

Rapid Response Networks and Frameworks

Julian Neylan DISARM Foundation

EU CYBERNET SUMMER SCHOOL

**Cyber Crisis Management:
Navigating Disinformation and
Cyber Attacks in the AI Era**



Federal Foreign Office



Funded by
the European Union

Key Actors

G7 Rapid Response Group

European External Action Service

Military (NATO, various MODs)

National institutions (Viginum, NASK, FCDO, MFAs etc.)

Civil Society (FIMI-ISAC, Fact-Checkers, NGOs etc.)

EU CYBERNET SUMMER SCHOOL

**Cyber Crisis Management:
Navigating Disinformation and
Cyber Attacks in the AI Era**



Federal Foreign Office



Funded by
the European Union

These Groups first detect an incident or campaign

After detection, analysis of the data is required

EU CYBERNET SUMMER SCHOOL

**Cyber Crisis Management:
Navigating Disinformation and
Cyber Attacks in the AI Era**



Federal Foreign Office



Funded by
the European Union



ABCDE Model

Actor Assets Account (who is it)

Behavior Actions by campaign (what are they doing)

Content Narrative Media (what are they saying)

Distribution Channel Target Reach (how far does their message go?)

Effect Engagement Impact (are there consequences?)

https://www.ivir.nl/publicaties/download/ABC_Framework_2019_Sept_2019.pdf

<https://www.brookings.edu/articles/adding-a-d-to-the-abc-disinformation-framework/>

[https://carnegie-production-assets.s3.amazonaws.com/static/files/Pamment - Crafting Disinformation 1.pdf](https://carnegie-production-assets.s3.amazonaws.com/static/files/Pamment_-_Crafting_Disinformation_1.pdf)



A-Attribution

Content Signals	Behavioral Signals	Technical Signals	Financial Signals
Text, images or video data contained within messages and/or websites are the focus of analysis	Rather than attending to what is communicated, behavioural analysis is concerned with how the distorting, distorting or deceptive activities are organized and conducted.	The focus here is upon the technological infrastructure used to mount an influence operation, in terms of IP addresses, particular servers and so forth	The financial data in making open-source attribution judgements. 19 'Following the money'

<https://adacio.eu/wp-content/uploads/2025/02/ADACio-D1.2-%E2%80%93-Open-Source-Investigations-Field-Guide.pdf>

B- DISARM Framework

Tactics

<https://www.disarm.foundation/framework>

Techniques

Plan Strategy 2 techniques	Plan Objectives 13 techniques	Target Audience Analysis 3 techniques	Develop Narratives 7 techniques	Develop Content 8 techniques	Establish Assets 16 techniques	Establish Legitimacy 5 techniques	Microtarget 4 techniques	Select Channels and Affordances 10 techniques	Conduct Pump Priming 5 techniques	Deliver Content 4 techniques
Determine Strategic Ends (0/4)	Cause Harm (0/3)	Identify Social and Technical Vulnerabilities (0/8)	Demand Insurmountable Proof	Create Hashtags and Search Artefacts (0/2)	Account Asset (0/7)	Co-Opt Trusted Sources (0/3)	Create Clickbait	Bookmarking and Content Curation	Seed Distortions	Attract Traditional Media
Determine Target Audiences	Cultivate Support (0/8)	Map Target Audience Information Environment (0/5)	Develop Competing Narratives	Develop Audio-Based Content (0/2)	Acquire/Recruit Network (0/2)	Establish Inauthentic News Sites (0/2)	Create Localised Content	Consumer Review Networks	Seed Kernel of Truth	Comment or Reply on Content (0/1)
	Degrade Adversary	Segment Audiences (0/5)	Develop New Narratives	Develop Image-Based Content (0/4)	Asset Origin (0/8)	Persona Legitimacy (0/4)	Leverage Echo Chambers/Filter Bubbles (0/3)	Digital Community Hosting Asset (0/17)	Trial Content	Deliver Ads (0/2)
	Dismay		Integrate Target Audience Vulnerabilities into Narrative	Develop Text-Based Content (0/7)	Build Network (0/3)	Persona Legitimacy Evidence (0/2)	Purchase Targeted Advertisements	Digital Content Creation Asset (0/2)	Use Fake Experts	Post Content (0/3)
	Dismiss (0/1)		Leverage Conspiracy Theory Narratives (0/2)	Develop Video-Based Content (0/2)	Cultivate Ignorant Agents	Present Persona (0/22)		Digital Content Delivery Asset (0/7)	Use Search Engine Optimisation	
	Dissuade from Acting (0/3)		Leverage Existing Narratives	Distort Facts (0/2)	Develop Owned Media Assets			Digital Content Hosting Asset (0/12)		
	Distort		Respond to Breaking News Event or Active Crisis	Obtain Private Documents (0/2)	Employ Commercial Analytic Firms			Formal Diplomatic Channels		
	Distract			Reuse Existing Content (0/4)	Establish Account Imagery (0/7)					
	Divide				Financial Instrument (0/9)					
	Facilitate State Propaganda				Infiltrate Existing Networks (0/2)					
	Make Money (0/6)				Leverage Content Farms (0/2)					
	Motivate to Act (0/3)									
	Undermine (0/4)									

EU CYBERNET COMMUNITY

Cyber Crisis Management:
Navigating Disinformation and
Cyber Attacks in the AI Era



Federal Foreign Office



Funded by
the European Union

Example from Recorded Future

Recorded Future Diamond Model for Influence Operations	
Influencer	Maduro Regime Venezuelan State-Sponsored Media Iranian State-Owned Media Iranian Proxy Media Organizations
Audience	International Government Officials Venezuelan Citizens Political Activists
Capabilities	T0002: Facilitate State Propaganda T0018: Purchase Targeted Advertisements T0101: Create Localized Content T0023: Distort Facts T0086: Develop Image-Based Content T0085.003: Develop Inauthentic News Articles T0084: Reuse Existing Content T0087: Develop Video-Based Content T0117: Attract Traditional Media T0126: Encourage Attendance at Events T0060: Continue to Amplify T0131.001: Legacy Web Content
Infrastructure	Venezuelan and Iranian Governments' Official Communications Graffiti Billboards/Posters teleSUR HispanTV Al Mayadeen



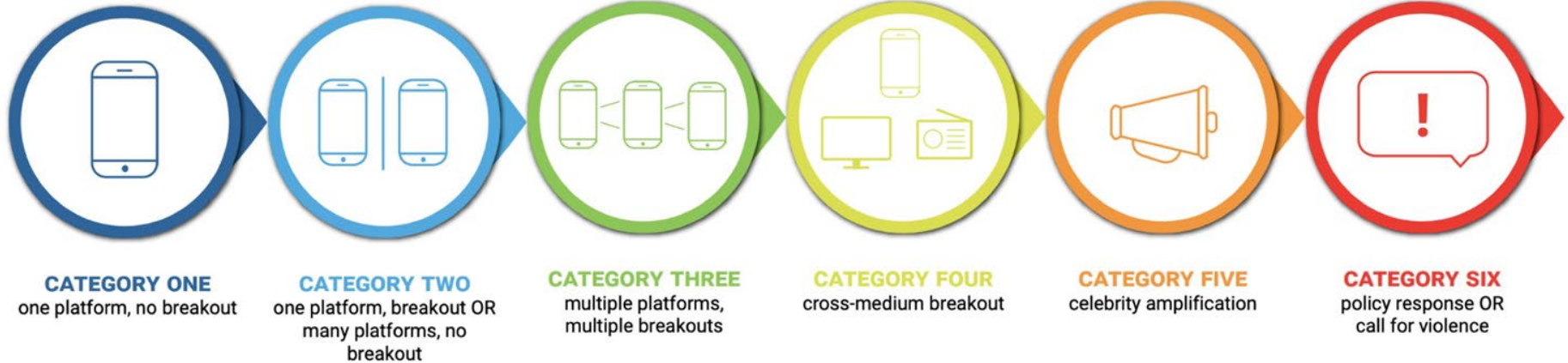
C- Content- INFOTESTER

INFOTESTER

1. **Cherry Picking** — presenting information using only data supporting a given thesis, while ignoring the wider context.
2. **Quote Mining** — using a short fragment of someone's longer speech in a way that significantly distorts its real, original tone.
3. **Anecdote** — the use of evidence in the form of personal experience or an isolated case, possibly rumour or hearsay, most often to discredit statistics.
4. **Whataboutism** — responding to a substantive argument not by addressing the heart of the matter, but by raising a new point that is unrelated to the topic under discussion. Often referred to as tossing a false lead to distract attention from the topic (*Red Herring*). Technique typical of Russian propaganda.

D- Breakout Scale

THE BREAKOUT SCALE

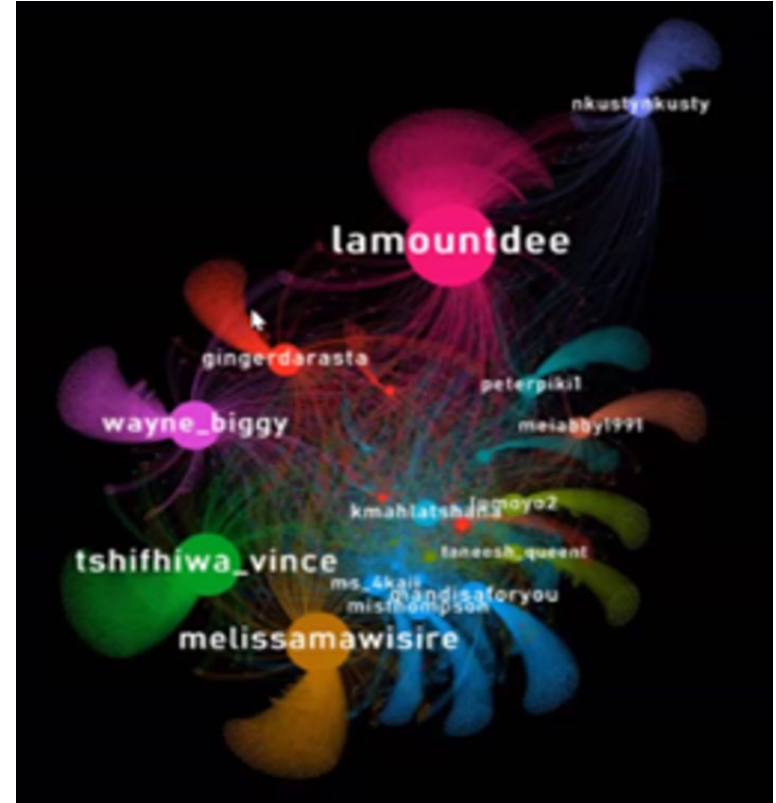


<https://www.brookings.edu/articles/the-breakout-scale-measuring-the-impact-of-influence-operations/>



Social Media Diffusion

- **Originators**
 - accounts that originate the narrative
- **Amplifiers**
 - accounts that amplify the narrative
- **True Believers**
 - legitimate users who buy into the narrative and amplify it
- **Opportunists**
 - accounts that hijack the narrative and amplify it for a different agenda



Example: Detecting Anomalous Engagement

- C = Coefficient of Traffic Manipulation
- $C = R/10 + F + U$
- R = % Retweets
- F = % traffic from top 50 users
- U = average posts per user

Phrase	R/10	F	U	Coefficient
Friday	4.966	1.48	1.22	7.666
#4thofJuly	5.834	1.78	1.17	8.784
Thursday	6.396	1.63	1.21	9.236
Covfefe	5.561	2.27	2.14	9.971
Wednesday	6.931	2.50	1.24	10.671
Davos 1	6.617	2.85	1.54	11.007
Davos 2	7.512	2.82	1.59	11.922
Tamim the Glorious	7.74	2.43	3.65	13.82
Qadhafi of the Gulf	9.495	4.11	2.94	16.545
#DigDoug	6.678	19.78	4.78	31.238
#LaFranceVoteMarine	8.484	24.23	5.31	38.024
#LePionMacron	8.928	27.82	5.94	42.688
#StopAstroturfing	8.727	34.99	6.61	50.327
#Marine2017	8.705	35.83	8.81	53.345

<https://demtech.oii.ox.ac.uk/research/posts/measuring-traffic-manipulation-on-twitter/>

E- Effect



Courtesy of Daniel J. Solove

- *Deception*: Did someone act upon deceptive content? how do you know? surveys? polls? focus groups?
- *Freedom of Speech*: is there censorship? who is being censored? is there a chilling effect? whose speech is suppressed?
- *Harm*: Is there a chance of real-world harm? to whom? how extensive? how would you know? user reports? news reports?

Frame Harms Through a Human Rights Lens



Threats to personal
and community safety



Violation of dignity



Harm to health and
well-being



Invasion of privacy



Hate and discrimination



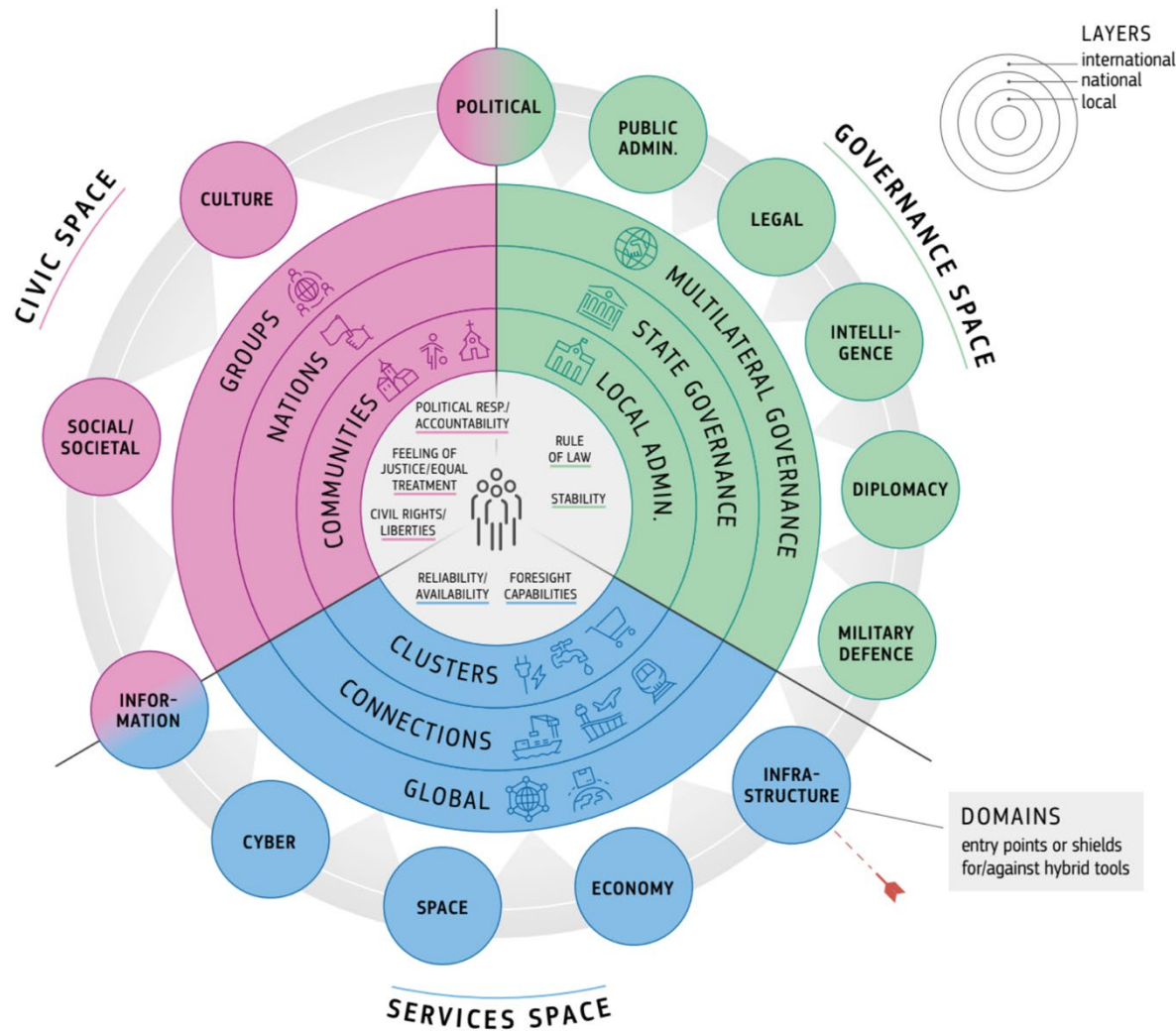
Deception and
manipulation



World Economic Forum Online Harms Typology

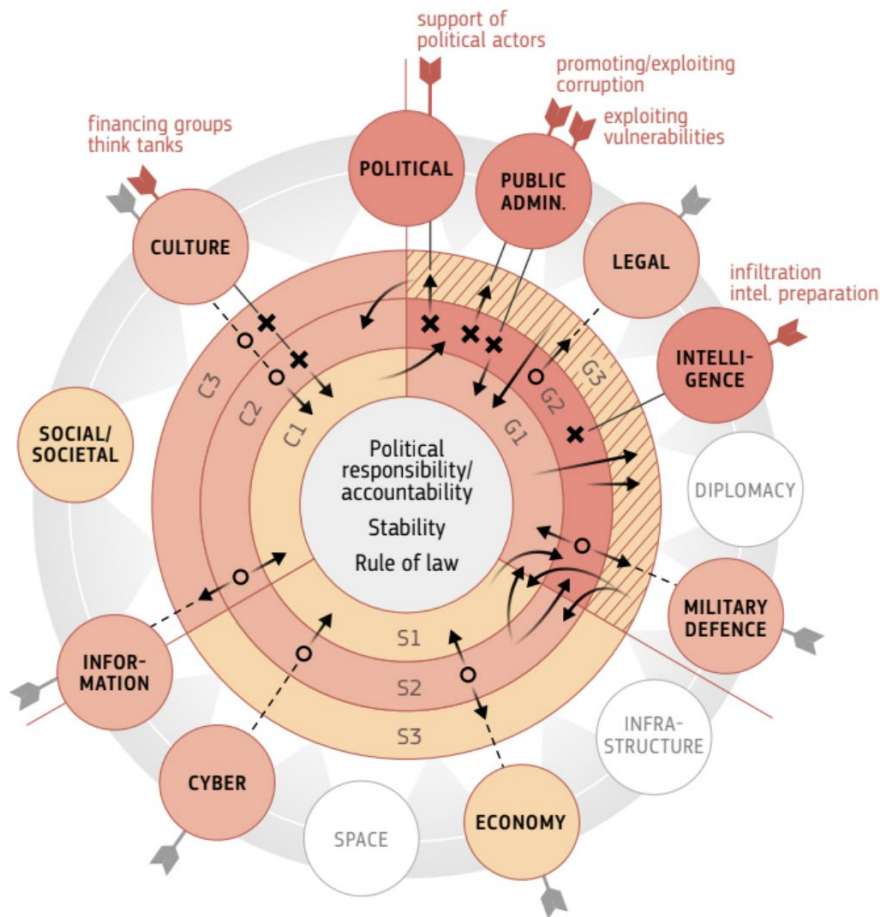
Other useful frameworks for contextualizing incidents





Source

RUSSIA IN WESTERN BALKANS (INFLUENCE)



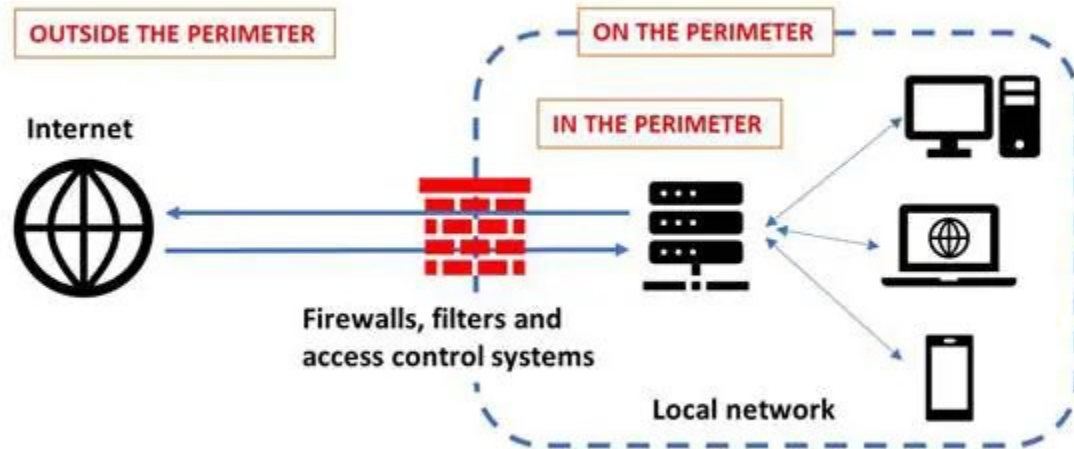
2018–2021

- **Financing cultural groups and think tanks** to spread Russophile views and culture.
C2/C3→C1→G2 Influenced cultural sensibilities at local level (C1) and societies and public administration at national level (G2).
- **Promoting and exploiting corruption** aimed to undermine the development.
G2→G3→S1/S2 It had a negative impact on the entire region (G3), harming local and national economic development (S1/S2).
- **Exploiting vulnerabilities in public administration** to undermine its efficiency.
G2→G1→G3 Poor governance led to discontent at the local level (G1), which affected the entire region (G3).
- **Support of political actors** who represent pro-Russia views.
G2→G3→C2 Affected Russia's influence in the region (G3) and polarised societies (C2).
- **Infiltration and intelligence preparation** by Russian services.
G2→G3 Covert expansion of military influence affected the entire region (G3).

Source



UNIDIR



<https://unidir.org/introducing-a-new-framework-to-analyze-ict-activities/>



Diamond Model for Influence

