

Identification and Characterization of Cross-Platform (Mis-)information Diffusion

Students

*William Lu
Zexun Yao
Aaron Tom
Duyen Nguyen
Jinyi Ye*

Mentors

*Luca Luceri
Goran Muric
Nikhil Jindal*



Mission: Curb misinformation spreading

Objective:

Identify and characterize the cross-platform diffusion and proliferation of (mis-)information during the Russia-Ukraine conflict.

Goals:

- Investigate how (mis-)information flows across diverse platforms
- Predict which stories will gain traction on mainstream media
- Uncover benevolent activities to share quality information and support suffering population



Study 1: Pathways to Virality

Study Case:

Online messages about the UK-RU conflict

Platforms:

- Facebook
- Twitter
- Instagram
- Bitchute
- Niche platforms
- Media outlets
- Misinformation sites

Task: Identify and characterize cross-platform information diffusion

Study case: Diffusion of links (URLs) within and across platforms

Research Syllabus:

1. Explore the diffusion of URLs and interaction network across multiple platforms (Social network analysis)
2. Investigate viral URLs shared across platforms, their growth and relation over time (Granger causality test and Hawkes process modeling)
3. Identify the sources, diffusion, and content employed to help and support suffering population (Mixed qualitative and quantitative analysis)

Study 2:

Problematic Content Detection

Task: Design an AI approach to automatically detect the diffusion of problematic content on different platforms

Study case: Screenshots of Twitter posts are often shared on diverse platforms, and might contain problematic content moderated by Twitter.

Research Syllabus:

1. Download Twitter Images and manufacture screenshots
2. Supervised machine learning approach based on Convolutional Neural Network
3. Unsupervised machine learning (VG-16 and k -means)

Study 1: Pathways to Virality

Study Case:

Online messages about the UK-RU conflict

Platforms:

- Facebook
- Twitter
- Instagram
- Bitchute
- Niche platforms
- Media outlets
- Misinformation sites

Task: Identify and characterize cross-platform information diffusion

Study case: Diffusion of links (URLs) within and across platforms

Research Syllabus:

1. Explore the diffusion of URLs and interaction network across multiple platforms (Social network analysis)
2. Investigate viral URLs shared across platforms, their growth and relation over time (Granger causality test and Hawkes process modeling)
3. Identify the sources, diffusion, and content employed to help and support suffering population (Mixed qualitative and quantitative analysis)

Study 2:

Problematic Content Detection

Task: Design an AI approach to automatically detect the diffusion of problematic content on different platforms

Study case: Screenshots of Twitter posts are often shared on diverse platforms, and might contain problematic content moderated by Twitter.

Research Syllabus:

1. Download Twitter Images
2. Supervised machine learning approach based on Convolutional Neural Network
3. Unsupervised machine learning (VG-16 and *k*-means)

Study 1: Pathways to Virality

Deliverable 1.1: The diffusion of shared URLs and interaction network across multiple platforms

Network attributes:

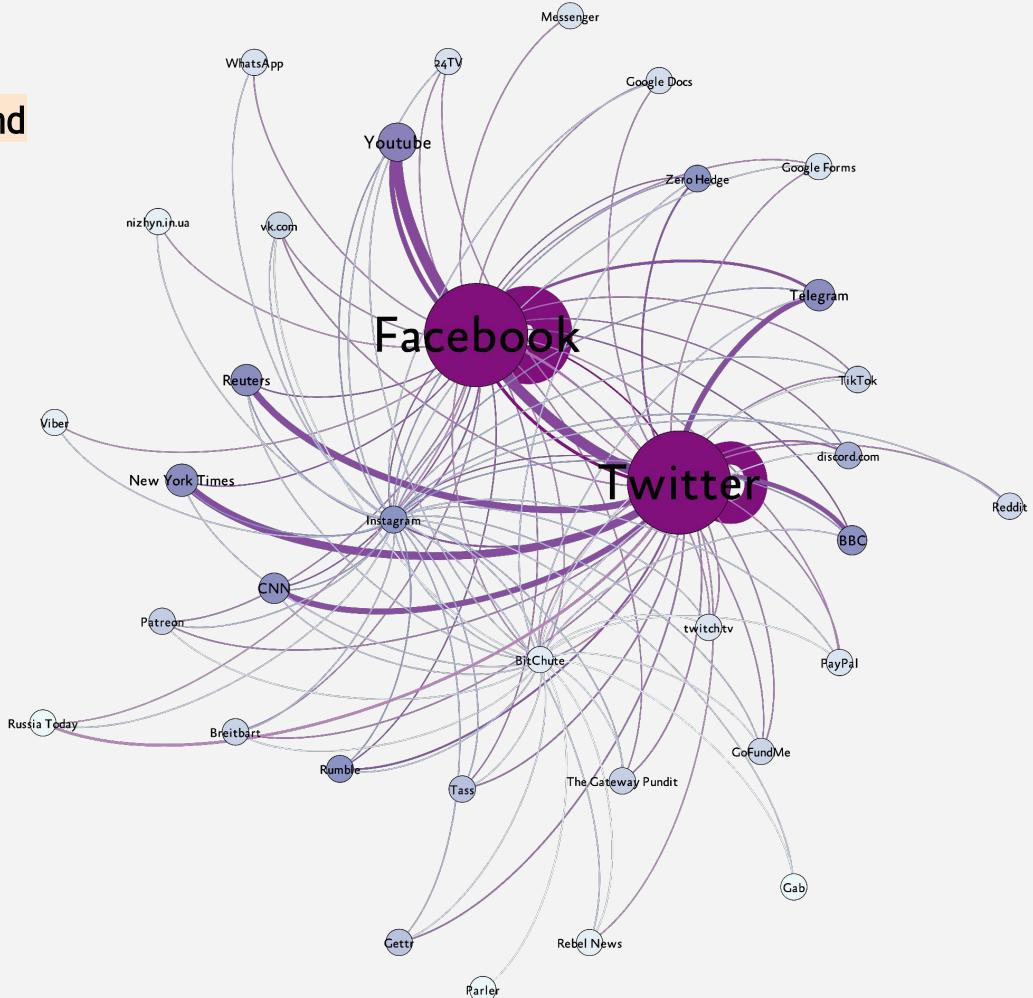
- Nodes: mainstream social media, niche social media, mainstream media, misinformation media, fundraising sites; size and color based on degree centrality
- Edges: the number of URLs linking one platform to the other, width and color based on weights

Data and Method:

- Posts with links related to Russian-Ukraine war from February to April 2022 on 4 platforms: Facebook (9.03M links), Twitter (8.16M links), Instagram (47K links) and Bitchute (3K links)
- Links in the posts are identified, extracted, expanded, parsed based on domain name, counted and ranked in each platform

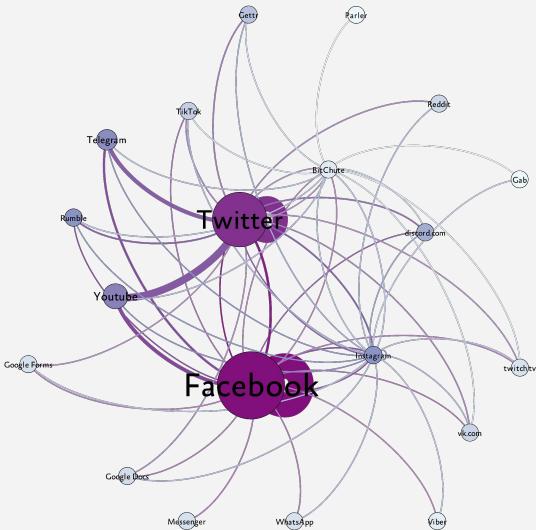
Results:

- Facebook and Twitter are the main hubs of URLs diffusion, with intra-platform links (Facebook 83.5%, Twitter 34.4%) much more than cross-platform links



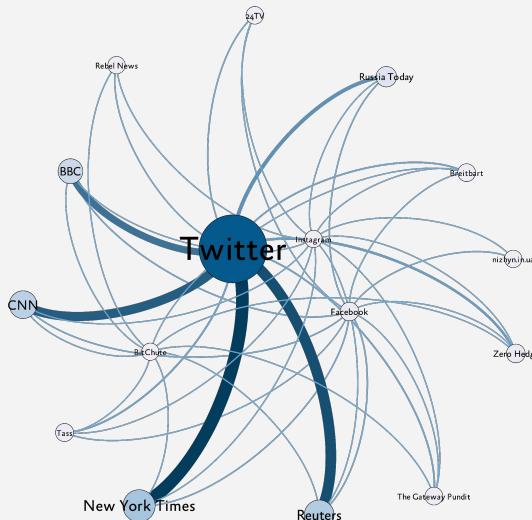
Study 1: Pathways to Virality

Deliverable 1.1: The diffusion of shared URLs and interaction network across multiple platforms



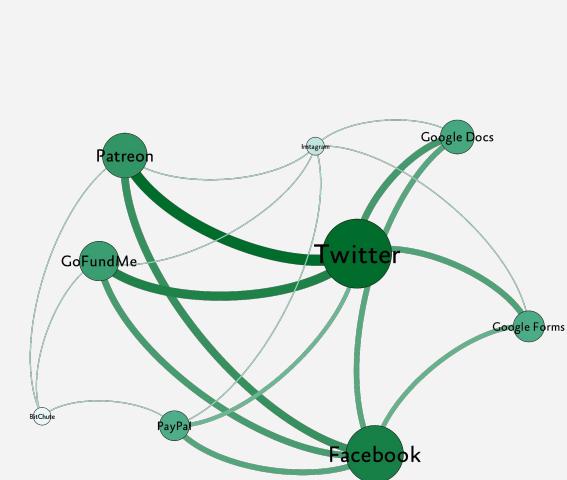
Social Media Network (Mainstream and Niche Platforms)

Niche platforms like Rumble, Gettr, Parler, Gab, Discord, Viber are identified. Messages originated from fringe communities can sometimes get mainstream traction. They account for roughly **2.7%, 9.1%, 36.6% and 5.6%** of all links in Facebook, Twitter, BitChute and Instagram.



Media Network (Mainstream and Misinformation Media Sites)

Misinformation sites like The Gateway Pundit, Rebel News, Zero Hedge and Breitbart are identified. Drawing from the list of low-credibility news sources from iffy.news, links to misinformation sites took up of **0.20%** and **0.58%** of all Facebook and Twitter links, respectively. Top low credibility sites are Rumble, ria.ru, Gettr, sputniknews, Sputnik News.



Online Support Sites

Online support and fundraising sites like Google Docs and Forms, Patreon, PayPal and GoFundMe are identified. Facebook and Twitter are the main venues for fundraising activities, accounting for **99.3%** of all links shared.

Study 1: Pathways to Virality

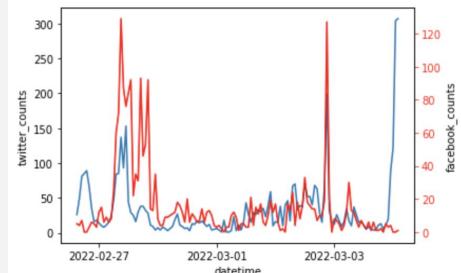
Deliverable 1.2: The diffusion of viral URLs across platforms, their growth, and relation over time

Method: Granger causality tests

Platform 1	Platform 2	Mutual Links	Over 100	Granger Causality Test (in hour, p < .05, maxlag = 10)			
				Direction	Count	Startplace	Peakplace
Twitter	Facebook	675986	219	Twitter - Facebook	33 (15.1%)	Facebook: 169 (77.2%) Twitter: 0	Facebook: 94 (43.0%) Twitter: 111 (50.7%)
				Facebook - Twitter	21 (9.6%)		
Twitter	Instagram	9226	0	Twitter - Instagram	/	/	/
				Instagram - Twitter	/		
Facebook	Instagram	34244	42	Facebook - Instagram	4 (9.5%)	Facebook: 35 (83.3%) Instagram: 1 (2.4%)	Facebook: 25 (59.5%) Instagram: 15 (35.7%)
				Instagram - Facebook	3 (7.1%)		

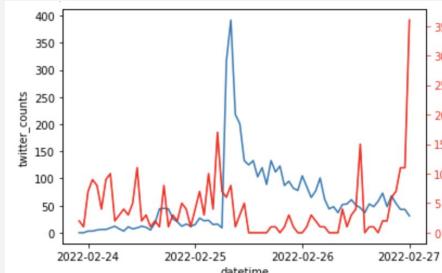
One twitter granger cause facebook case:

$lag_tf = 1, F = 5.308, p < .05$



One facebook granger cause twitter case:

$lag_tf = 2, F = 5.140, p < .01$



Data Processing:

- Mutual links with over 100 shares on Twitter, Facebook and Instagram
- Time frame: from the first appearance to the second peak
- Unit of analysis: hour
- Identify the place where first appearance and the peak occur
- Perform granger causality test within the selected time frame
- Check if the data is stationary before testing

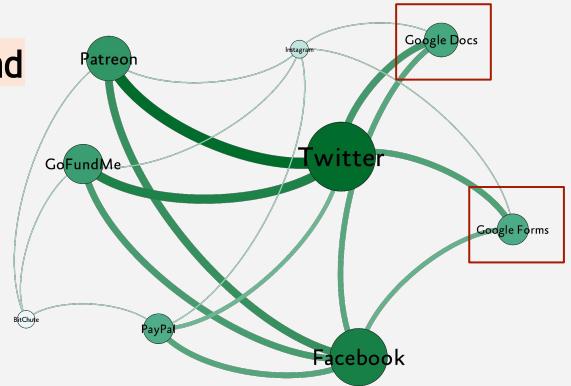


Study 1: Pathways to Virality

Deliverable 1.3: The sources, diffusion, and content employed to help and support suffering population

Methodology:

- Extract web domains from the Twitter and Facebook datasets
- Look at top 50 most frequent links from each data set
- Qualitatively classify each link into categories



Google Forms

- 1 - study
- 2 - sign ups
- 3 - requests
- 4 - petitions
- 5 - evacuation
- 6 - support forms
- 7 - misc
- 8 - deleted/inaccessible links

Google Docs

- 1 - next step
- 2 - donation
- 3 - activity
- 4 - support
- 5 - petitions
- 6 - misc
- 7 - deleted/inaccessible links

expandedLink	expandedLink_Count	domain	Categorization
0 https://forms.gle/rPv7mJau8	10743	forms.gle	6
1 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	783	docs.google.com	5
2 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform?usp=link	467	docs.google.com	4
3 https://forms.gle/1EHX0hNyyg5z7	441	forms.gle	7
4 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform?edit_requested=true	433	docs.google.com	4
5 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform?usp=link	422	docs.google.com	7
6 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	399	docs.google.com	1
7 https://forms.gle/Yaku3NtKwC1h4Z	302	forms.gle	6
8 https://forms.gle/2YGS3wqEUeWw448	257	forms.gle	6
9 https://forms.gle/JLTGzHn4qo4f16	252	forms.gle	7
10 https://forms.gle/mIbzbQnDwPRhLBjW8	240	forms.gle	2
11 https://forms.gle/RhNpxwsZDF9Gw4k8	232	forms.gle	6
12 https://forms.gle/vQdLmk0Bm0mU3	204	forms.gle	6
13 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	196	docs.google.com	6
14 https://forms.gle/P3S2J4Z7B7Dmtm1	157	forms.gle	6
15 https://forms.gle/S2H2UrJnJwMM1PTA	157	forms.gle	6
16 https://forms.gle/NSPfJ6oas5kV9GA	152	forms.gle	6
17 https://forms.gle/WWWqH7C88pY7Njus	121	forms.gle	2
18 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform?usp=ef_link	112	docs.google.com	5
19 https://forms.gle/Z5TfJXenea7	108	forms.gle	5
20 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	108	docs.google.com	2
21 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	90	docs.google.com	7
22 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	84	docs.google.com	6
23 https://forms.gle/4A7DfLQzQxJt0Z	83	forms.gle	1
24 https://forms.gle/kNWhMvYtakJ5C	82	forms.gle	8
25 https://forms.gle/T7mUuM892XqgP9	80	forms.gle	8
26 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	78	docs.google.com	2
27 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	75	docs.google.com	2
28 https://forms.gle/nmm3avXQDQfTu8	75	forms.gle	6
29 https://forms.gle/Q7tNBLETWey7N7	69	forms.gle	3 suspicious
30 https://forms.gle/Lt6mMmSM1SNhlyk	65	forms.gle	5
31 https://docs.google.com/forms/u/1/t/1ApQLScxVhPy5EcHfDM5c_yQ7EcPLH_68reKmWb3D159Eew/viewform	64	docs.google.com	8

Study 1: Pathways to Virality

Deliverable 1.3: The sources, diffusion, and content employed to help and support suffering population

Google Docs: Sheets | Docs | Slides

Twitter Google Docs

- 1369 unique Google docs/slides/spreadsheets
- 14472 instances of links

Facebook Google Docs

- 1883 unique Google docs/slides/spreadsheets links
- 4423 instances of links

DEAR READERS: This is a live, dynamic document. There will be formating anomalies and irregular spacing due to the nature of the document. Please do not mind the styling. GitHub pull requests are welcome to see if anything is new. Sources, updates or new information and constructive feedback is welcomed via GitHub. All content inclusion is up to the editors. Best efforts will be made to keep this document up to date. Please note that this is a live document.

Reference 1	Reference 2	Reference 3	Reference 4	Reference 5

Лист за відставку Гетьманцева

Заступнику Володимиру Олександровичу!

Шановнику Валентину Семеновичу!

Моє підписання єдине, якщо буде уявленням інтересів України.

Підписано мною пропозиція Генералі України з часу нафбасу розглянути сюди згідні. І проханою діяти спільно — із згадкою, що розумію, що у подальшому виникніть, якщо не будуть перед Україною.

Вам відомо, що більш з першого годин видав підтримку Збройних сила і Терористичну оборону, надігруючи на це.

Ця пансько звернення — пропозиція Яну узяти до катастрофічних пошуків чи складом руйнівний економічний та соціальний розвиток України.

Ви відмінно для нас і новий прогрес — сконструюєте. Так, на поточний момент 49% підприємств не працюють (або майже зупинили). Але за останні місяці переважно майже всі МСБ (хоча останнім часом зупинилися) працюють з колись меншою, якщо створюють ринок для бізнесу. І змусили близько 10% підприємств вже проводити зупинки відсутності в корпорації.

Цей ринок створюється за кордоном. Це — економіка війни. Оскільки, Україна посідає 110-е місце у рейтингу економічного сектора — що з інспекцією неприватності корінно для видення бізнесу.

Спогади людів в Україні вони в 1,5 раз більше чиїх, ніж у середньому у світі. Конгруєтно, що неможливо без експортної. А прорахунки цифрової митниці дійшли членів правління, що обирають обговорювати чи требує про громадські.

Після презентації, як мані сознання членів нам довести сукупність економіку дому, яку відійде переважно більшість освіти і розвинутого сектора, в Україні, а також українським бізнесом.

Спогади людів в Україні вони в 1,5 раз більше чиїх, ніж у середньому у світі. Конгруєтно, що неможливо без експортної. А прорахунки цифрової митниці дійшли членів правління, що обирають обговорювати чи требує про громадські.

Тому вільне не можна відкладати — після Гетьманцева переважно сказати, — що він з активним працюючим маркетплейсом, соціальних ційл та радикального способу життя, противним ринковим

показникам та методом життя від Данило Гетьманцева.

Він з постійними прогресивними економічніми секторами, працюючим коректним державним регулюванням та підтримкою соціальної справедливості.

Наукові праці та професійні діяльність пана Гетьманцева переважно сказати, — що він з активним

Digital Humanitarians Organising for Ukraine

Listen to the Twitter session live here:

- Sat, 16 Apr 2022
- 9 AM Pacific
- 12 PM Eastern
- 5 PM London
- 7 PM Kyiv

Twitter Spaces Briefing: 16 Apr 2022

Study 1: Pathways to Virality

Deliverable 1.3: The sources, diffusion, and content employed to help and support suffering population

Google Forms

Twitter Google Forms

- 1358 unique Google Form links
- 24088 instances of links

Facebook Google Forms

- 8191 unique Google Form links
- 24292 instances of links

Hauptbahnhof: Volunteer Sign Up

Thank you for joining our effort to support BIPOC people fleeing Ukraine.

You are volunteering in a joint initiative of Each One Teach One (EOTD) e.V., International Women's Space, Institute of Black People in Germany (IBG) and MyRefugee Berlin e.V. By collaborating Berlin Arrival Support and other organizations and initiatives, we connect BIPOC refugees fleeing the war in Ukraine with housing, advice, meals, legal resources and information so everyone can make the best possible decisions for themselves. We place special emphasis on keeping families and refugee groups together and countering the effects of racism that BIPOC individuals may have experienced while fleeing war.

There are two locations of this project you can volunteer in:

- Berlin central station (Hauptbahnhof) for first arrival support, e.g. providing them with basic supplies at the station, assisting them to travel further, handing people the most important info or directing them to different places in Berlin for further support. Sign up under: <http://tiny.cc/VolunteerHBB>

- CUSBU Project Office, which is a community support office for matching people with accommodations, assisting with paperwork, accompanying to government offices, help with accessing legal advice, warm meals and more. Sign up under: <http://tiny.cc/VolunteerM574>

At both locations you will be briefed and supervised, so don't worry if this is your first time supporting in an emergency response and social support project.

Just pick the location, date and time that suits you best to volunteer and we'll be happy to work with you!

Please find further instruction on completion of the form. And again: Thank you!

[actom@usc.edu \(not shared\)](#) [Switch account](#)

* Required

The screenshot shows a Twitter post with a blue header banner. The banner features two women in red shirts, one smiling and one pointing, with the text "与其练1万种招式，不如把这一招练一万次！" (Rather than practicing 10,000 moves, why not practice this one 10,000 times!). Below the banner, the main text reads: "5天学会的股票投资策略，终身免费1对1的复习". A "FREE STOCK TRADING CLASS, FREE LIFETIME REVISION" button is visible. Below this, there is a "Required" field for an email address: "actom@usc.edu (not shared)" with a "Switch account" link. At the bottom, there is a video thumbnail showing a person speaking in Chinese, with the caption "2022年第6期的交易... 来了解吧！" (2022 Year 6 Trading... Come and learn!).



Refugees Token | Airdrop - Donation

Refugees Token has launched its airdrop program, where it is giving away free RFG Tokens to anyone who joins the campaign and performs a series of simple social activities.

We will be donating \$0.5 from each participant to the people affected from the Ukraine-Russia war!
(Example: 10,000 participants = \$5,000 donation)

Complete all airdrop round tasks.

Airdrop Date: Available until 21/04/22 - Airdrop Amount: \$10,000 in RFG tokens.

Everyone who buys tokens during the Airdrop campaign will be rewarded an extra amount.

Perform the following tasks to complete the campaign and win the rewarding RFG tokens.

[actom@usc.edu \(not shared\)](#) [Switch account](#)

* Required



Registration form for Indian citizens stranded in Ukraine

Embassy of India, Kyiv requests all Indian citizens stranded in Ukraine, and are yet to depart from their respective addresses and want to be evacuated urgently to fill up this form.

[actom@usc.edu](#) [Switch account](#)

* Required

Email *

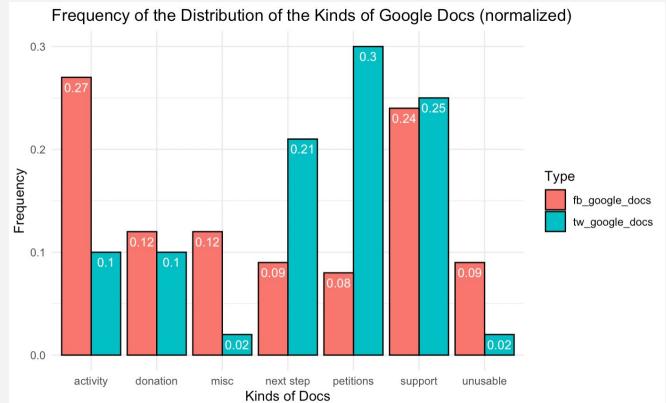
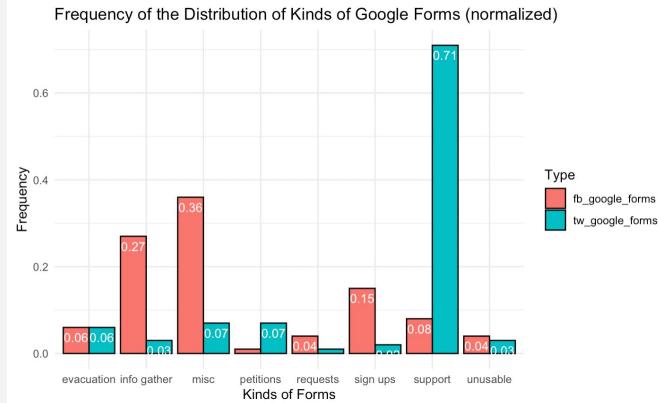
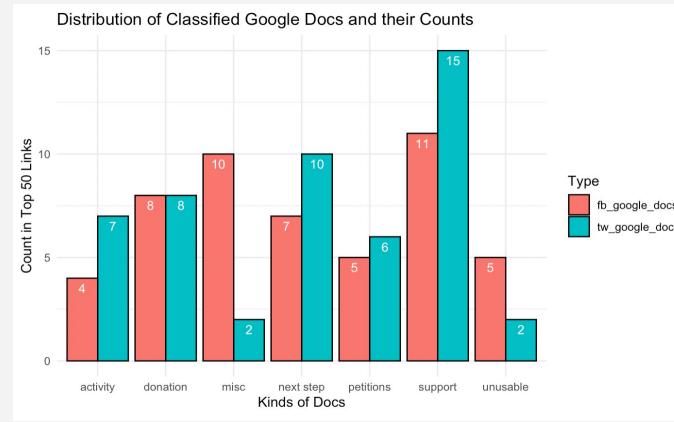
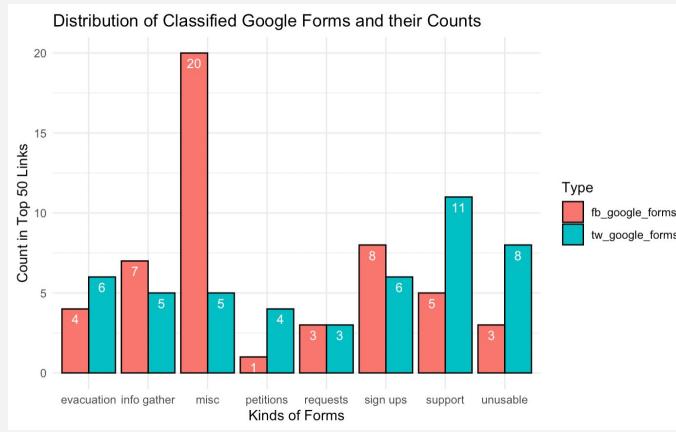
Your email

Full Name (As in passport, Surname followed by Given Name) *

Your answer

Study 1: Pathways to Virality

Deliverable 1.3: The sources, diffusion, and content employed to help and support suffering population



Study 1: Pathways to Virality

Study Case:

Online messages about the UK-RU conflict

Platforms:

- Facebook
- Twitter
- Instagram
- Bitchute
- Niche platforms
- Media outlets
- Misinformation sites

Task: Identify and characterize cross-platform information diffusion

Study case: Diffusion of links (URLs) within and across platforms

Research Syllabus:

1. Explore the diffusion of URLs and interaction network across multiple platforms (Social network analysis)
2. Investigate viral URLs shared across platforms, their growth and relation over time (Granger causality test and Hawkes process modeling)
3. Identify the sources, diffusion, and content employed to help and support suffering population (Mixed qualitative and quantitative analysis)

Study 2:

Problematic Content Detection

Task: Design an AI approach to automatically detect the diffusion of problematic content on different platforms

Study case: Screenshots of Twitter posts are often shared on diverse platforms, and might contain problematic content moderated by Twitter.

Research Syllabus:

1. Download Twitter Images and manufacture screenshots
2. Supervised machine learning approach based on Convolutional Neural Network
3. Unsupervised machine learning (VGG-16 and k -means)

Study 2: Moderated Posts Detection

Deliverable 2.1: Download Twitter images and manufacture screenshots

Data gathering

- Download 1.5 million Twitter images
- Manufacture screenshots using the 'tweet-capture' package
- Leverage Kaggle screenshots dataset

t 1 Dave Retweeted

The Spectator Index
SPECTATOR INDEX
@spectatorindex

BREAKING: France says EU will impose sanctions against Russia and calls Putin's speech 'paranoid'

2:43 PM · Feb 21, 2022 · TweetDeck

775 Retweets 200 Quote Tweets 3,993 Likes

t 1 Torquemada's Kat 🐱.cat 🇺🇦 🇷🇺 🇺🇦 Retweeted

Oryx
@oryxspioenkop

In today's false flag operation Russia used a BTR-70M APC to depict a Ukrainian armoured vehicle.

Ukraine doesn't operate the BTR-70M. They couldn't even get that right.

oryxspioenkop.com/2022/02/false-...



2:21 PM · Feb 21, 2022 · Twitter for iPhone

977 Retweets 113 Quote Tweets 2,300 Likes

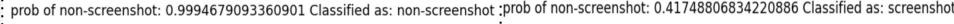
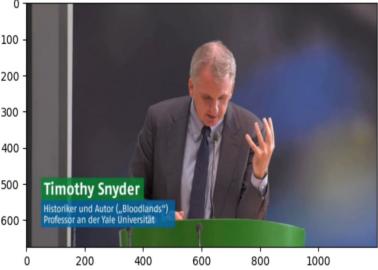
   

Study 2: Moderated Posts Detection

Deliverable 2.2: Design a supervised learning model to detect Twitter screenshots using CNN

	First scenario	Second scenario
Train dataset	<ul style="list-style-type: none"> • 2808 tweet screenshots (tweet-capture) • 2808 non-tweet screenshots (twitter) 	<ul style="list-style-type: none"> • 3510 tweet screenshot • 3509 non-tweet screenshots
Test dataset	<ul style="list-style-type: none"> • 702 screenshot (tweet-capture) • 701 non-tweet screenshots (twitter) 	<ul style="list-style-type: none"> • 200 unlabeled Twitter images
Result	<ul style="list-style-type: none"> • Accuracy: 0.989 • F1: 0.989 • Recall: 0.983 • Roc-Auc: 0.989 	<p>Use 100 images of each class for evaluation</p> <ul style="list-style-type: none"> • Accuracy: 0.76 • F1: 0.689 • Recall: 0.981 • Precision: 0.53

prob of non-screenshot: 0.9996232986450195 Classified as: non-screenshot

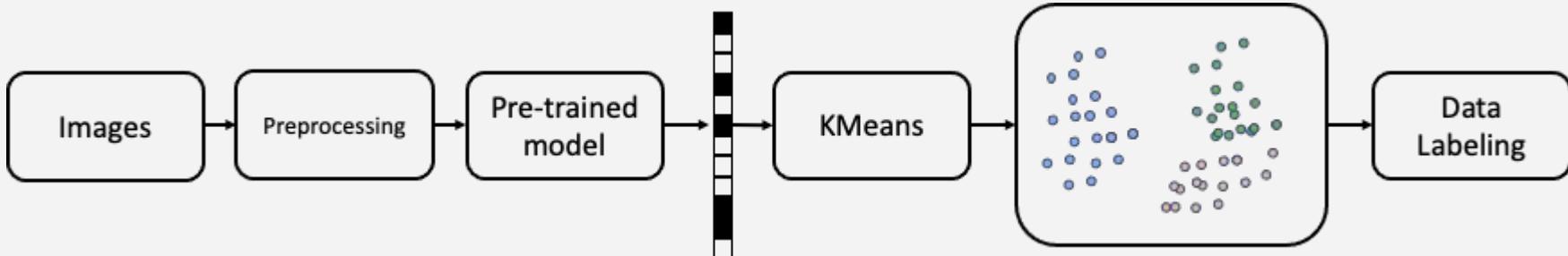


Study 2: Moderated Posts Detection

Deliverable 2.3: Design an unsupervised learning model to detect Twitter screenshots

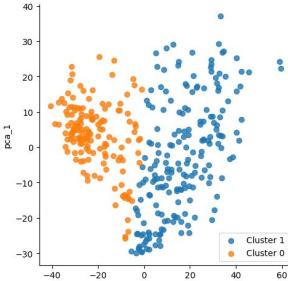
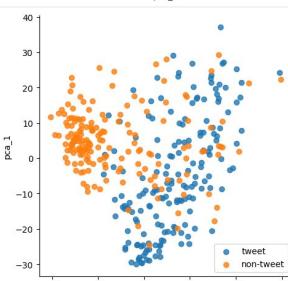
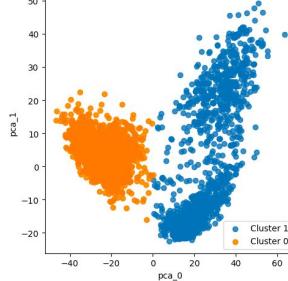
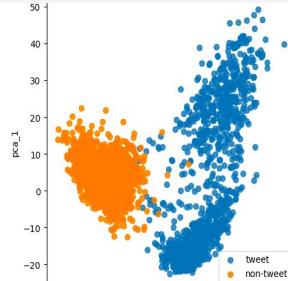
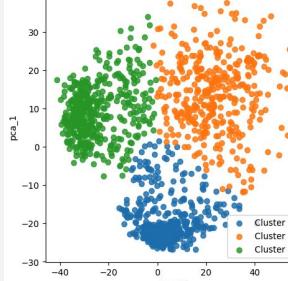
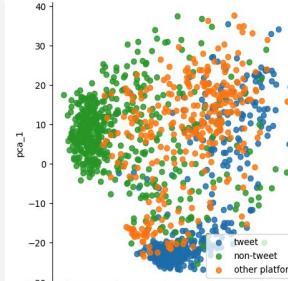
Methodology: Cluster images based on visual similarity to identify tweet screenshot.

- Clean and preprocess images.
- Extract features from the images using pre-trained model VGG16.
- Use PCA to reduce dimensions for visualization.
- Cluster the features with K-means.
- Assign labels to data points based on cluster results.



Study 2: Moderated Posts Detection

Deliverable 2.3: Design an unsupervised learning model to detect Twitter screenshots

	Binary clustering	Multiclass clustering	
Data	<ul style="list-style-type: none">■ 174 tweet screenshots (Kaggle)■ 174 random images	<ul style="list-style-type: none">■ 2000 tweet screenshots (Tweet-Capture)■ 2000 random images	
Result	  <p>ROC AUC: 0.804 Accuracy: 0.804 F1: 0.816 Precision: 0.77 Recall: 0.867</p>	  <p>ROC AUC: 0.995 Accuracy: 0.995 F1: 0.995 Precision: 0.998 Recall: 0.993</p>	  <p>ROC AUC: 0.824 Accuracy: 0.715 F1: 0.7 Precision: 0.7 Recall: 0.7</p>

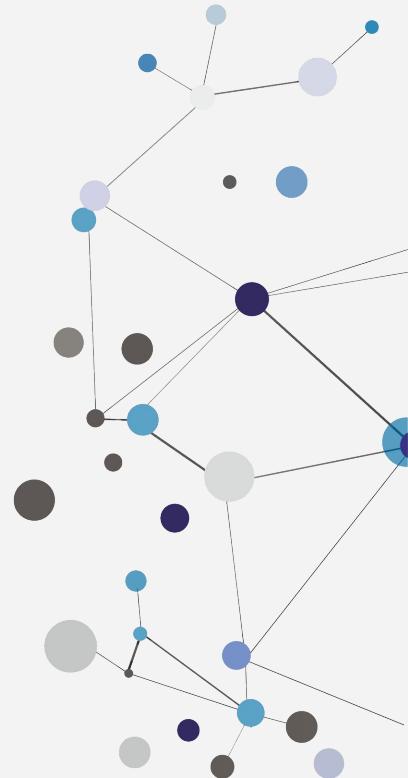
Discussion

Conclusion

- We explored (mis-)information sharing across the information ecosystems
- We identified means and resources for supporting suffering population in times of crisis
- We designed and tested AI approaches to detect Twitter screenshots

Future works

- Expand our study to a larger information ecosystem by including more niche platforms and instant messaging apps
- Enlarge the suite of statistical techniques used to assess causality (ongoing: Hawkes process modeling)
- Enhance AI approaches and validate them into the wild to identify screenshots across platforms



Our Research Team



Student Researcher
Jinyi Ye



Student Researcher
William Lu



Student Researcher
Zexun Yao



Student Researcher
Duyen Nguyen



Student Researcher
Aaron Tom



Mentor
Luca Luceri

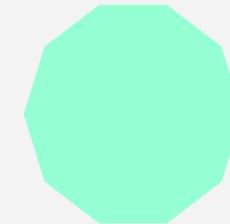


Mentor
Goran Muric



Mentor
Nikhil Jindal

Thanks!



Visit Our Project Website:

https://zexunyao.github.io/crossplatform_misinfo.github.io/

Contact us on:

