# Tactics, Threats & Targets: Modeling Disinformation & its Mitigation

Shujaat Mirza, Liang Niu

Paolo Papotti

Labeeba Begum, Sarah Pardo, Azza Abouzied, Christina Pöpper

Network and Distributed System Security (NDSS)
March 2nd, 2023

# Disclaimer

- In the course of this talk, we may be presenting views that challenge our beliefs or perspectives especially if we were exposed to certain narratives that are legitimized by our communities, our leaders, our loved ones, etc.

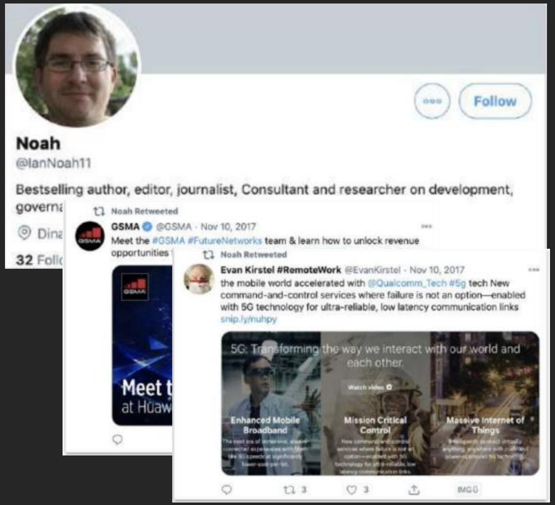
- In the presentation of disinformation case studies, we are not taking sides in any discourse

- There are no angels: different parties on different sides take part in disinformation

- We are all victims of disinformation: our goal is to understand it and not to favor a side or a viewpoint

# Fake cluster attacks Belgian Government & boosts Huawei

https://graphika.com/reports/fake-cluster-boosts-huawei

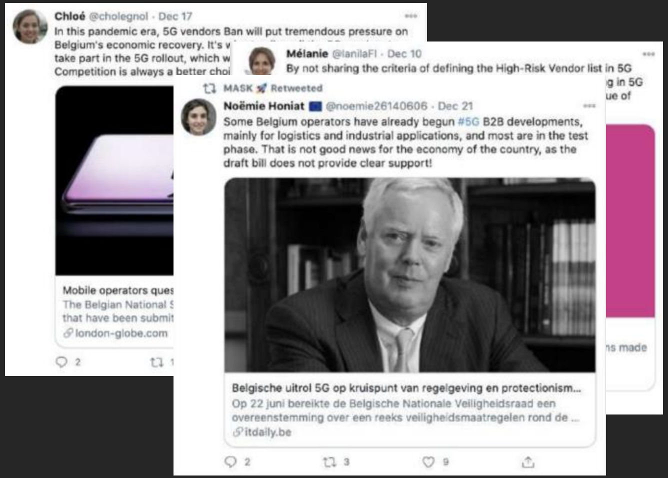


Noah
@IanNoah11
Bestselling author, editor, journalist, Consultant and researcher on development, governa
Dina
32 Follo

Noah Retweeted
GSMA @GSMA · Nov 10, 2017
Meet the #GSMA #FutureNetworks team & learn how to unlock revenue opportunities

Noah Retweeted
Evan Kirstel #RemoteWork @EvanKirstel · Nov 10, 2017
the mobile world accelerated with @Qualcomm_Tech #5g tech New command-and-control services where failure is not an option—enabled with 5G technology for ultra-reliable, low latency communication links snip.ly/nuhpy

Follow
Meet t
at Huaw
5G: Transforming the way we interact with our world and each other.
Watch video
Enhanced Mobile Broadband
Mission Critical Control
Massive Internet of Things
3
3



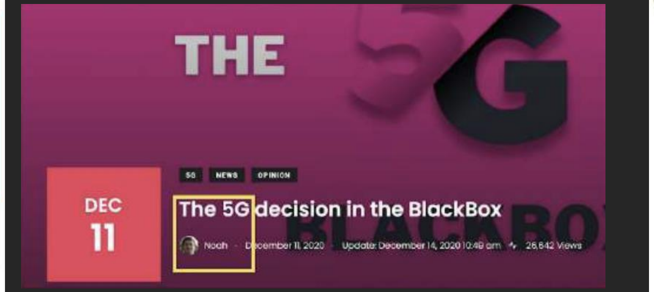
Chloé @cholegnol · Dec 17
In this pandemic era, 5G vendors Ban will put tremendous pressure on Belgium's economic recovery. It's w
take part in the 5G rollout, which w
Competition is always a better choi

Mobile operators ques
The Belgian National S
that have been submit
london-globe.com

Mélanie @IanilaFl · Dec 10
By not sharing the criteria of defining the High-Risk Vendor list in 5G
g in 5G
ue of
ns made

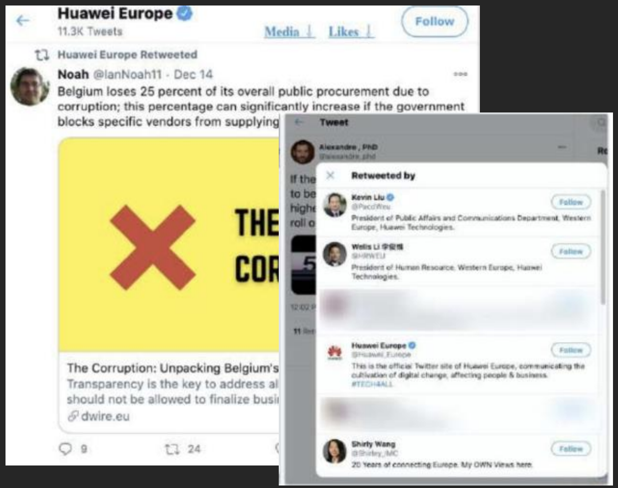
MASK 🚀 Retweeted
Noëmie Honiat 🇪🇺 @noemie26140606 · Dec 21
Some Belgium operators have already begun #5G B2B developments, mainly for logistics and industrial applications, and most are in the test phase. That is not good news for the economy of the country, as the draft bill does not provide clear support!

Belgische uitrol 5G op kruispunt van regelgeving en protectionism...
Op 22 juni bereikte de Belgische Nationale Veiligheidsraad een overeenstemming over een reeks veiligheidsmaatregelen rond de ...
itdaily.be



The 5G decision in the BlackBox

Noah · December 11, 2020 · Update: December 14, 2020 10:49 am · 26,642 Views



Huawei Europe
11.3K Tweets

Huawei Europe Retweeted
Noah @IanNoah11 · Dec 14
Belgium loses 25 percent of its overall public procurement due to corruption; this percentage can significantly increase if the government blocks specific vendors from supplying

The Corruption: Unpacking Belgium's
Transparency is the key to address al
should not be allowed to finalize busi
dwire.eu

Tweet
Alexandre , PhD
If the
to be
high
roll o

Retweeted by
Kevin Liu
President of Public Affairs and Communications Department, Western Europe, Huawei Technologies.
Wells Li
President of Human Resource, Western Europe, Huawei Technologies.
Huawei Europe
@Huawei_Europe
This is the official Twitter site of Huawei Europe, communicating the cultivation of digital change, affecting people & business. #TECH4ALL
Shirly Wang
20 Years of connecting Europe. My OWN Views here.

July 2017

Noah, a best selling author & journalist, makes Twitter account & gains 2000+ followers

2017-18

Noah retweets tech-related content

Nov/Dec 2020

Multiple Brussel based accounts start tweeting articles about why Belgium's 5G decision is a bad idea

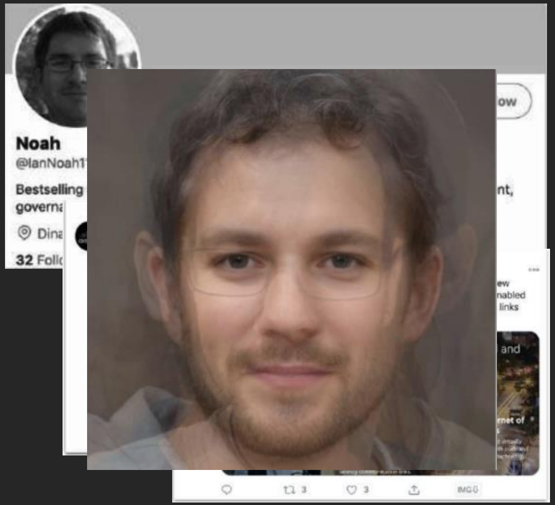
Some of these articles are written in Dutch languages

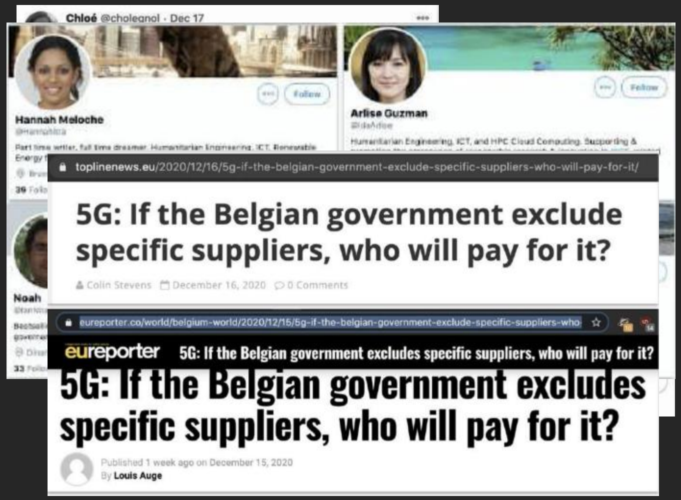
Early Dec 2020

- Noah authors articles titled 'The 5G decision in the BlackBox' published on Brussel based website dwire

Mid Dec 2020

The content finds widespread attention on the platform as highlighted by retweets by many accounts, including those belonging to Huawei




Chloé @cholegnol · Dec 17

Hannah Meloche

Arlisa Guzman

toplinenews.eu/2020/12/16/5g-if-the-belgian-government-exclude-specific-suppliers-who-will-pay-for-it/

5G: If the Belgian government exclude specific suppliers, who will pay for it?

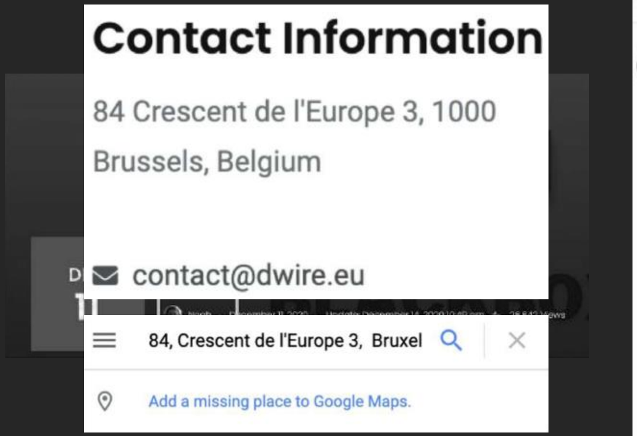
Colin Stevens · December 16, 2020 · 0 Comments

eureporter.co/world/belgium-world/2020/12/15/5g-if-the-belgian-government-exclude-specific-suppliers-who

eureporter 5G: If the Belgian government excludes specific suppliers, who will pay for it?

5G: If the Belgian government excludes specific suppliers, who will pay for it?

Published 1 week ago on December 15, 2020
By Louis Auge

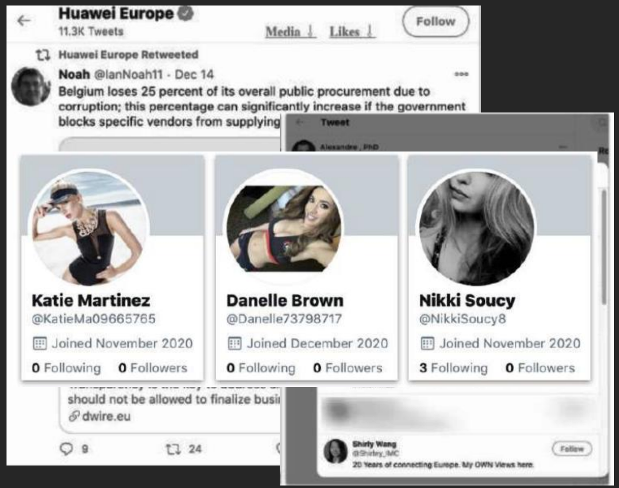

Contact Information

84 Crescent de l'Europe 3, 1000
Brussels, Belgium

✉ contact@dwire.eu

84, Crescent de l'Europe 3, Bruxel

Add a missing place to Google Maps.


Huawei Europe
11.3K Tweets
Follow

Huawei Europe Retweeted
Noah @IanNoah11 · Dec 14
Belgium loses 25 percent of its overall public procurement due to corruption; this percentage can significantly increase if the government blocks specific vendors from supplying
should not be allowed to finalize busi
dwire.eu

| Katie Martinez | Danelle Brown | Nikki Soucy |
|---|---|---|
| @KatieMa09665765 | @Danelle73798717 | @NikkiSoucy8 |
| Joined November 2020 | Joined December 2020 | Joined November 2020 |
| 0 Following 0 Followers | 0 Following 0 Followers | 3 Following 0 Followers |

July 2017

2017-18

Nov/Dec 2020

Early Dec 2020

Mid Dec 2020

The identity of Noah was made up!

His profile used GAN generated photo

Noah's account was part of 14 accounts that has similar activity pattern

Same script, different domains were used

Dwire is a fake news site with made up address

Network of bots used to boost amplification

Huawei executives interactions with posts boosted amplification

Insufficient evidence on source of campaign

Discovered by Graphika

# Fake Cluster Boosts Huawei

# Disinformation Campaign

A disinformation campaign or operation is a coordinated effort by individuals or groups to manipulate public opinion and change how people perceive events in the world by intentionally producing or amplifying disinformation [1]

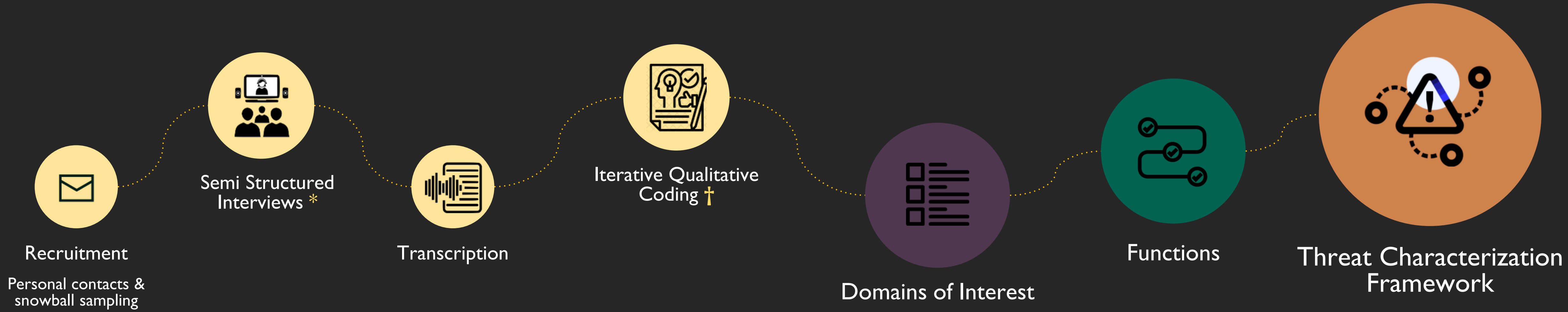
[1] Wilson, T., & Starbird, K. (2020). Cross-platform disinformation campaigns: lessons learned and next steps. *Harvard Kennedy School Misinformation Review.*

# Modeling Attacks

Why think of disinformation through a cybersecurity lens?

- Systemization
- Detection
- Prioritization
- Countermeasures

# Constructing the Modeling Framework



Recruitment

Personal contacts & snowball sampling

Semi Structured Interviews *

Transcription

Iterative Qualitative Coding †

Domains of Interest

Functions

Threat Characterization Framework

* All interviews took place over Zoom between July and November 2021; the interviews lasted from 30 minutes up to 1 hour

† Four authors reviewed the interviews independently

# Constructing the Modeling Framework

An interview study of 22 Mitigators from 19 different organizations

Security

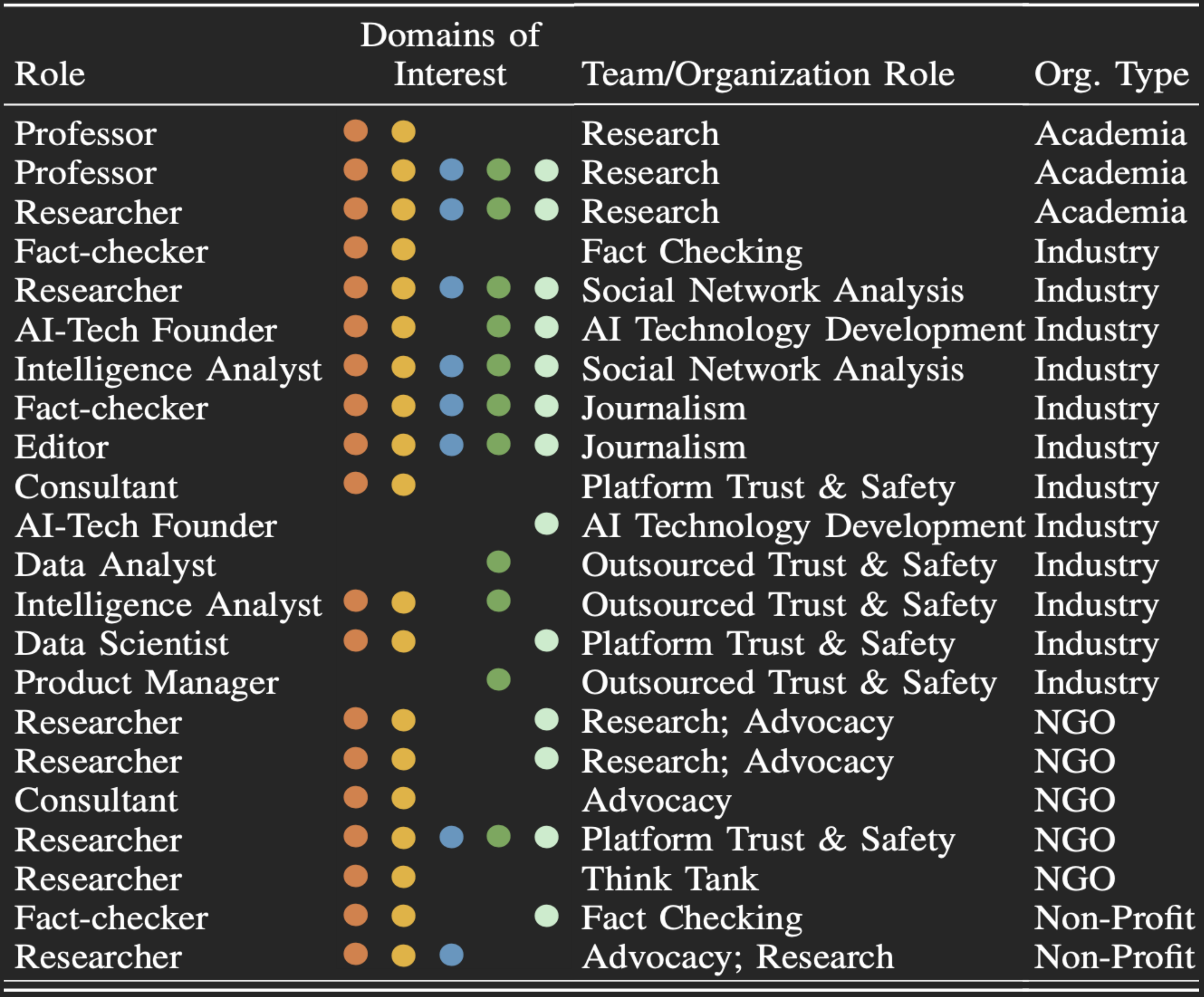
Democracy

Economy

Safety

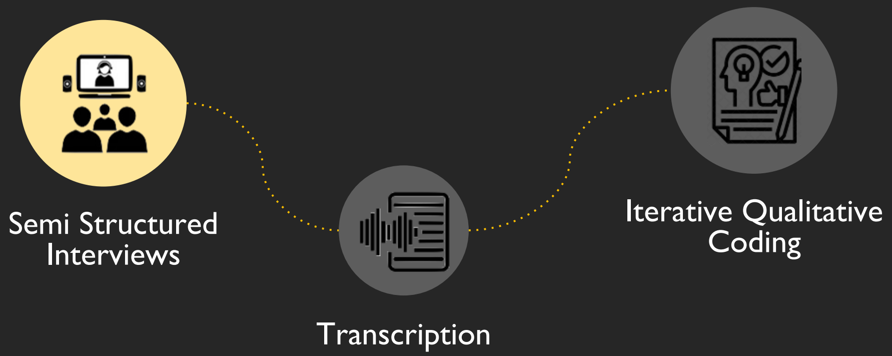
Health

Recruitment



| Role | Domains of Interest | | | | | Team/Organization Role | Org. Type |
|---|---|---|---|---|---|---|---|
| Professor | ● | ● | | | | Research | Academia |
| Professor | ● | ● | ● | ● | ● | Research | Academia |
| Researcher | ● | ● | ● | ● | ● | Research | Academia |
| Fact-checker | ● | ● | | | | Fact Checking | Industry |
| Researcher | ● | ● | ● | ● | ● | Social Network Analysis | Industry |
| AI-Tech Founder | ● | ● | | ● | ● | AI Technology Development | Industry |
| Intelligence Analyst | ● | ● | ● | ● | ● | Social Network Analysis | Industry |
| Fact-checker | ● | ● | ● | ● | ● | Journalism | Industry |
| Editor | ● | ● | ● | ● | ● | Journalism | Industry |
| Consultant | ● | ● | | | | Platform Trust & Safety | Industry |
| AI-Tech Founder | | | | | ● | AI Technology Development | Industry |
| Data Analyst | | | | ● | | Outsourced Trust & Safety | Industry |
| Intelligence Analyst | ● | ● | | ● | | Outsourced Trust & Safety | Industry |
| Data Scientist | ● | ● | | | ● | Platform Trust & Safety | Industry |
| Product Manager | | | | ● | | Outsourced Trust & Safety | Industry |
| Researcher | ● | ● | | | ● | Research; Advocacy | NGO |
| Researcher | ● | ● | | | ● | Research; Advocacy | NGO |
| Consultant | ● | ● | | | | Advocacy | NGO |
| Researcher | ● | ● | ● | ● | ● | Platform Trust & Safety | NGO |
| Researcher | ● | ● | | | | Think Tank | NGO |
| Fact-checker | ● | ● | | | ● | Fact Checking | Non-Profit |
| Researcher | ● | ● | ● | | | Advocacy; Research | Non-Profit |

# Constructing the Modeling Framework


Semi Structured Interviews

Transcription

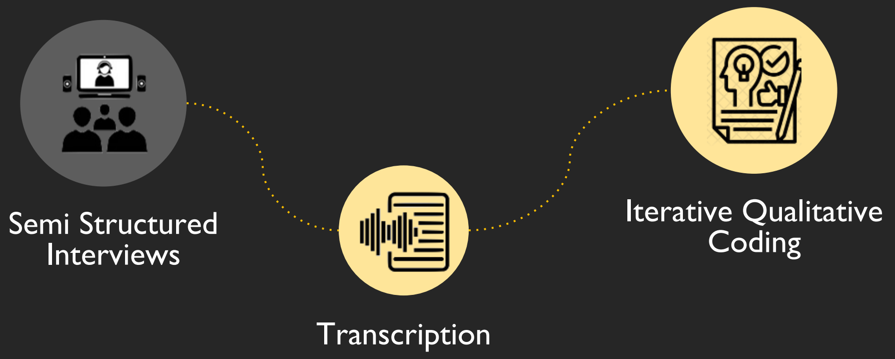
Iterative Qualitative Coding

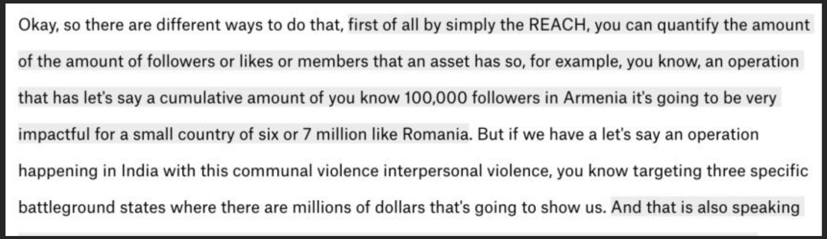
| Roles | # |
|---|---|
| Platform Trust & Safety Specialist | 3 |
| Outsourced Trust & Safety Specialist | 4 |
| Fact Checker | 3 |
| Journalist | 2 |
| Academic Researcher | 3 |
| AI-Tech Founder | 2 |
| Advocacy Researcher | 5 |

## Expert Interviews Slide Deck - Themes

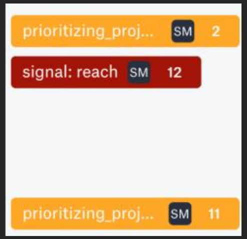
- Background: Role/Team/Organization
- Surfacing & Prioritizing Projects
- Assessing Projects
- Characterization of Threat Actors
- Challenges & Wishlist

# Constructing the Modeling Framework





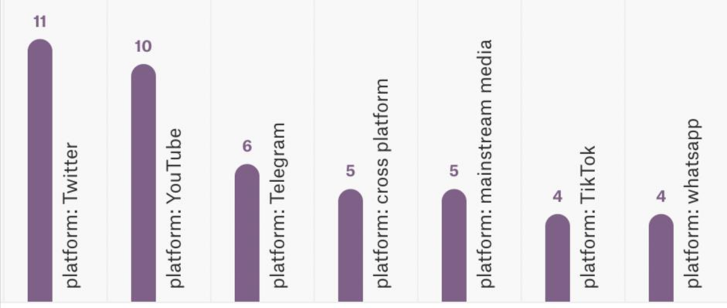
Okay, so there are different ways to do that, first of all by simply the REACH, you can quantify the amount of the amount of followers or likes or members that an asset has so, for example, you know, an operation that has let's say a cumulative amount of you know 100,000 followers in Armenia it's going to be very impactful for a small country of six or 7 million like Romania. But if we have a let's say an operation happening in India with this communal violence interpersonal violence, you know targeting three specific battleground states where there are millions of dollars that's going to show us. And that is also speaking



prioritizing_proj... SM 2

signal: reach SM 12

prioritizing_proj... SM 11



## Iterative Qualitative Coding Process

- Familiarization
- Open-coding
- Analytical memo writing
- Framework development
- Indexing

Shujaat Mirza, Labeeba Begum, Liang Niu, Sarah Pardo, Azza Abouzied, Paolo Papotti, Christina Pöpper

# Domains of Interest

We use participants’ areas of focus to identify five primary domains where the contest between mitigation teams and disinformation actors takes place

01

### National Security

Campaigns target international relations and conflicts between states, often supplementing traditional warfare

02

### Democracy

Campaigns target democratic processes such as elections, censuses, referenda, and ballot initiatives

03

### Economy

Campaigns target financial interests to disrupt market activity, or abuse the financial incentives of platforms to make a profit

04

### Public Safety

Campaigns aim to cause civil unrest or violence, often utilizing hate speech to target vulnerable groups

05

### Public Health

Campaigns diminish trust in science, leading to vaccine hesitancy & delay in the provision of health care during crisis events

# Threat Characterization Framework

In our framework, disinformation events or campaigns are characterized by the following four elements:

01

## Threat Actor

Who creates, spreads or amplifies disinformation?

02

## Attack Patterns

How do the actors effectively disinform?

03

## Attack Channels

On which platforms and media do the actors disinform?

04

## Target Audience

Who are the targets of the actors' attacks?

# Threat Characterization Framework

In our framework, disinformation events or campaigns are characterized by the following four elements:

01

### Threat Actor

Who creates, spreads or amplifies disinformation?

02

### Attack Patterns

How do the actors effectively disinform?

03

### Attack Channels

On which platforms and media do the actors disinform?

04

### Target Audience

Who are the targets of the actors' attacks?

# Attack Patterns

Flood: bots, cyborgs, copypasta

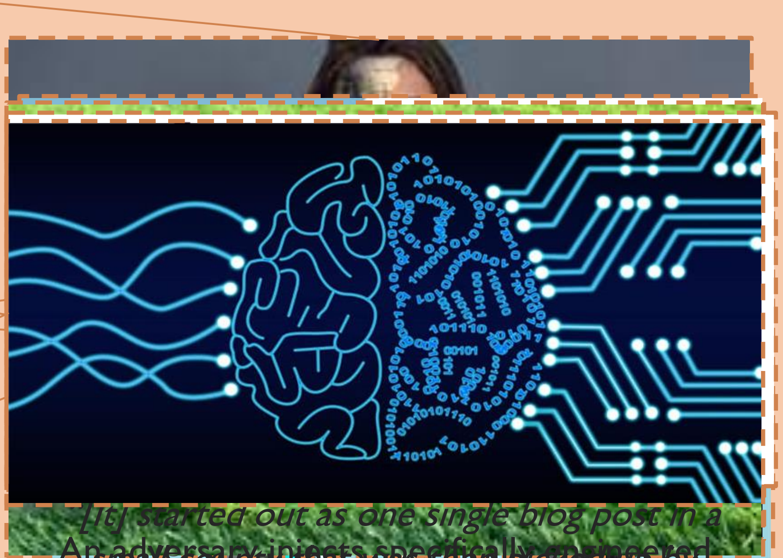
Drown: trolls, hijacking

Counterfeit: pseudo entities, astroturfing, pseudo content

Infiltrate: seed-invite-amplify, mainstream

Evade Detection: gaming heuristics, poisoning attacks, crowdsource

Evade Attribution: proxy companies, dark PR firms



*“[it] started out as one single blog post in a small county that was then picked up by* [illegible]

An adversary injects specifically engineered [illegible]

bot-assisted human, inheriting characteristics from both. It is [illegible] *like a bot, but then if someone responds to them, a person will* [illegible]

[illegible]

*corporation” - P15.*

as tobacco, energy and insurance companies.

*the goals of those communities.”*

# Modeling Fake Cluster Attacks Belgian Government

actor

pattern

channel

target

Insufficient evidence

GAN generated profiles

Dwire news site

Creating western personas & languages

Network of twitter bots

Invite Huawei engagement

Repurpose existing content

Traces of multi campaign infrastructure

Twitter

News domains

| | | |
|---|---|---|
| | | corporate |
| actor | corporate? | ideological |
| pattern | counterfeit :: pseudo entities | |
| | counterfeit :: pseudo content | |
| | counterfeit :: astroturfing | |
| | flood :: bots | |
| | infiltrate :: seed-invite-amplify | |
| | evade detection :: crowdsource | |
| | evade attribution :: PR firms | |
| channel | social media | |
| | web | |
| target | Belgian audiences | |

actor

pattern

channel

target

# External Validity

Gr

Fak
Hu

Accou
Over

Graph

01.20

15
o
si

+550 website domain name registered

+750 fake media in 116 countries

Imp
EU

Maxin
negati
about
advers

EU DisinfoLab

Gra

Face
VDA

Coordin
Boosted

05.2020

Graphika

Myanmar Military Network

Coordinated Inauthentic Behavior Traced to Members of Myanmar Military Before Elections

Ben Nimmo, Léa Ronzaud,
C. Shawn Eib, Rodrigo Ferreira

10.2020

Takedowns

# Utility and Anticipated Usage

### Standardized, structured analysis

Organize unstructured information into a compact form communicable to a diverse set of stakeholders

### Threat severity based ranking for triage

Inspired by CVSS, rank campaigns as a means of triage to guide mitigators work by prioritizing incidents by severity

### Tackling cross-platform campaigns

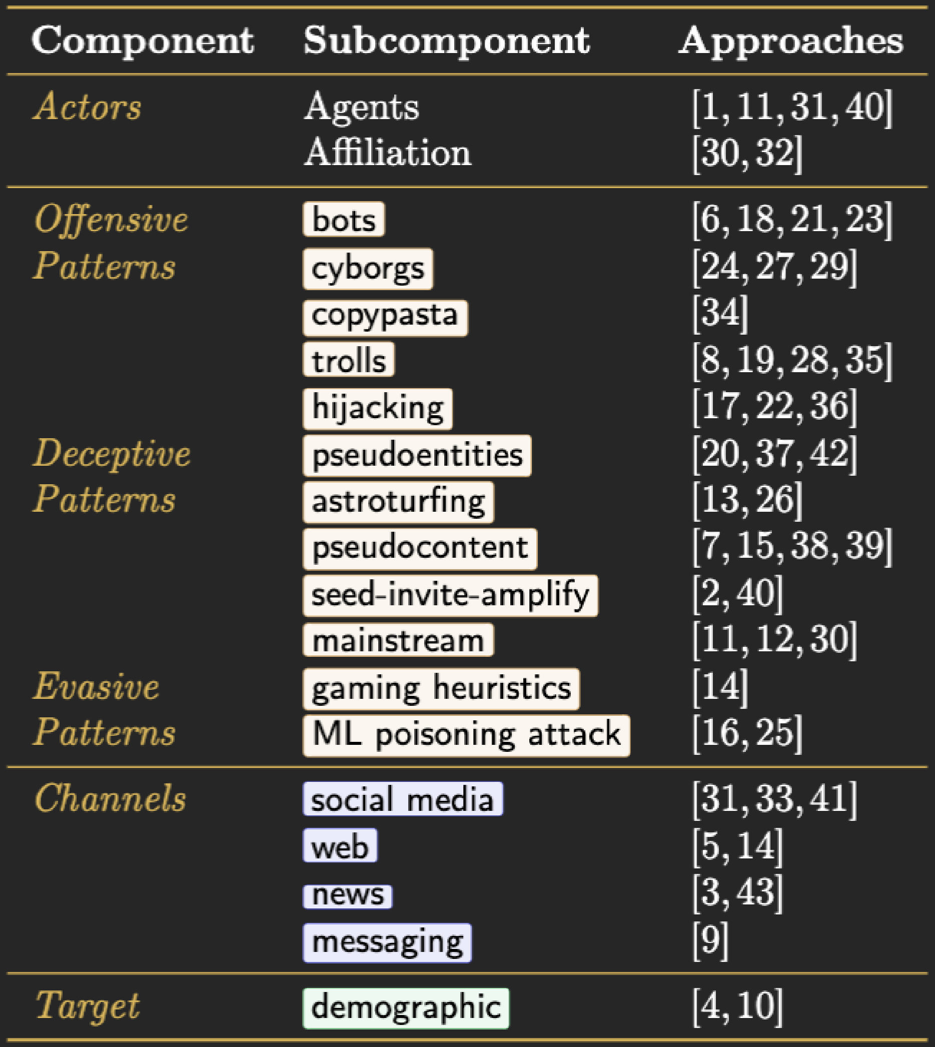
Take a broader view in their mitigation effort by capturing different channels involved in cross platform operations

### Bended patterns & tactics

Draws parallel to malware operations as tactics are used in combination to achieve desired goals of the operation

# Towards an Automated Procedure

- ➢ Harmful content can go viral faster than teams can intervene due to resource constraints
- ➢ To develop threat assessment and triage systems on top of our framework, automation will be essential to implement at large scale
- ➢ We identify framework components with potential for automation and related work on relevant methods



| Component | Subcomponent | Approaches |
|---|---|---|
| *Actors* | Agents | [1, 11, 31, 40] |
| | Affiliation | [30, 32] |
| *Offensive Patterns* | bots | [6, 18, 21, 23] |
| | cyborgs | [24, 27, 29] |
| | copypasta | [34] |
| | trolls | [8, 19, 28, 35] |
| | hijacking | [17, 22, 36] |
| *Deceptive Patterns* | pseudoentities | [20, 37, 42] |
| | astroturfing | [13, 26] |
| | pseudocontent | [7, 15, 38, 39] |
| | seed-invite-amplify | [2, 40] |
| | mainstream | [11, 12, 30] |
| *Evasive Patterns* | gaming heuristics | [14] |
| | ML poisoning attack | [16, 25] |
| *Channels* | social media | [31, 33, 41] |
| | web | [5, 14] |
| | news | [3, 43] |
| | messaging | [9] |
| *Target* | demographic | [4, 10] |

*“We are now increasingly seeing that [disinformation] is seen as a cyber threat, and certain approaches that we’ve been taking to tackle cybersecurity issues might be used for disinformation as well. We’re seeing quite a lot of overlap starting to emerge between these two areas”* — *One of the participants*

**Direct your questions at**
**shujaat.mirza@nyu.edu**

Shujaat Mirza

Labeeba Begum

Liang Niu

Sarah Pardo

Azza Abouzied

Paolo Papotti

Christina Pöpper

CITIES

CENTER FOR
CYBER SECURITY

EURECOM

جامعة نيويورك أبوظبي
NYU | ABU DHABI