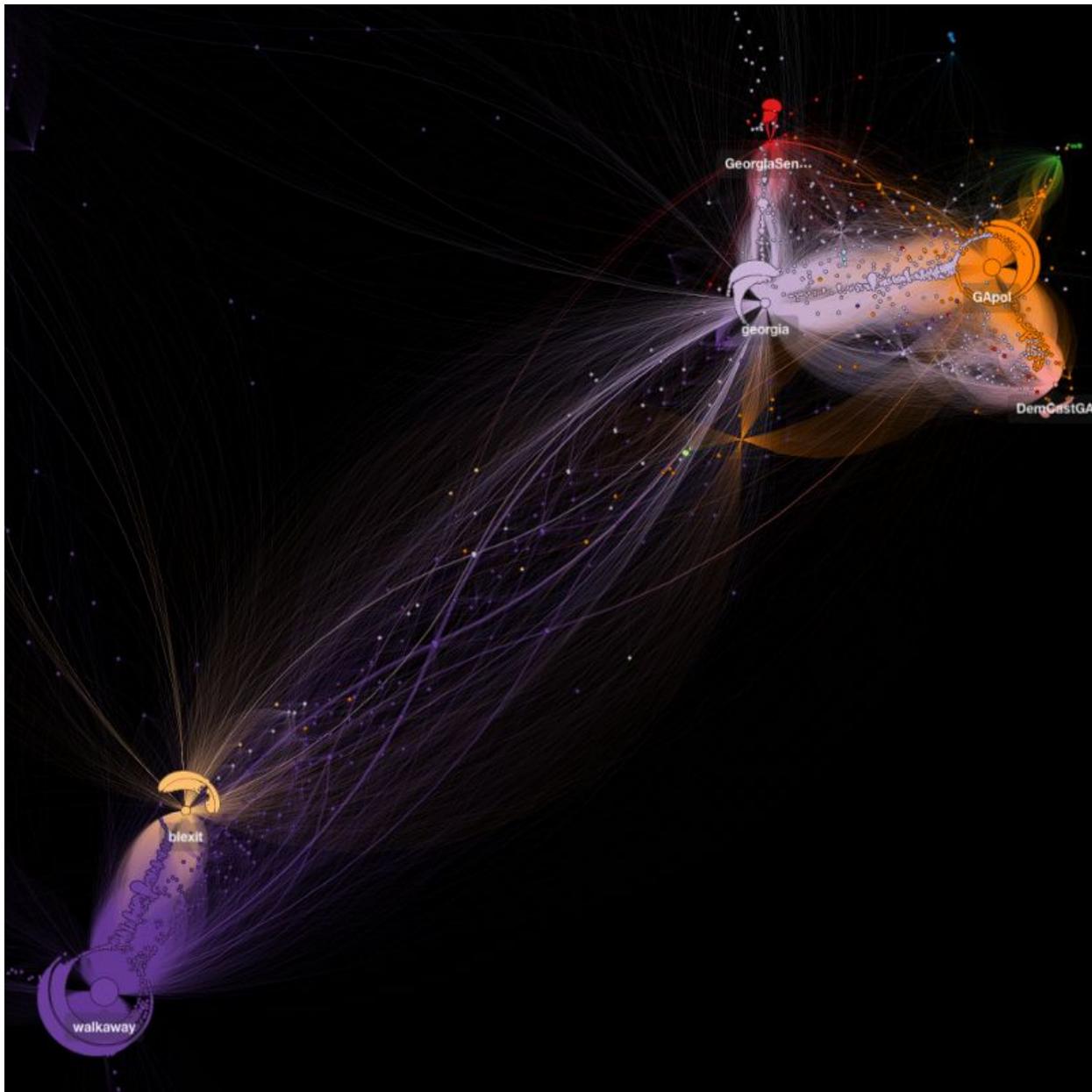
(Above: #Georgia, #GApol section of the network analysis)



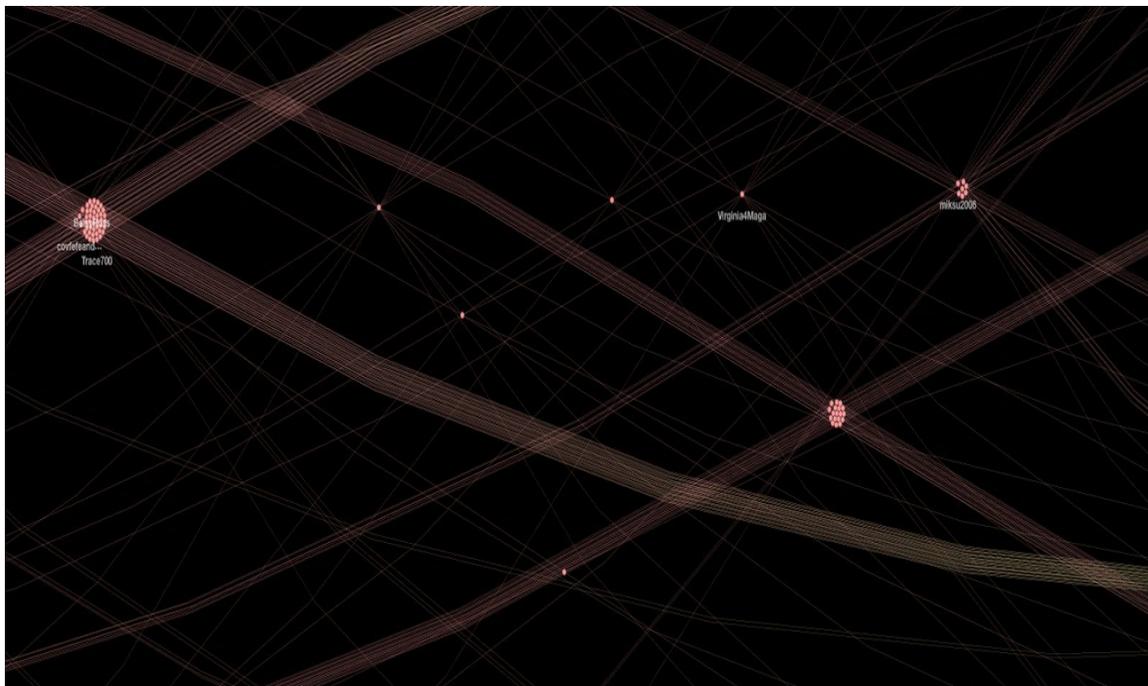


(Above: #walkaway section of the network analysis)



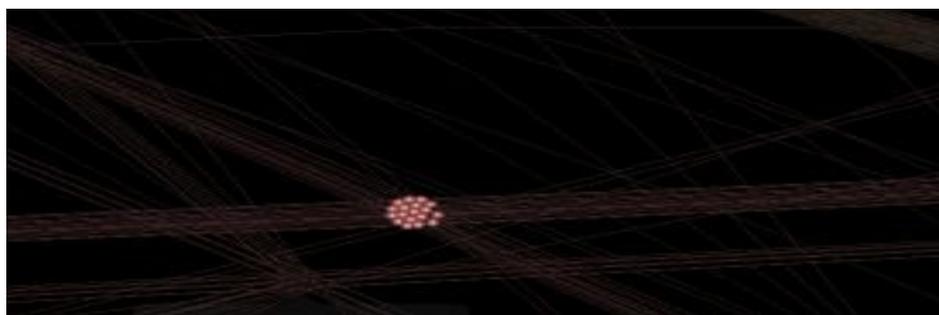


(Above: #Georgia, #Gapol section of the network analysis and the #walkaway section of the network analysis and the shared connections between each)



(Above: groupings of Accounts posting about #walkaway and #georgia or #gapol)

Let's take a look at one of the "x" patterns that we see in the network. The distinct "x" pattern occurs when a single point provides a complete path to and from two or more separate points in a network. In this case, these are groups of accounts cross-pollinating topics in between an active Information operation ("#walkaway) and the target community (#georgia, #GaPol).



(Above: Distinct "x" pattern seen in between an active information operation (#walkaway) and a community (#GAPol, #georgia).)



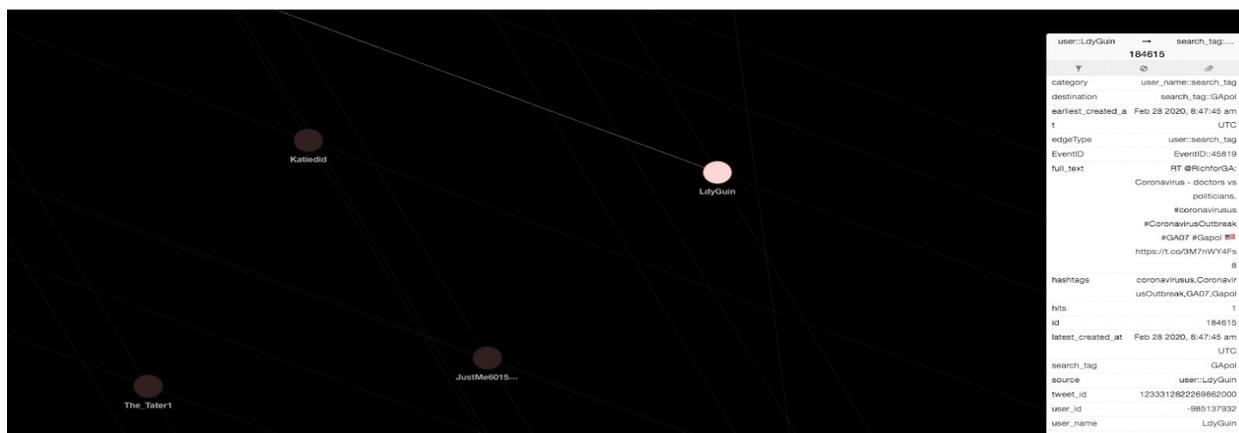
(Above: Grouping of accounts involved in providing messaging to a community from an active information operation.)

The connectedness of the groupings of points in between two or more different networks is proportional to the distance between the grouping and the networks it is connected to; this causes the grouping to position itself to the middle between the networks. Notably, that brings this particular grouping of accounts so close is that they all were **retweeting the same exact post and providing a similar pro #walkaway, anti-democratic narrative in subsequent tweet.**

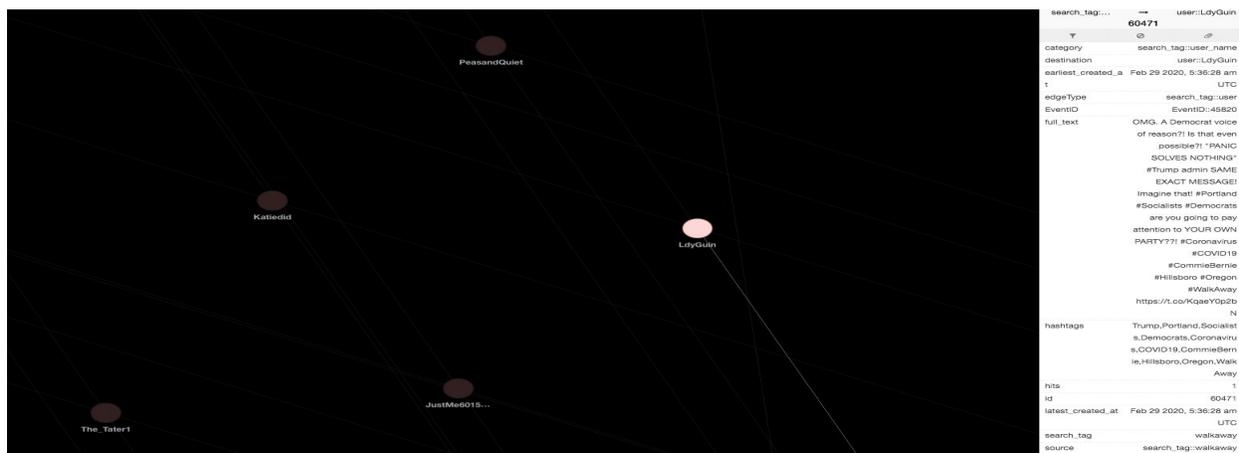
In the example below, we see how one specific user collected in our set is participating in two different narratives. LdyGuin is witnessed here amplifying a #GApol tweet while concurrently tweeting a #WalkAway message which seems to be directed at those living in Portland, Oregon. The cross-over analysis allows us to see how users fall into specific communities. **The specific nodes that appear as an 'X' given their relationships across main topics provide the analysts with an effective way to spot users that are spreading similar narratives and doing so at the same relative frequencies.**

```

full_text          RT @RichforGA:
                  Coronavirus - doctors vs
                  politicians.
                  #coronavirus
                  #CoronavirusOutbreak
                  #GA07 #Gapol 🇺🇸
                  https://t.co/3M7nWY4Fs
    
```



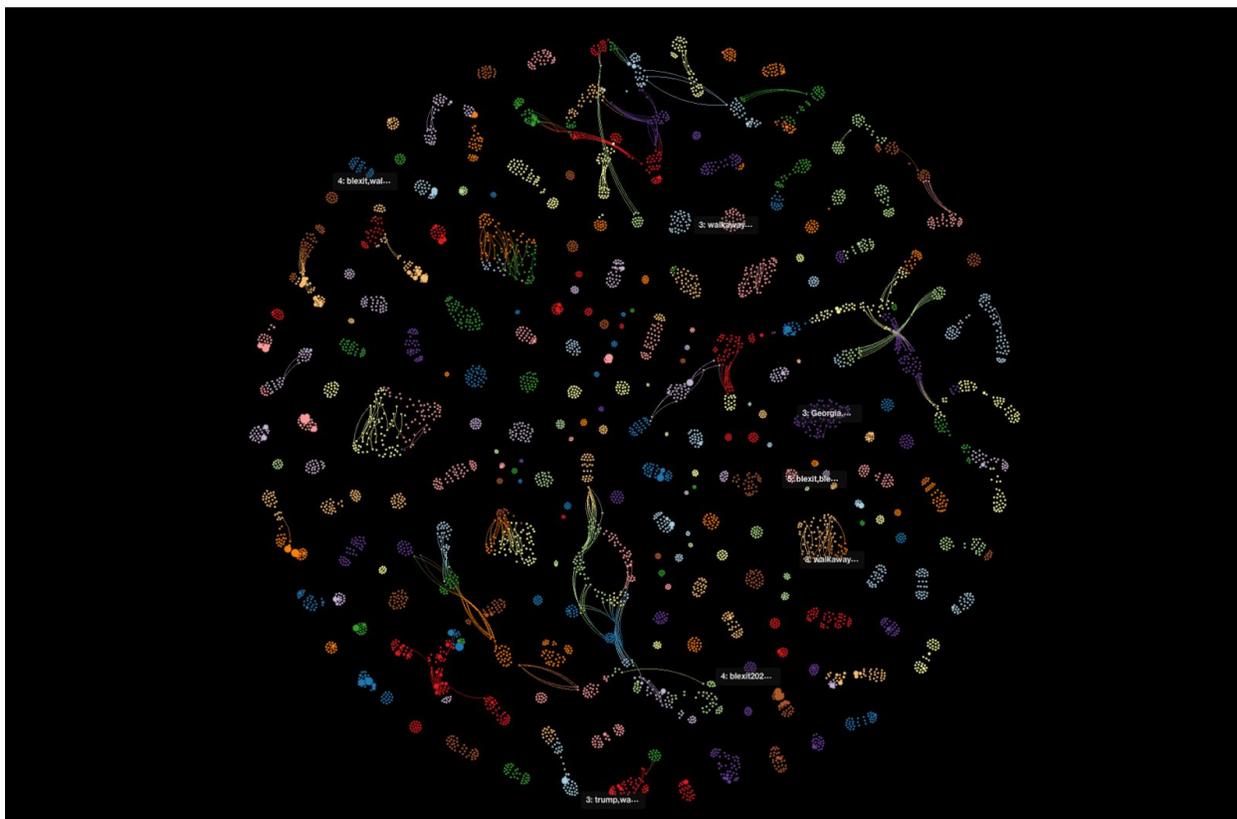
(Above: an example of contrast messaging provided by an account in the grouping between #walkaway and #georgia, #gapol)



(Above: Another example of contrast messaging from the same suspect account in the grouping between #walkaway and #georgia, #gapol)

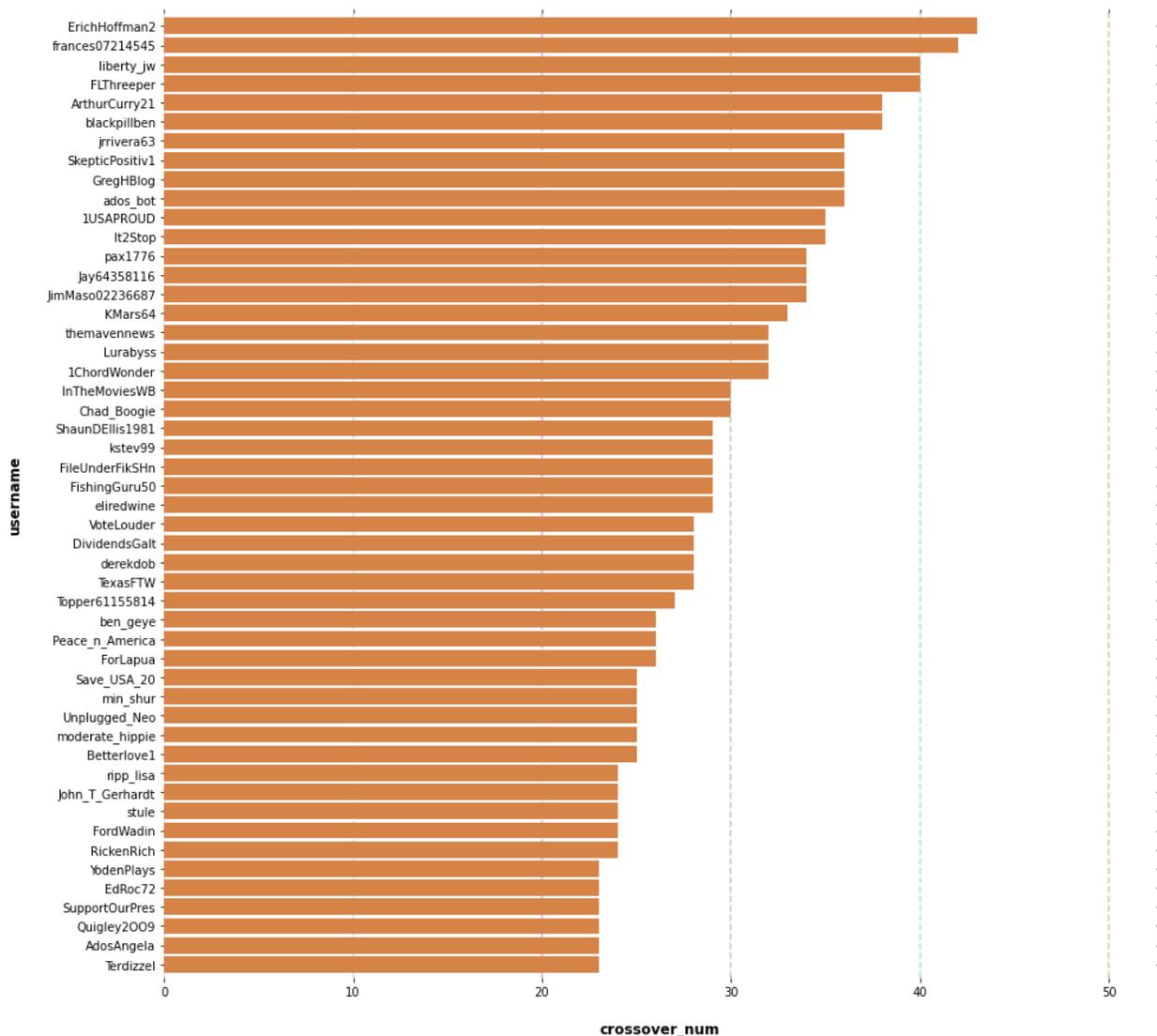


UMAP Analysis of Hashtags



In the case of this collection, only Tweets containing keywords like WalkAway, ADOS, GAPOL, and Georgia were collected. Therefore we witness an overwhelming amount of these specific hashtags present in the dataset. Yet with UMAP^[A1] analysis we are able to see the specific hashtags each user is tweeting, and their overlapping audience with shared characteristics. In the end, we are given a visual representation of coordinated accounts pushing specific narratives. A closer look at the largest accounts in this set gives us the most influential and active spammers in the set.



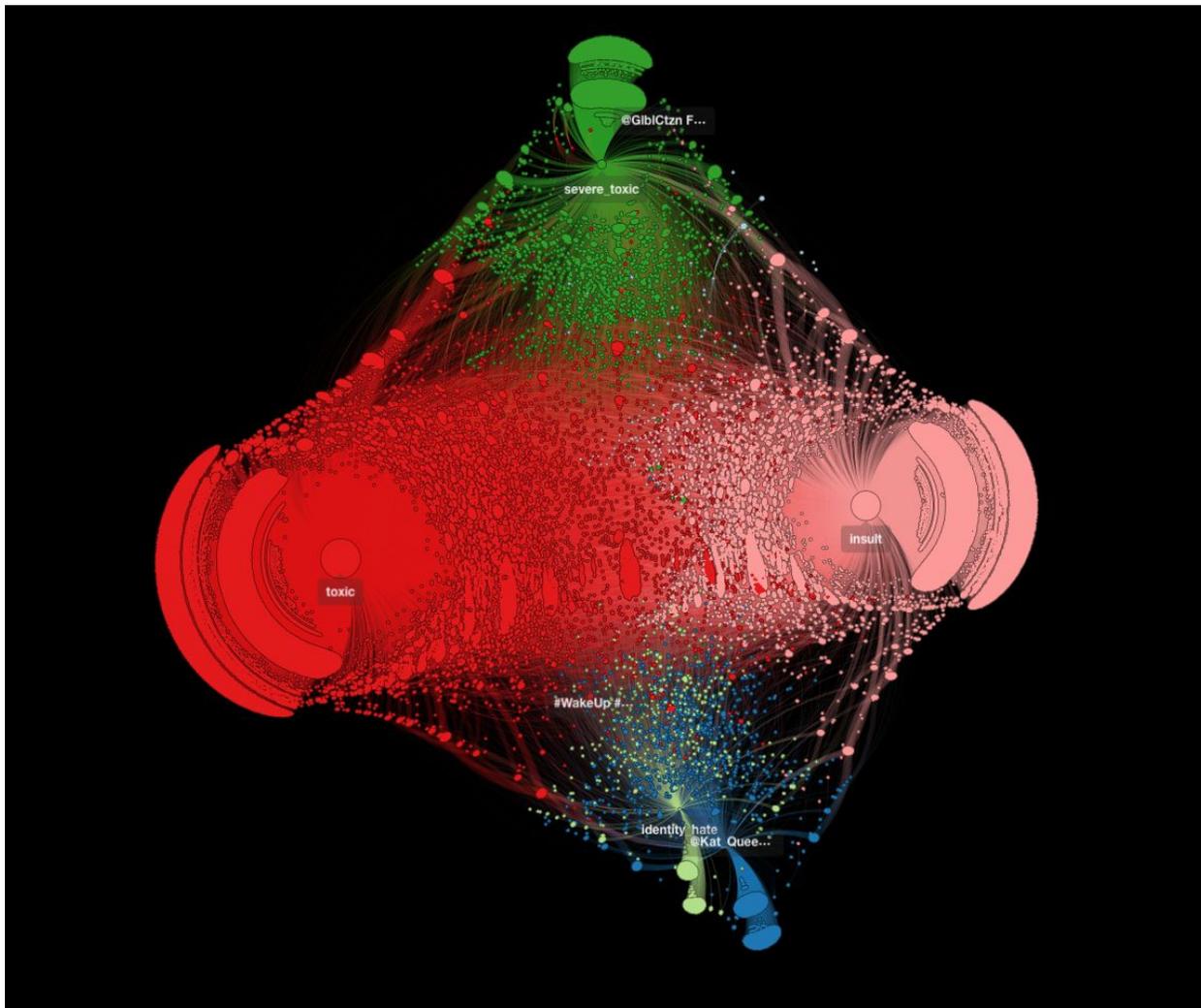


(Above: Top 50 accounts by crossover number^[A].)

The crossover number is how many different hashtags were accompanying the user within the dataset. We can use this as a heuristic to determine the likelihood that an account is spamming hashtags or topics.

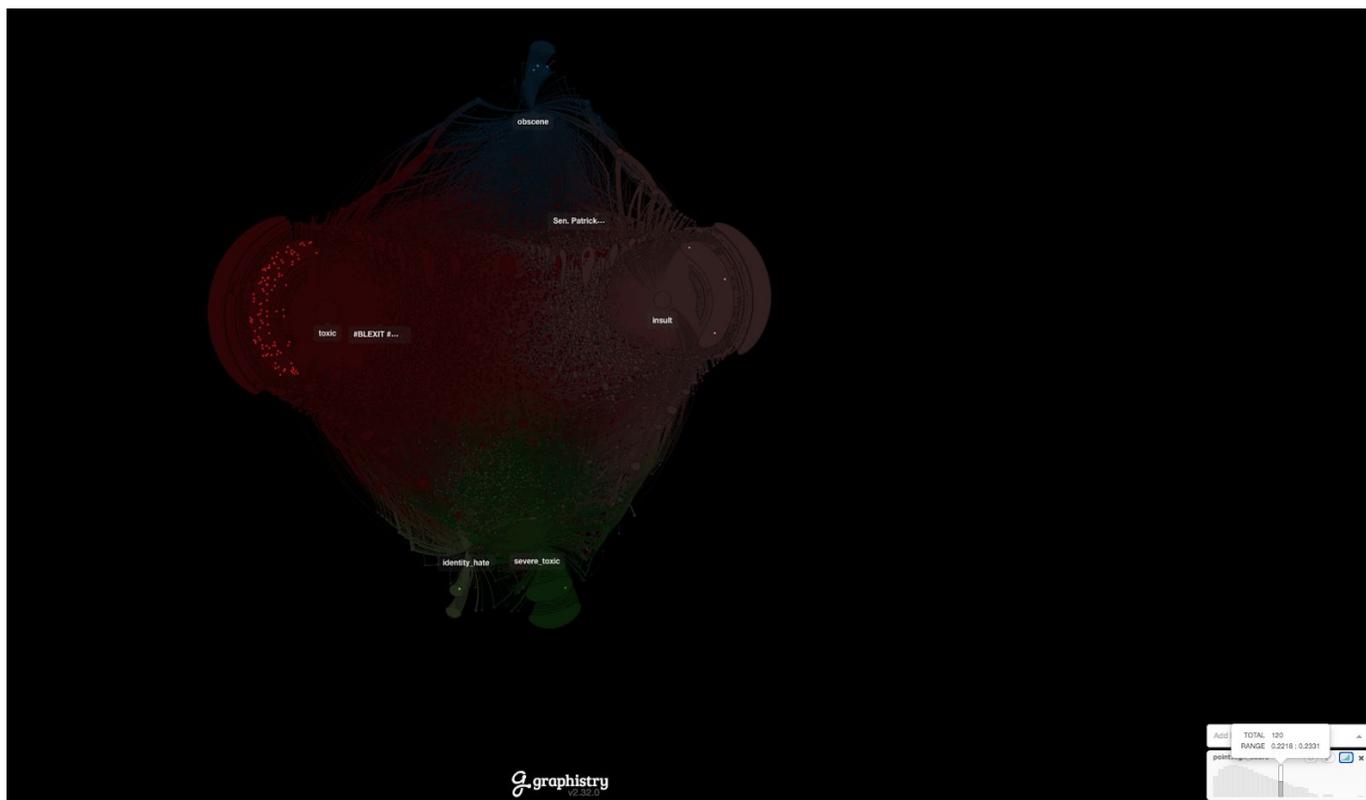


Toxicity Analysis



Toxicity analysis^B is useful when analyzing a large textual dataset as it allows for an easy heuristic for finding content that proved to be in violation of Twitter's term of services. Accounts ranking high in these categories are often 'Trolls,' i.e. people dedicated to instigating emotions surrounding specific topics. In terms of the Georgia general election, we have seen the Trolls captured in these analyses pushing voter fraud narratives and harassing other users in the process.



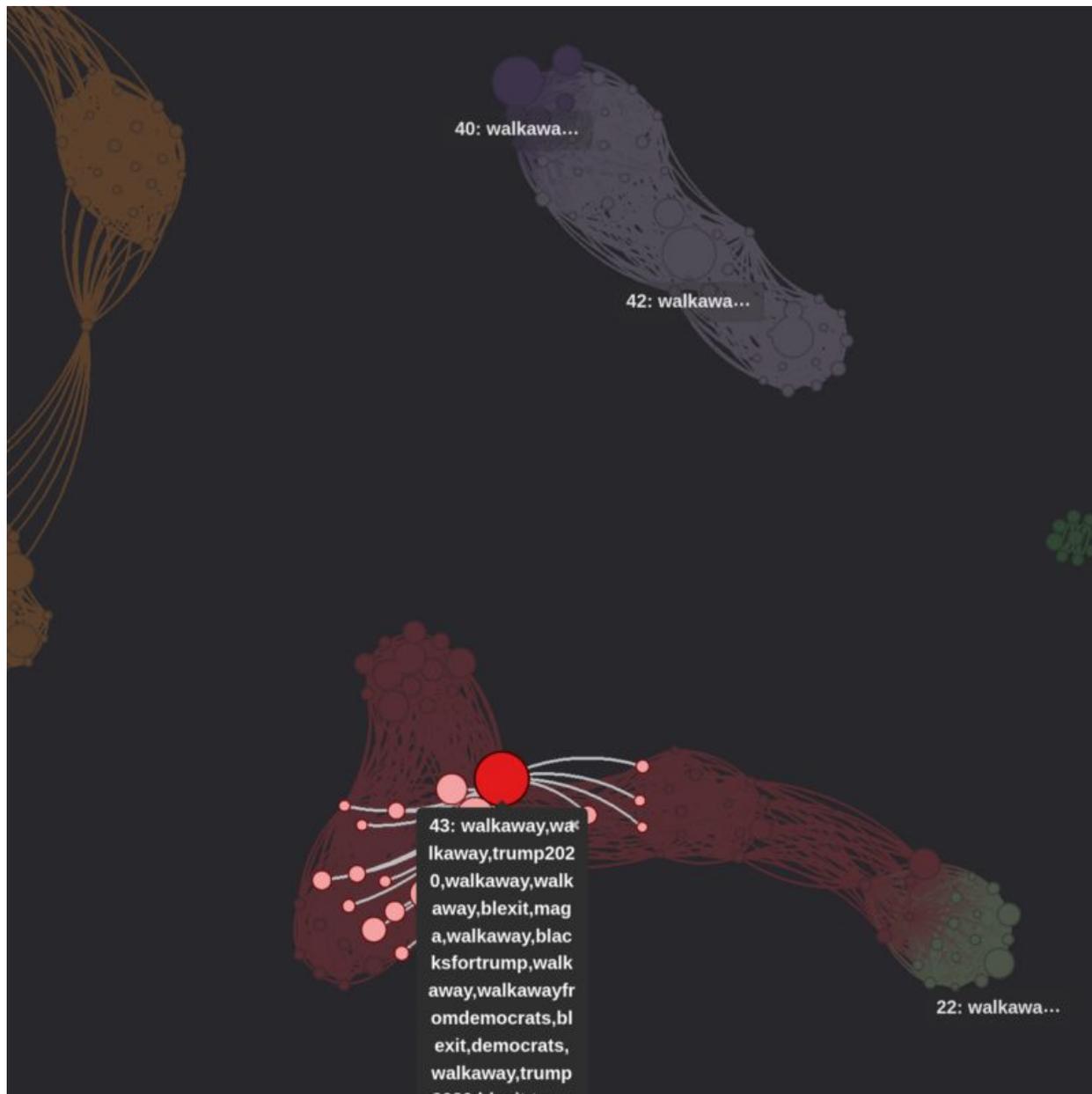


(Above: Isolating tweets and hashtags by range of Toxicity Score^[B].)



Top Spammers — UMAP^A Account Analysis

#1: ErichHoffman2



(Above: Top spam account by crossover number easily represented in our visualization.)

The top spam account measured using the UMAP^B model was @ErichHoffman2, ranking at a rather large rate of 42. We are able to interpret this ranking and placement in the network chart to see what hashtags this specific account is pushing. Studying these hashtags allows for us to understand the most common hashtags in the whole set, as the applied crossover analysis demonstrates how these hashtags have the highest rate of overlap with other accounts' content.

What is most useful in UMAP is that it is very effective at identifying top spammers, which in turn, enables us to discover the exact methods in which these narratives are being applied.

Upon further investigation we note @ErichHoffman2 has a small following but is very active in spreading potent disinformation (e.g.comparing the Democratic Party to rats and constantly pushing a fraudulent election theory surrounding Joe Biden's landslide win). Despite being a small account, it is clear how dangerous the narratives being shared are.





Erich Hoffman @ErichHoffman2 · Nov 2

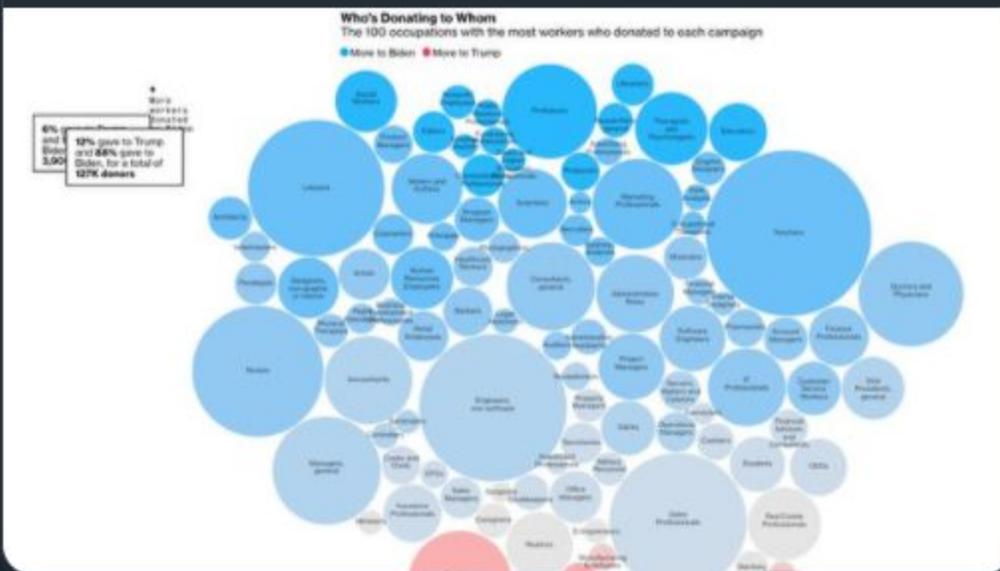
Facts matter. American Patriots vs. Big TECH & corrupt Demrats!

- #MAGA2020LandslideVictory
- #WalkAwayFromDemocrats
- #TRUMP2020Landslide
- #RedWave2020
- #WomenForTrump
- #BlacksForTrump
- #LatinosForTrump
- #WalkAway



Robert Barnes @Barnes_Law · Nov 2

When people ask why I back @realDonaldTrump, this chart sums it up well. For the same reason construction workers, custodians, factor workers, truckers, drivers, and laborers do: he kept his promises to working class America.





(Above: Examples of tweets from our Top account.)





Erich Hoffman @ErichHoffman2 · Nov 21

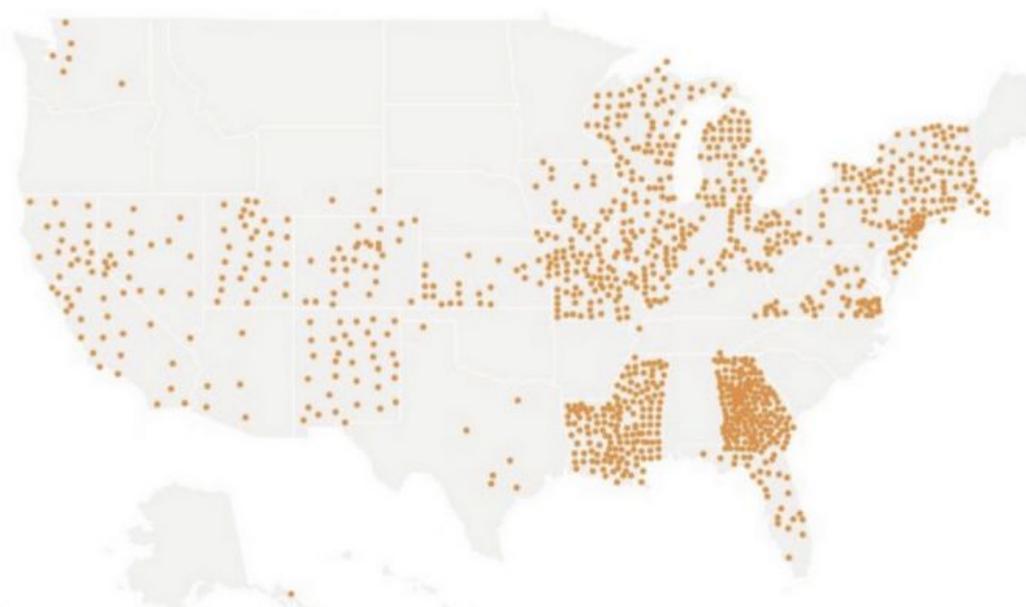


End the madness remove all controversial Dominion software from US voting machines! This is a 1st step to bring back fair & free elections. Next step, NO MORE mail-in ballots. Show up and vote!! Only military or overseas Americans can use absentee ballots with proper voter ID.



Melissa Tate #StopTheSteal @TheRightMelissa · Nov 20

National Distribution of dominion software .. look at **Georgia!** Governor Kemp might very well be one of the politicians that have been uncovered to be getting millions\$ in kickbacks for setting up Dominion in their State. That would make sense why he stabbed Georgians in the back



Erich Hoffman @ErichHoffman2 · Nov 21



Truth to Power!



Donald J. Trump @realDonaldTrump · Nov 21

Big voter fraud information coming out concerning **Georgia**. Stay tuned!





Erich Hoffman @ErichHoffman2 · Nov 17



Truth to Power!! The voting irregularities are ALWAYS in favor of Demrats. Wonder why? Look yourself in the mirror and ask do I want to be with a corrupt political party?



Brandon Straka  @BrandonStraka · Nov 16

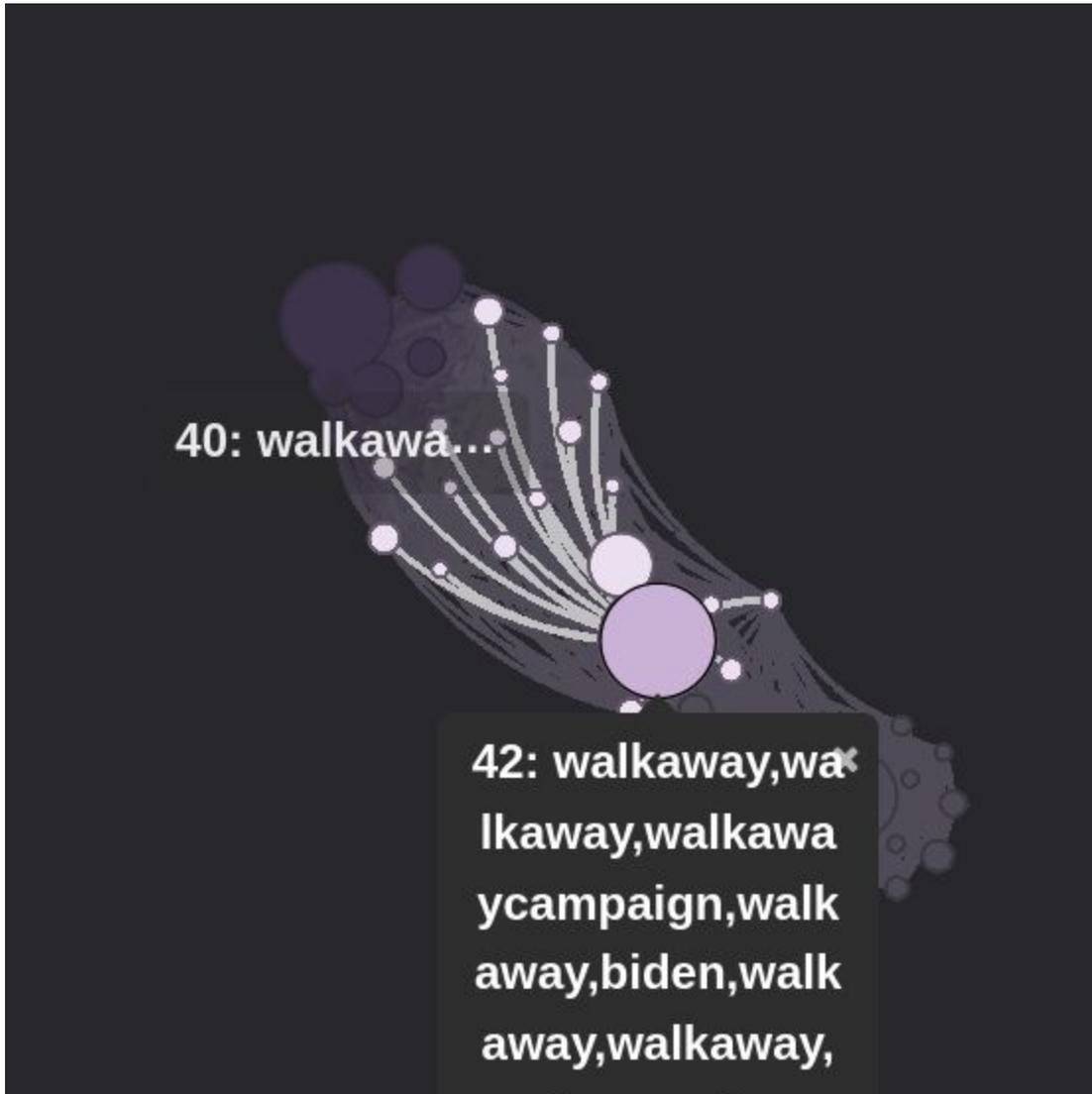
2,600 ballots found in **Georgia** which were uncounted. 2/3 were Trump votes. Again and again- why are the mistakes always *coincidentally* at the cost of Republicans and never the other way around? 🤔🤔



(Above: Erich spamming topics related to Georgia)



Account #2: Frances07214545



(Above: #2 spam account by crossover number easily represented in our visualization.)





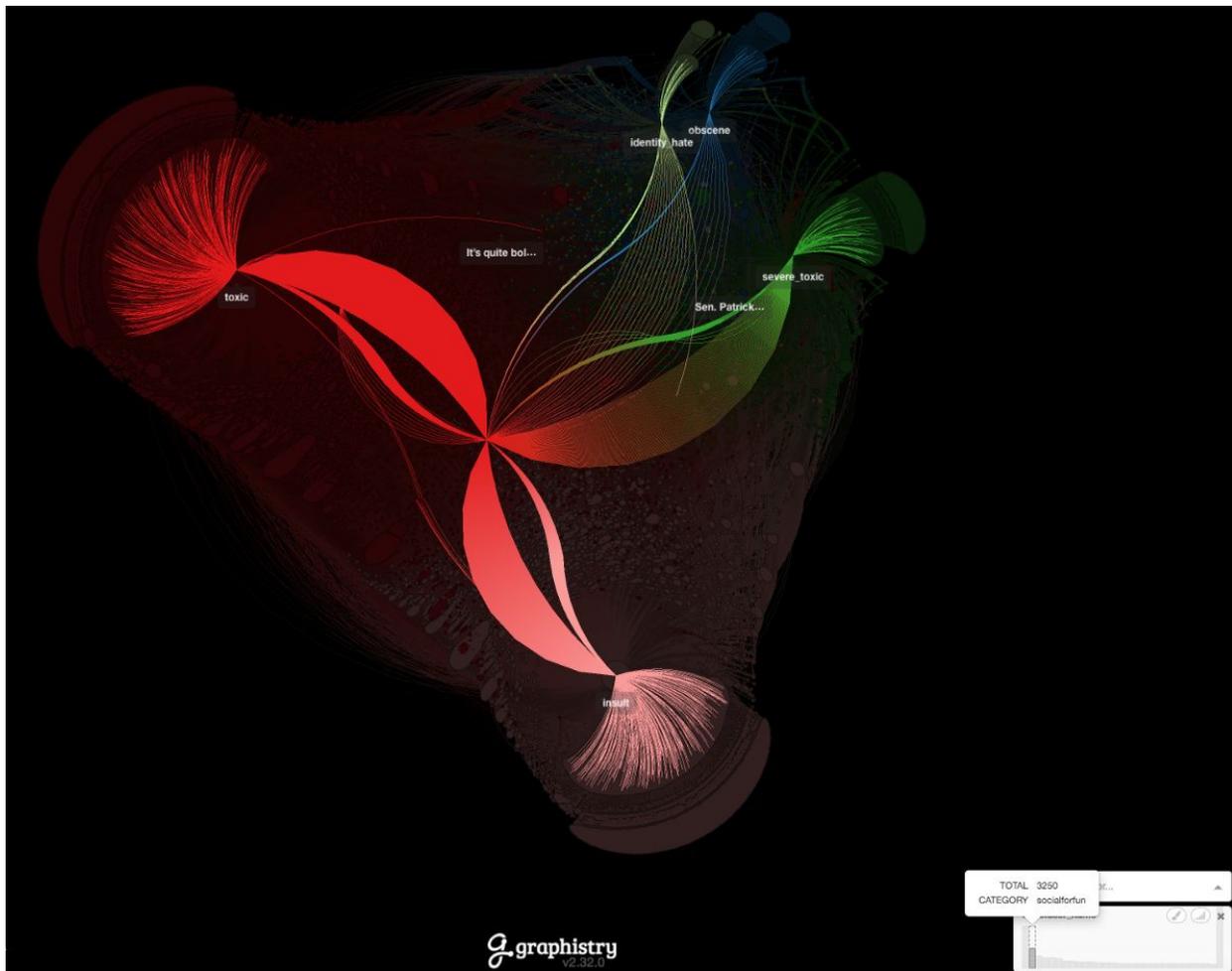


(Above: Examples of tweets spamming hashtags from our Top account.)



Top Accounts By Volume, With Tweets Labeled by Toxicity

#1: Socialforfun



 **Black women are the backbone is an insulting trope**
@socialforfun

I may be blocked but I said what I said- offensive. All skinfolk ain't kinfolk . [@WTFAMERICA20171](#) #ADOS



1:08 PM · Nov 9, 2020 · Twitter Web App

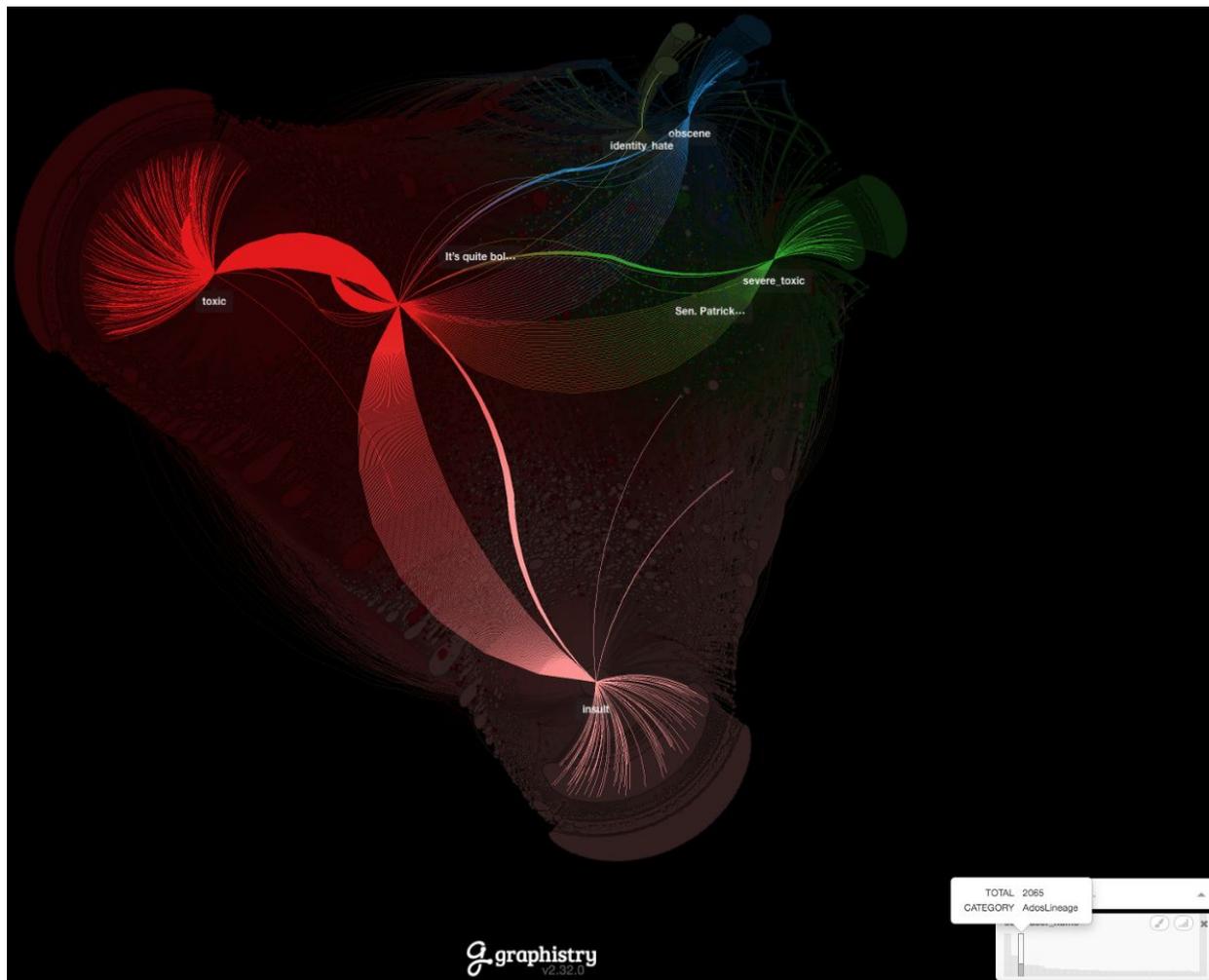
2 Retweets 3 Likes





#2: AdosLineage



 **The ADOS app** @AdosLineage · Nov 10

Anybody that thinks Joe Biden is going to be the 46th president of the United States I have a bridge to nowhere I want to sell you.
[#Reparations](#) [#ADOCs](#) [#adosapp](#) [#ADOS](#)

 **Matt Couch** @RealMattCouch · Nov 9

BREAKING: CNN has rescinded Arizona from former Vice President Joe Biden..
[Show this thread](#)

   2 

 **The ADOS app** @AdosLineage · Nov 18

Look at those Asian thugs

 **Michael Morgan** @mikewhoatv · Nov 17

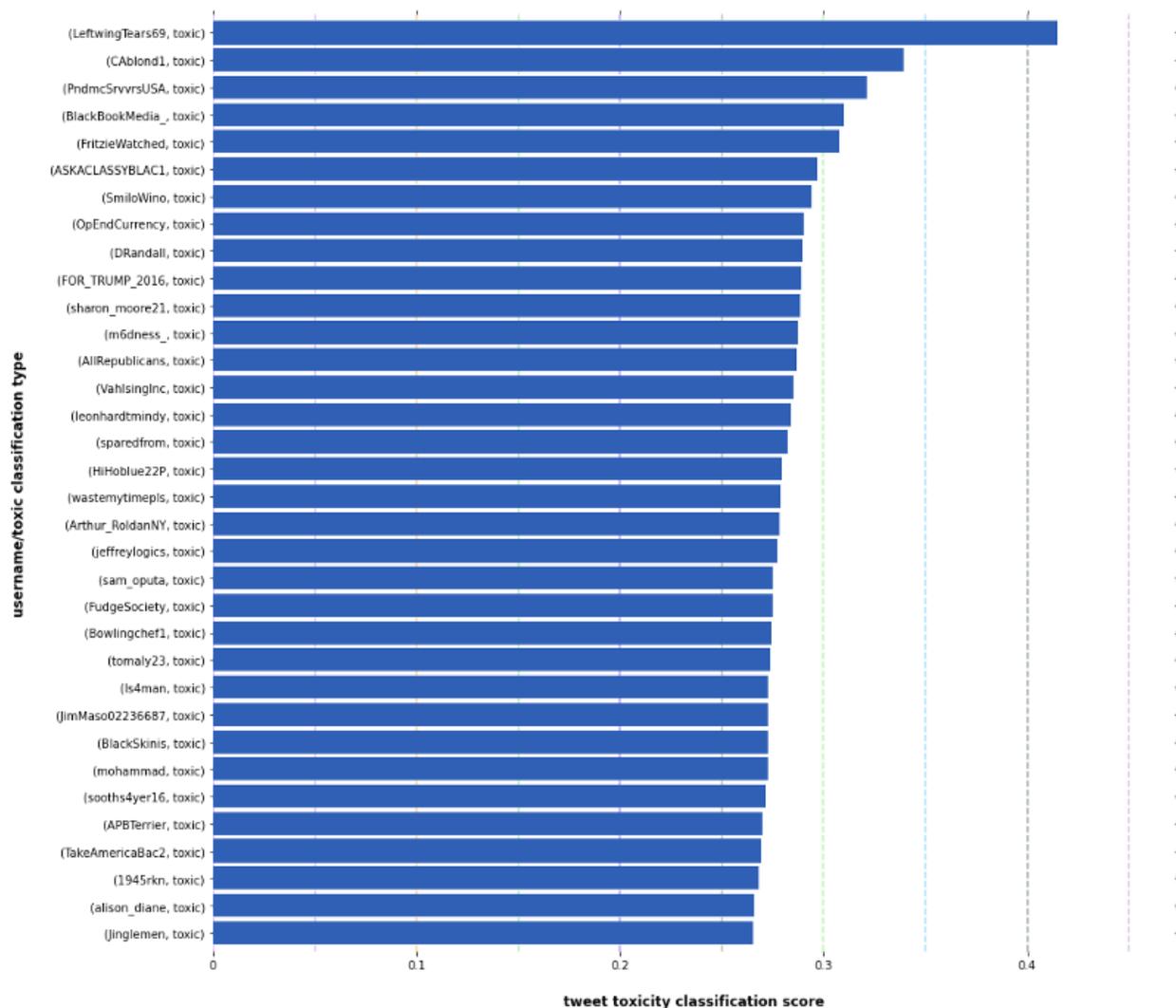
Queens residents take road rage to another level. The twist at the end wasn't what I expected...





Top Trolls by Tweet Toxicity



(Above: Top usernames by tweet Toxicity Score^B.)

The toxicity classification score^B is the highest rated classified tweet accompanying the user within the dataset. We can use this as a heuristic to determine the likelihood that an account is a troll, as well as the degree and type of toxicity the account exhibits via its behaviors.



#1 LeftwingTears69

Molly Jong-Fast 🇺🇸 @MollyJongFast · 7h
Imagine if hunter spent thanksgiving at camp david on the tax payer dime



2.6K 5.5K 26.3K

#LiberalsmisAmentalDisorder
@LeftwingTears69

Replying to @MollyJongFast

Jesus Christ molly? It's like u go out of ur way to prove how biased you are. You are why media has lower approval ratings than Congress. Hacks like you that put your hate for Trump over your love for this country. Instead of joking, why not try actually doing your job?



11:45 AM · Nov 27, 2020 · Twitter for Android

classified by our tweet toxicity analysis) The most toxic tweet of this dataset has with a severe toxic rating of 0.414674 and came from LeftwingTears69. Upon further analysis and collection, we have determined LeftwingTears69 to be an extremely toxic user who only harasses people, and astroturfs the current narrative that is trending in the far-right disinformation spaces. The account frequently harasses journalists like Molly Jong Fast editor of The Daily Beast.

Molly Jong-Fast 🇺🇸 @MollyJongFast · Oct 29
1016 Americans died of coronavirus today.

Acyn Torabi @Acyn · Oct 29
Jr. claims Coronavirus death numbers are down to "almost nothing"



2.1K 8.9K 33K

#LiberalsmisAmentalDisorder
@LeftwingTears69

Replying to @MollyJongFast and @joshrogin

How many of those died with it and how many from it? And how many had prior health issues? How old were most of them? All these important factors, once again, left out by our hack media mouthpieces for the democrats party. After China, alt left media is the most responsible.

9:47 AM · Nov 2, 2020 · Twitter for Android



Senator Mazie Hirono @maziehirono · Nov 18
@evelyndouek asks an excellent question. @YouTube is doing a terrible job stopping election-related misinformation.

That @SenateGOP didn't invite them to yesterday's hearing just shows they don't really care about protecting our elections and only about their metrics.

WIRED @WIRED · Nov 17
Opinion: YouTube's general public relations and governance policy over the years has been to be more opaque, keep its head down, keep quiet, and let the other platforms take the heat.
It's generally gotten away with it wired.trib.al/7aQFIEz

33 46 163

#LiberalismIsAmentalDisorder @LeftwingTears69 · Nov 19
I'm sorry you hate free speech and you hate how free ppl are free to express their hatred towards despicable, warmongering, racists like you. WE HATE YOU! MOST AMERICANS HATE YOU! YOU ARE AN ANTI AMERICAN TRAITOR. #retire #now

1 46 163

Mike Smith! It's a nom de guerre. @BiotechLottery · Nov 19
Replying to @tstorm83_13 and @scrowder
Buy a MAGAT flag and get a lobotomy.

1

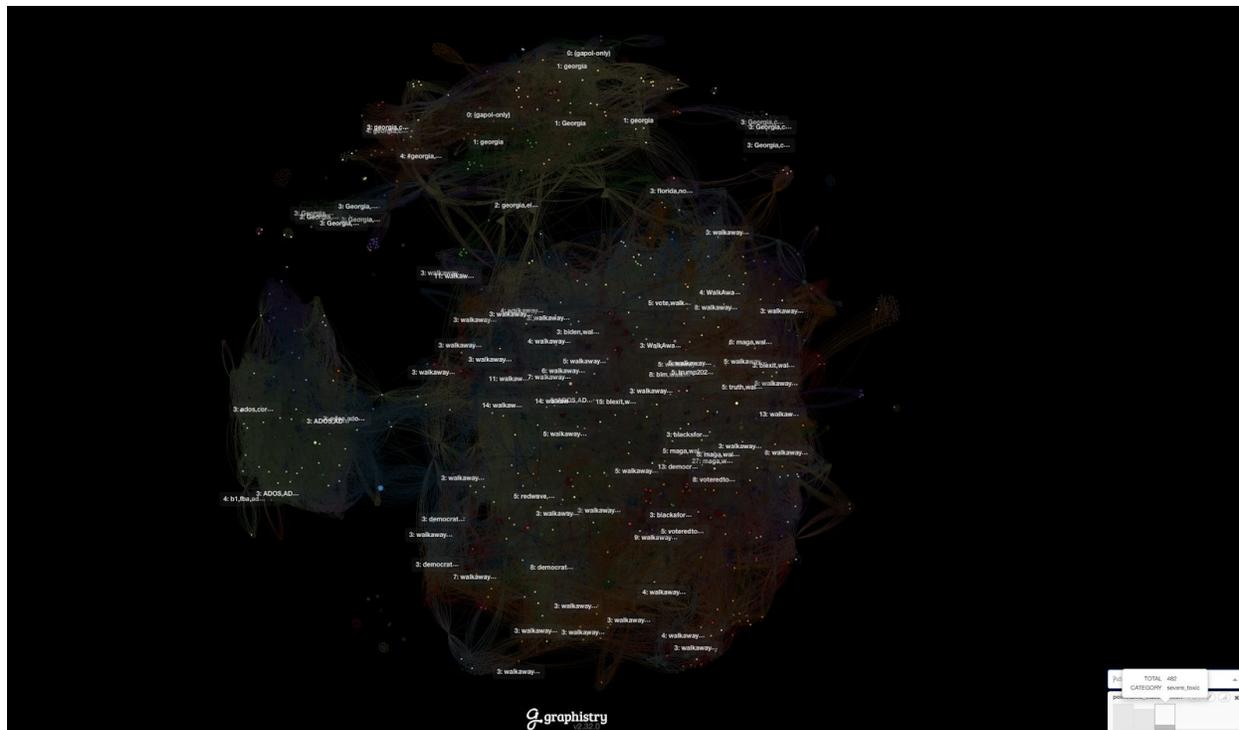
#LiberalismIsAmentalDisorder @LeftwingTears69 · Nov 19
Uh oh. Looks we triggered another alt left snowflake!!! What's a matter? You mad you can't diddle kids in peace anymore? The demonrats not protecting your perversions anymore? #lol #DemsAreDestroyingAmerica



GIF

1





(Above: Filtering users based on toxicity score, Point size is crossover number.)

Combining the analyses allows us to compare hashtag activity with the level of account engagement. The level of overlap presents nodes with toxicity that are connected to hashtags/topic intensity. This gives us a clear visualization of users' toxicity to compare within the network above. We can collect data on which accounts are the most toxic and how often they crossover into other corresponding hashtags.



Conclusion

The major goal of this analysis has been to look at accounts that are spammers, sharing information at high rates, and Trolls, accounts sharing measurably toxic and harmful information. Big Data Analytics and leading Machine Learning models allow for the collecting and analyzing of the exceedingly large amount of data that is produced on platforms such as Twitter. The capabilities of Project Domino's system can be applied to a specific topic or campaign, and paired with disinformation analysts, data scientists, and subject matter experts to present a visual picture of the data. However, there are many challenges when it comes to preventing disinformation, including: 1) a complex focus on censorship and free speech, 2) the need to flip the ratio of the current disinformation ecosystem (6:1, with disinformation in the lead), and 3) the lack of understanding of disinformation's actual impact on victims, democracy, and our society as a whole.

Before we as citizen data scientists can effectively combat and correct disinformation, we must study how environments that promote them exist in the digital universe. Thus far, much of the focus has been on identifying automated accounts (also known as 'bots), yet this type of behavior can be difficult to discern, particularly since cyborgs who produce both organic and amplified content make it ever more challenging to indiscriminately label.

Project Domino and Stop Online Violence Against Women Inc. continue to work together to determine how this tool can be used by journalists, academics, and other researchers wanting to study further about these details, network groupings, as well as the rate of information being shared on specific topics or narratives that impact communities, elections, and the overall future of the disinformation culture we are currently experiencing.



References

1. A threat to an American Democracy: Digital Voter Suppression

<https://stoponlinevaw.com/wp-content/uploads/2020/02/7.pdf>

Convolutional Neural Networks for Twitter Text Toxicity Analysis :

https://link.springer.com/chapter/10.1007/978-3-030-16841-4_38

Quantifying Toxicity and Verbal Violence on Twitter:

<https://dl.acm.org/doi/abs/10.1145/2818052.2869107>



Appendix

A. UMAP HASHTAG ANALYSIS- Each user tweets about different hash tags at different intensities; this is similar to how some songs have a lot of bass, others have a lot of treble, and others have a rich mix at a variety of levels. We calculate the hashtag intensities for each user (akin to an audio pattern fingerprint, but for tweet topic levels) and then run UMAP to automatically group users with similar hashtag behaviors. We connect accounts when they're highly similar, and when many similar ones, pick only the top 20 nearest-neighbors. Our algorithms automatically describe each account and group based on the top hashtags they are promoting. When many accounts are coordinated, they connect into big structures that instantly pop out and reveal how they are being coordinated and for what influence campaigns they're being used for. Our settings normally filter out normal organic users, but if we turn up the gain, we'll find tiny clusters of users interested in the same topics. When visually analyzed, this network visualization can spot the accounts that have the highest rate (crossover number) and largest variety of hashtags shared, which is often behavior portrayed with spam accounts.

B. TOXICITY ANALYSIS- We use BERT uncased with transformers and XLNet on multi-GPU for multi-label classification where we classify each tweet text with a toxicity score for each of 5 labels ("toxic", "severe_toxic", "insult", "obscene", "threat", "identity_hate"). We then visualize the resultant dataset. This Analysis enables the user to be able to analyze who the top Trolls are in a community. The visualization then separates the ranges of the top toxicity score derived from the labeling ("toxic","severe_toxic","insult","obscene","threat","identity_hate") for each tweet. This process allows us to look at the overall dataset and find the most toxic tweets contained within and filter the tweets based on severity of the toxicity rating.

Stop Online Violence Against Women Inc. is a nonprofit that addresses inadequate laws and policies that lack protections for women in particular women of color. We focus on online violence against women, laws, and policy changes needed at the local, state, and federal levels. We also focus on technology and social media company's accountability. SOVAW serves as a resource of services and options for women and women of color, based on their level of targeted harassment or violence. We report on the diverse issues and impacts for those who are willing to share their stories. This includes digital voter suppression. We are a nonpartisan organization.

Project Domino is an open source nonprofit research platform for citizen data scientists to research disinformation across the internet at scale. Disinformation is an insidious threat to our democracy and way of life, we at Project Domino have been working tirelessly to develop tools to understand the source and spread of disinformation at scale in order to shine a light on the growing problem and give citizen data scientists and communities at risk the information they need to make informed decisions in real-time. We are a nonpartisan organization.

