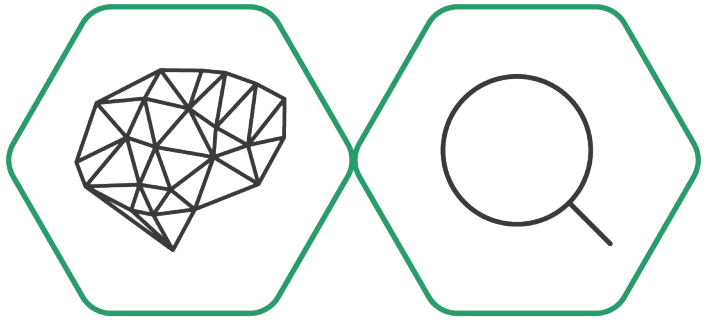
Detecting Coordinated Link Sharing Behaviour with CooRnet

Fabio, Giglietto (University of Urbino), Roberto Mincigrucci (University of Urbino), Giada Marino (University of Sassari)

Amsterdam, January 9 2023

Tutorial at the Digital Media Initiative Winter School



vera.ai

Summary

1. Definitions: from CiB to CLSB;
2. Getting started: the Global CLSB List;
3. Create CrowdTangle lists with groups and Pages from the Global CLSB List;
4. Collect all the posts created by these accounts from 2022-02-20 2022-12-03 with a set of Ukraine invasion keywords;
5. From the CrowdTangle CSV, extract all the links in the posts;
6. Used this list of URLs to retrieve all the Facebook/Instagram shares of these URLs (all posts that shared these URLs);
7. Detected coordinated accounts and clusters by analyzing the dataset of shares;
8. Output extraction and exploration;
9. CLSB map of Ukraine Invasion.

Definitions: from CiB to CLSB

- Coordinated Inauthentic Behavior (CiB) is described by Meta “as the use of multiple Facebook or Instagram assets, working in concert to engage in Inauthentic Behavior, where the use of fake accounts is central to the operation”
- Coordinated Link Sharing Behaviour (CLSB) is a form of CiB performed by a network of social media accounts that **repeatedly** share the same link in a **very short period** of time from each other
- CooRnet is an R package that detected CLSB

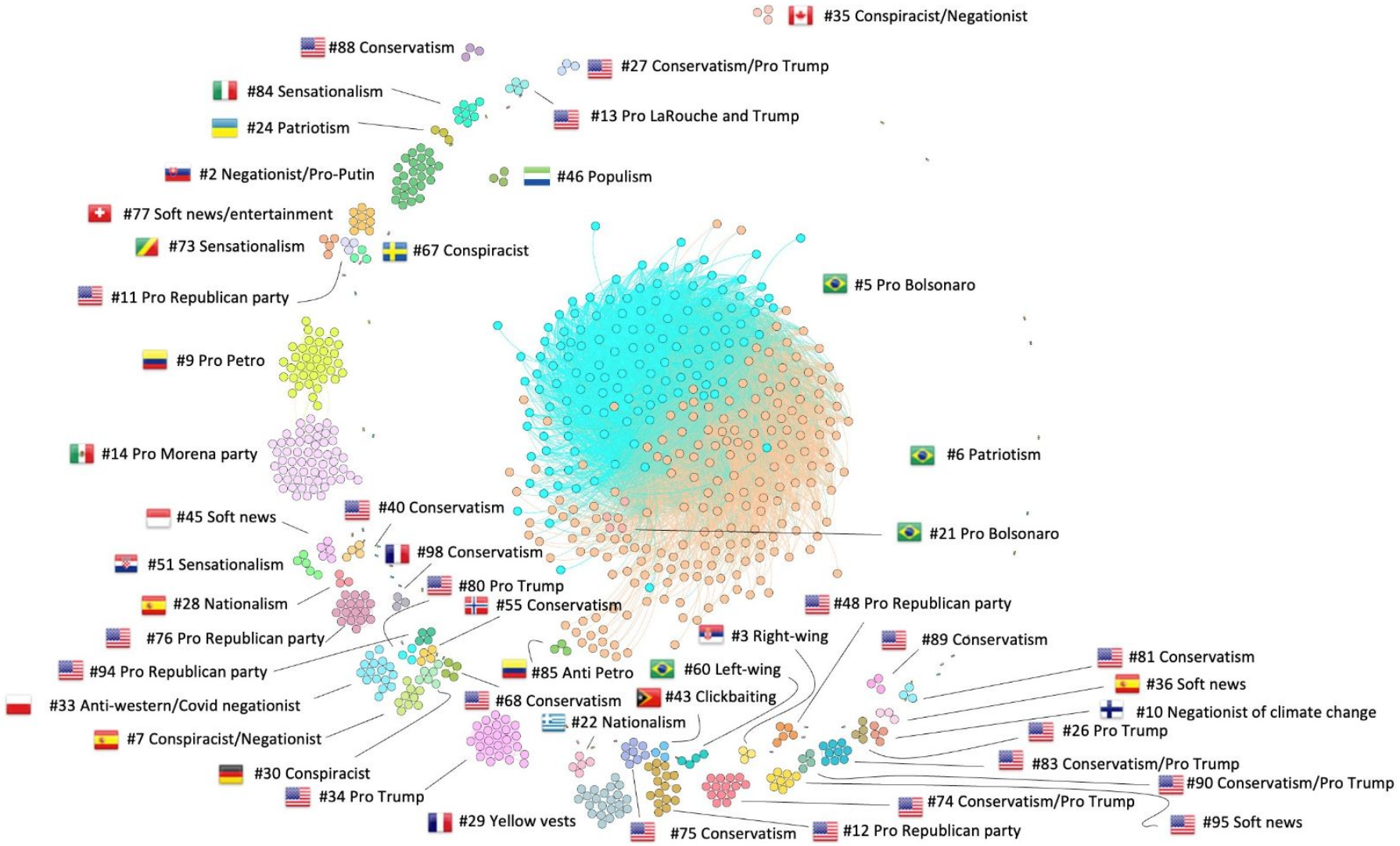
Tip

A cycle of CLSB detection starts from a set of links. For this tutorial we collected links from posts shared by a list of already known coordinated actors. This process is called “CooRnet iteration”

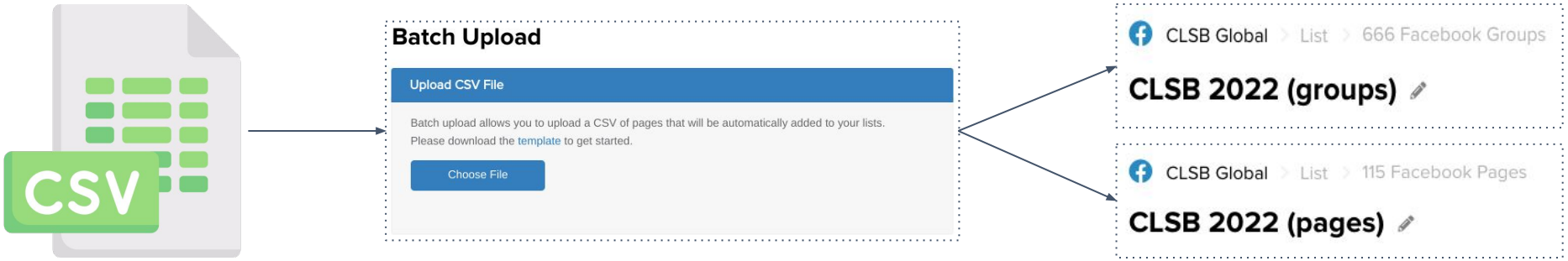
Getting Started: the Global CLSB List (1/2)

- For the Global CLSB List we started from 25,870 news stories rated as problematic (“false”, "missing context", "mixture or false headline" or “missing context”) by Facebook’s third-party fact-checkers from January 2017 to December 2021
- Using CooRnet, we detected 818 CLSB accounts that shared at least four different problematic news stories

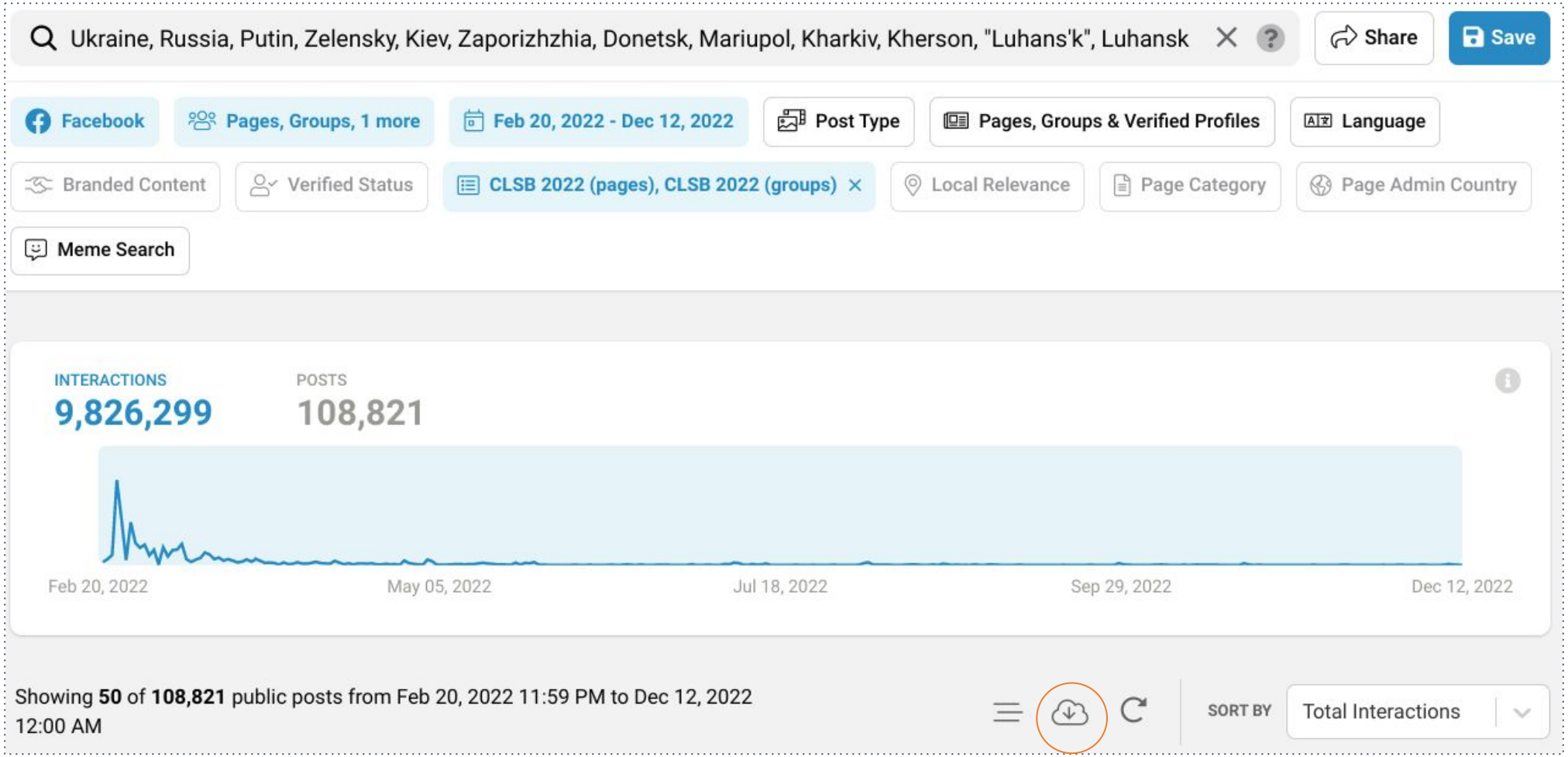
Getting Started: the Global CLSB List (2/2)



CrowdTangle lists



Collecting posts



Extracting links

112,119 posts

```
allpostsfile <-
"https://github.com/fabiogiglietto/CooRnet_at_DMI_WS_2023/blob/main/rawdata/allpost
s.csv?raw=true"
```

```
urls <- CooRnet::get_urls_from_ct_histdata(ct_histdata_csv = allpostsfile,
                                           newformat = TRUE)
```

31,578 unique URLs

Retrieving Facebook posts that shared our links

```
ct_shares.urls <- CooRnet::get_ctshares(urls,
                                        sleep_time = 1,
                                        get_history = FALSE,
                                        clean_urls = TRUE)
```


1,128,826 posts

Tip

This process takes a long time. For sake of time, please download the ct_shares rds file from the tutorial repository on GitHub

For your future projects, you may want to request an increase of the standard rate limit associated with your CrowdTangle token

Detecting and marking coordinated shares

```
CooRnet::estimate_coord_interval(ct_shares.df = ct_shares.urls) # 23 secs
```



```
output <- CooRnet::get_coord_shares(ct_shares.df = ct_shares.urls,
                                    coordination_interval = "23 secs",
                                    parallel = FALSE,
                                    percentile_edge_weight = 0.995,
                                    keep_ourl_only = TRUE,
                                    clean_urls = TRUE)
```

Output extraction and exploration (1/2)

```
CooRnet::get_outputs(coord_shares_output = output,
                     component_summary = TRUE,
                     cluster_summary = TRUE,
                     top_coord_share = TRUE,
                     top_coord_urls = TRUE,)
```

highly_connected_coordinated_entities

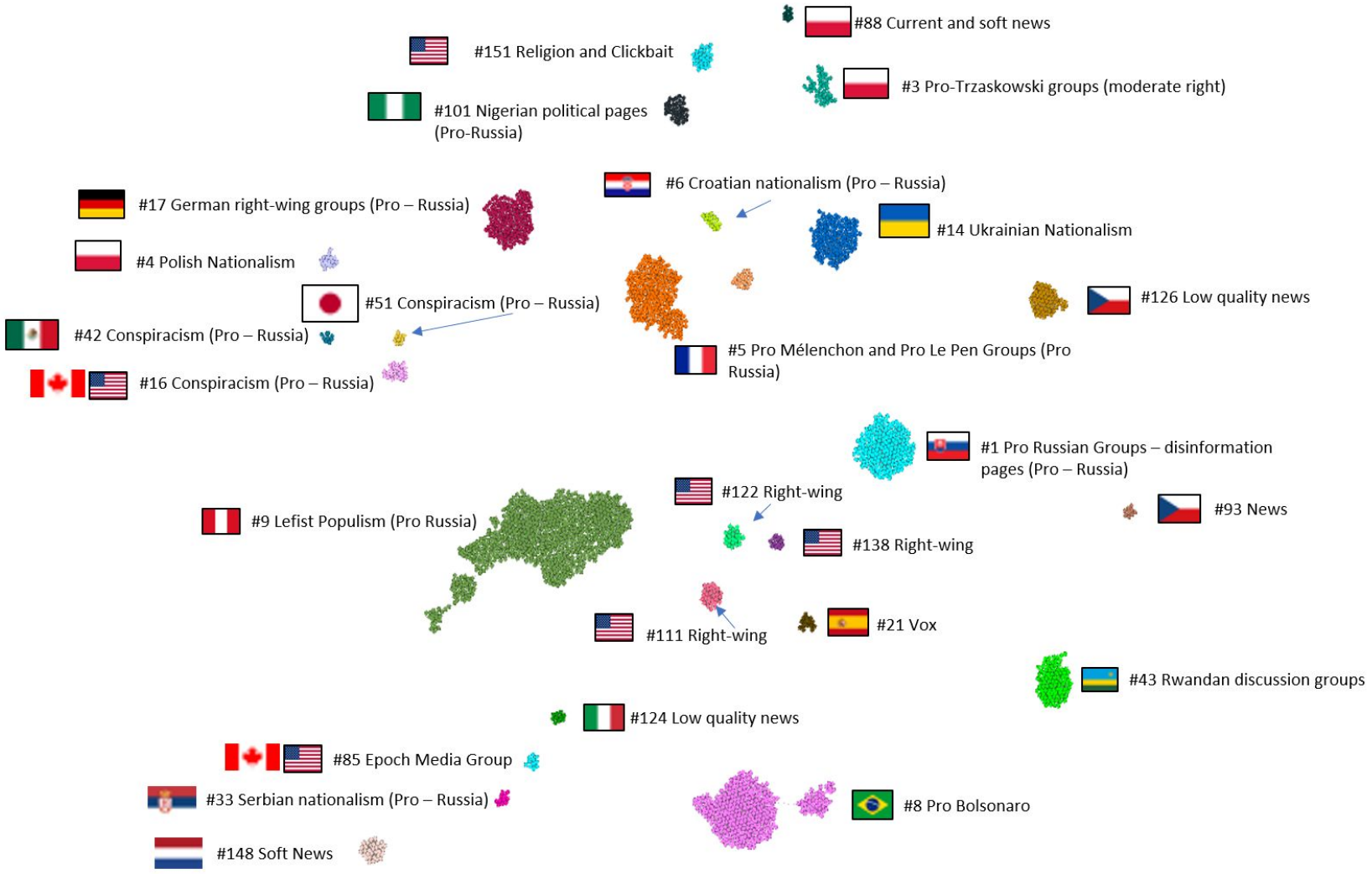
cluster_summary

highly_connected_g.graphml

Output extraction and exploration (2/2)

1. Highly_connected_coordinated_entities: lists and describe all the accounts detected as coordinated (2,501)
2. Cluster_summary: lists and describes all coordinated networks (161)
3. Highly_connected_g.graphml: file to be imported in Gephi to visually explore the map

CLSB map of Ukraine Invasion



Outline of the work sessions

1. Explore the outputs to identify a network of your interest;
2. Join a group based on your interest in the network;
3. Take a look to the following examples
   https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3743531
   https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3775469
4. Analyze the network (using social media analysis and OSINT techniques) from the actors (A), behaviors (B) and content (C) perspectives
5. Prepare a presentation (and eventually a report) to describe your work

Fabio Giglietto / University of Urbino

Contact: https://scholar.social/@fabiogiglietto

vera.ai

CERTH
CENTRE FOR RESEARCH & TECHNOLOGY HELLAS

The University Of Sheffield.

ATC
ATHENS TECHNOLOGY CENTER

KINIT

ontotext

Deutsche Welle

Università degli Studi di Napoli Federico II

1506 Università degli Studi di Urbino Carlo Bo

AFP

Universiteit van Amsterdam

école normale supérieure paris—saclay

université PARIS-SACLAY

EBU
OPERATING EUROVISION AND EURORADIO

EU DisinfoLab

Fraunhofer IDMT

Follow us on Twitter: @veraai_eu
Website: https://www.veraai.eu/

Co-financed by the European Union, Horizon Europe programme, Grant Agreement No 101070093.

Additional funding from Innovate UK grant No 10039055 and the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract No 22.00245

Work session 1

1. Individually familiarize and explore CooRnet outputs.
2. Start creating the coordinated accounts map.

Work session 2

1. Work on the coordinated accounts map.
2. Classify the various networks identified.

Work session 3

1. Finalize and edit the coordinated accounts map.
2. Identify one or more network that deserves a more in depth look.

Work session 3

Networks of interest:

1.

2.

3.

…

Work session 4

1. Groups formation;
2. Analysis of the expected outputs.

Work session 4

Groups:

1.

2.

3.

…

Work session 5

1. Examples of research work with CooRnet data

   https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3743531
   https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3775469

2. Setting the analysis based on the actors, behaviors, content approach.

Work session 6 and 7

Case description: Exploration and analysis of **actors** in the network. Identification of malicious actors behind the networks, if any. Answering, for example, to the question: Is it this network related to a specific or multiple news sources? Is this/these news source/s problematic or untrustworthy?

Work session 8 and 9

Case description: Exploration and analysis of the network sharing **behaviors**. Identification of patterns and overlapping trends, if any. For example, type of posts published or links posted in comments, etc.

Work session 10 and 11

Case description: Exploration and analysis of **content** published by pages and groups in the network with examples of posts (screenshots)

Work session 12 and 13

Case description: Observation and analysis, if possible, of networks' engagement patterns.

Work session 14 to 16

Work sessions focused on final presentation, final report writing and conclusion.