



# **An exposition on THE forEign information mAnipulation and interference**

**Tactics, Techniques, and Procedures (TTPs)**

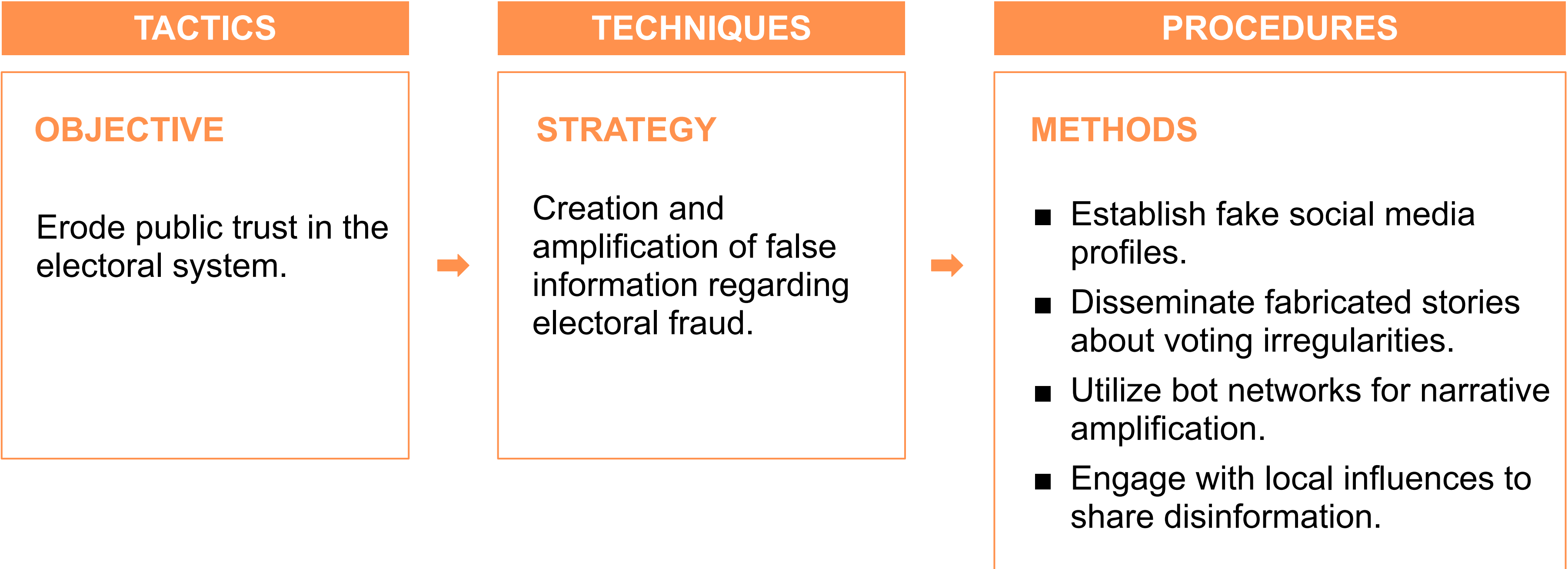
ATHENA Stakeholder Board Meeting

19 March 2025

Marios Dikaiakos (University of Cyprus)

# Influence Operation Scenario

A foreign entity aims to disrupt the democratic process of country X.



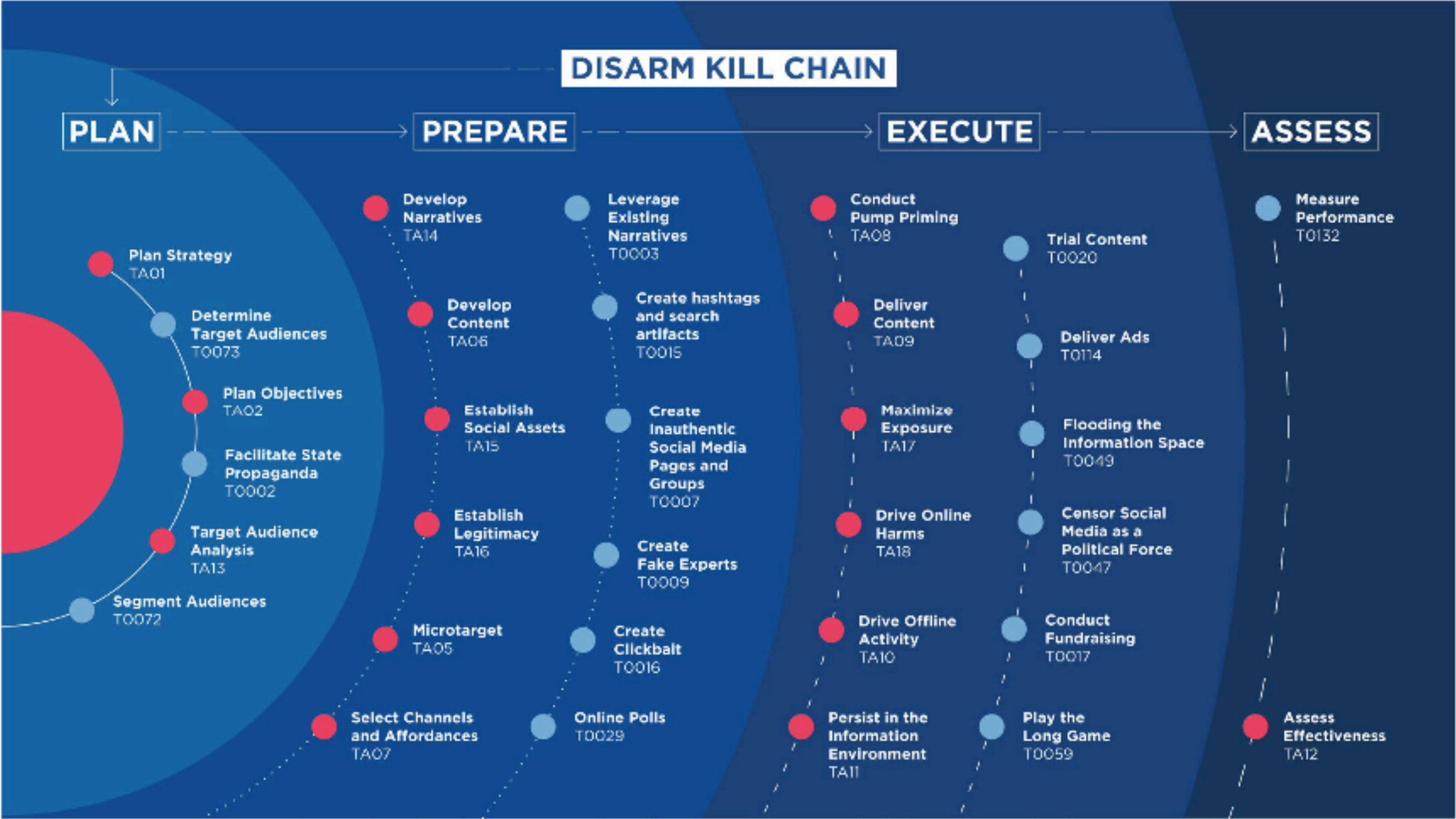
# Tactics, Techniques, and Procedures (TTPs)



- ❑ **Tactics:** High-level objectives that threat actors aim to achieve, such as **undermining democratic institutions**, **creating social divisions**, or **manipulating public opinion**.
- ❑ **Techniques:** Specific methods employed to accomplish these objectives, including **coordinating inauthentic behavior on social media**, **manipulating search engine algorithms**, or **using bots to disseminate disinformation**.
- ❑ **Procedures:** Detailed processes **combining various techniques across multiple tactics**, tailored to specific campaigns to **maximize impact** and **evade detection**.



# DISARM Framework



**Kill chain:** a framework that comprises the sequential stages involved in orchestrating and executing disinformation campaigns



Funded by the European Union (grant number 101132686). UK participants in Horizon Europe Project ATHENA are supported by UKRI grant number 10107667 (Trilateral Research). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA) or UKRI. Neither the European Union nor the granting authority nor UKRI can be held responsible for them. Grant Agreement 101132686 ATHENA HORIZON-CL2-2023-DEMOCRACY-01.



# DISARM Taxonomy



Plan			Prepare						Execute						Assess
Plan Strategy	Plan Objectives	Target Audience Analysis	Develop Narratives	Develop Content	Establish Social Assets	Establish Legitimacy	Microtarget	Select Channels and Affordances	Conduct Pump Priming	Deliver Content	Maximize Exposure	Drive Online Harms	Drive Offline Activity	Persist in the Information Environment	Assess Effectiveness
Determine Target Audiences	Facilitate State Propaganda	Segment Audiences	Leverage Existing Narratives	Create hashtags and search artifacts	Create Inauthentic Social Media Pages and Groups	Create fake experts	Create Clickbait	Online polls	Trial content	Deliver Ads	Flooding the Information Space	Censor social media as a political force	Conduct fundraising	Play the long game	Measure Performance
Determine Strategic Ends	Degrade Adversary	Geographic Segmentation	Develop Competing Narratives	Generate information pollution	Cultivate ignorant agents	Utilize Academic/Pseudoscientific Justifications	Purchase Targeted Advertisements	Chat apps	T0039 : Bait legitimate influencers	Social media	Trolls amplify and manipulate	Harass	Conduct Crowdfunding Campaigns	Continue to Amplify	People Focused
	Dismiss	Demographic Segmentation	Leverage Conspiracy Theory Narratives	Create fake research	Create inauthentic websites	Compromise legitimate accounts	Create Localized Content	Use Encrypted Chat Apps	Seed Kernel of truth	Traditional Media	Hijack existing hashtag	Boycott/"Cancel" Opponents	Organize Events	Conceal People	Content Focused
	Discredit Credible Sources	Economic Segmentation	Amplify Existing Conspiracy Theory Narratives	Hijack Hashtags	Prepare fundraising campaigns	Create personas	Leverage Echo Chambers/Filter Bubbles	Use Unencrypted Chats Apps	Seed distortions	Post Content	Bots Amplify via Automated Forwarding and Reposting	Harass People Based on Identities	Pay for Physical Action	Use Pseudonyms	View Focused
	Distort	Psychographic Segmentation	Develop Original Conspiracy Theory Narratives	Distort facts	Raise funds from malign actors	Backstop personas	Use existing Echo Chambers/Filter Bubbles	Livestream	Use fake experts	Share Memes	Utilize Spameflauge	Threaten to Dox	Conduct Symbolic Action	Conceal Network Identity	Measure Effectiveness
	Distract	Political Segmentation	Demand insurmountable proof	Reframe Context	Raise funds from ignorant agents	Establish Inauthentic News Sites	Create Echo Chambers/Filter Bubbles	Video Livestream	Use Search Engine Optimization	Post Violative Content to Provoke Takedown and Backlash	Conduct Swarming	Dox	Sell Merchandise	Distance Reputable Individuals from Operation	Behavior changes
	Dismay	Map Target Audience Information Environment	Respond to Breaking News Event or Active Crisis	Edit Open-Source Content	Prepare Physical Broadcast Capabilities	Create Inauthentic News Sites	Exploit Data Voids	Audio Livestream	Employ Commercial Analytic Firms	One-Way Direct Posting	Conduct Keyword Squatting	Control Information Environment through Offensive Cyberspace Operations	Sell Merchandise	Launder Accounts	Content
	Divide	Monitor Social Media Analytics	Develop New Narratives	Reuse Existing Content	Create Inauthentic Accounts	Leverage Existing Inauthentic News Sites		Social Networks		Comment or Reply on Content	Inauthentic Sites Amplify News and Narratives	Delete Opposing Content	Encourage Attendance at Events	Change Names of Accounts	Awareness
		Evaluate Media Surveys	Integrate Target Audience Vulnerabilities into Narrative	Use Copy-paste	Create Anonymous Accounts	Prepare Assets Impersonating Legitimate Entities		Mainstream Social Networks		Post inauthentic social media comment	Amplify Existing Narrative	Block Content	Call to action to attend	Conceal Operational Activity	Knowledge
		etc			etc	etc		etc			etc	etc	etc	etc	etc



# DISARM Taxonomy



Plan Strategy 2 techniques	Plan Objectives 13 techniques	Target Audience Analysis 3 techniques	Develop Narratives 7 techniques	Develop Content 8 techniques	Establish Assets 14 techniques	Establish Legitimacy 6 techniques	Microtarget 4 techniques	Select Channels and Affordances 12 techniques	Conduct Pump Priming 5 techniques	Deliver Content 4 techniques	Maximise Exposure 7 techniques	Drive Online Harms 5 techniques	Drive Offline Activity 5 techniques	Persist in the Information Environment 6 techniques	Assess Effectiveness 3 techniques
Determine Strategic Ends (1/4)	Cause Harm (2/3)	Identify Social and Technical Vulnerabilities (2/4)	Demand Insurmountable Proof	Create Hashtags and Search Artefacts	Acquire Compromised Asset (2/2)	Co-Opt Trusted Sources (2/3)	Create Clickbait	Blogging and Publishing Networks	Seed Distortions	Attract Traditional Media	Amplify Existing Narrative	Censor Social Media as a Political Force	Conduct Fundraising (2/3)	Conceal Information Assets (2/3)	Measure Effectiveness (2/3)
Determine Target Audiences	Cultivate Support (4/6)	Map Target Audience Information Environment (2/3)	Develop Competing Narratives	Develop Audio-Based Content (2/2)	Acquire/Recruit Network (2/2)	Create Fake Experts (2/1)	Create Localised Content	Bookmarking and Content Curation	Seed Kernel of Truth	Comment or Reply on Content (2/1)	Bait Influencer	Control Information Environment through Offensive Cyberspace Operations (2/4)	Encourage Attendance at Events (2/2)	Conceal Infrastructure (2/3)	Measure Effectiveness Indicators (or KPIs) (2/2)
	Degrade Adversary	Segment Audiences (1/3)	Develop New Narratives	Develop Image-Based Content (1/4)	Build Network (1/3)	Create Personas (2/1)	Leverage Echo Chambers/Filter Bubbles (1/3)	Chat Apps (2/2)	Trial Content	Deliver Ads (2/2)	Cross-Posting (2/3)	Direct Users to Alternative Platforms	Organise Events (2/2)	Conceal Operational Activity (2/2)	Measure Performance (2/3)
	Dismay		Integrate Target Audience Vulnerabilities into Narrative	Develop Text-Based Content (1/4)	Create Inauthentic Accounts (2/4)	Establish Inauthentic News Sites (2/2)	Purchase Targeted Advertisements	Consumer Review Networks	Use Fake Experts	Post Content (2/3)		Flood Information Space (2/4)	Physical Violence (2/2)	Continue to Amplify	
	Dismiss (1/3)		Leverage Conspiracy Theory Narratives (2/2)	Develop Video-Based Content (2/2)	Create Inauthentic Social Media Pages and Groups	Fabricate Grassroots Movement		Discussion Forums (2/1)	Use Search Engine Optimisation			Incentivize Sharing (2/2)	Sell Merchandise	Exploit TDS/Content Moderation (2/2)	
	Dissuade from Acting (2/3)		Leverage Existing Narratives	Distort Facts (1/2)	Create Inauthentic Websites	Impersonate Existing Entity (2/3)		Email				Manipulate Platform Algorithm (2/1)		Play the Long Game	
	Distort			Obtain Private Documents (2/2)	Cultivate Ignorant Agents			Formal Diplomatic Channels							
	Distract			Reuse Existing Content (2/4)	Develop Owned Media Assets			Livestream (2/2)							
	Divide				Employ Commercial Analytic Firms			Media Sharing Networks (2/3)							
	Facilitate State Propaganda				Infiltrate Existing Networks (2/2)			Online Polls							
	Make Money (2/4)				Leverage Content Farms (2/2)			Social Networks (2/4)							
	Motivate to Act (2/2)				Prepare Fundraising Campaigns (2/2)			Traditional Media (2/3)							
	Undermine (2/4)				Prepare Physical Broadcast Capabilities										
					Recruit Malign Actors (2/3)										

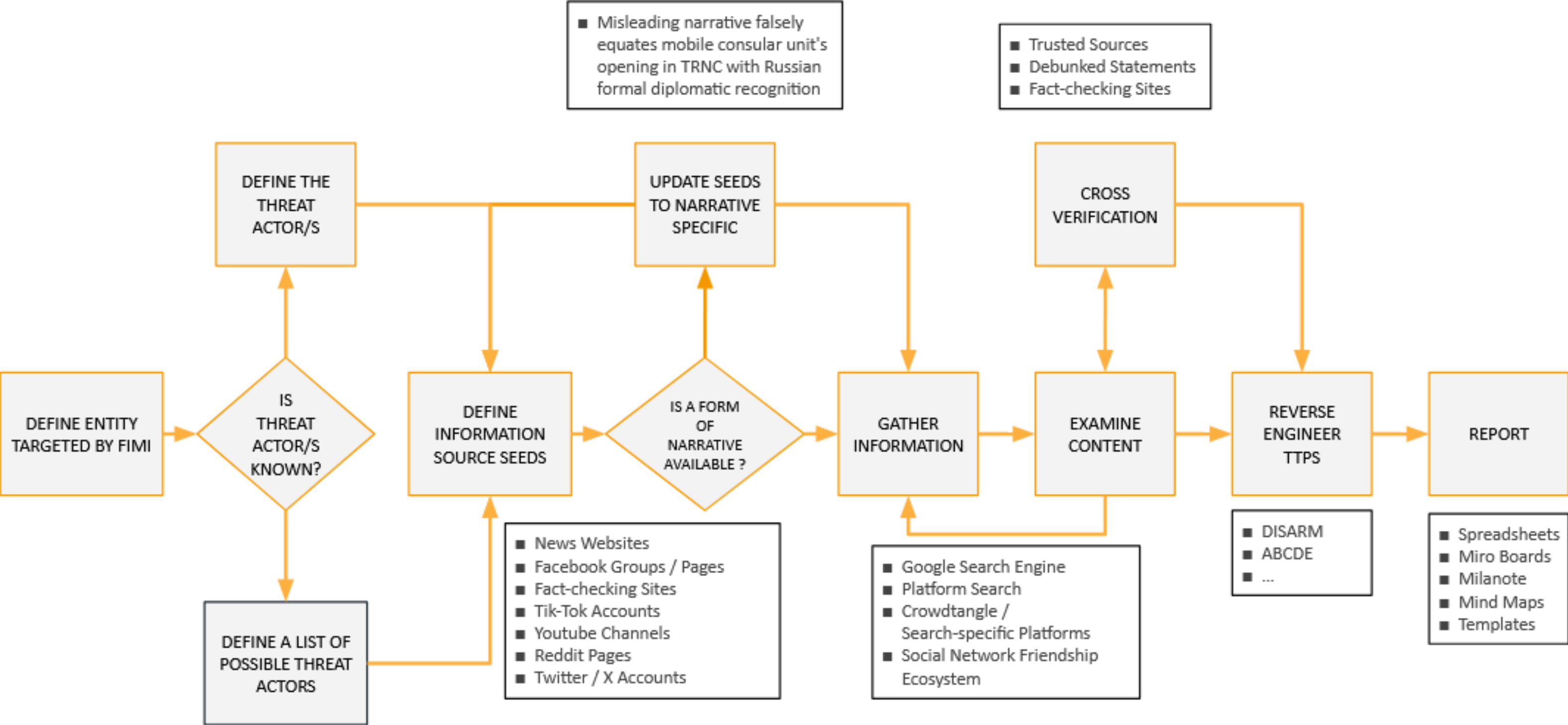


Funded by the European Union (grant number 101132686). UK participants in Horizon Europe Project ATHENA are supported by UKRI grant number 10107667 (Trilateral Research). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA) or UKRI. Neither the European Union nor the granting authority nor UKRI can be held responsible for them. Grant Agreement 101132686 ATHENA HORIZON-CL2-2023-DEMOCRACY-01.





# ATHENA FIMI Analysis Methodology





# Participatory Analysis














# Participatory Analysis Canvases



## "Observable" Analysis Canvas

<b>1. Actor/Provenance</b>  Actual individuals, groups, or organizations as well as classes of individuals, organizations, systems or groups (e.g., the finance sector) who publish/circulate the Observable and/or contribute to its content. Channel used to publish the content. The purpose of this component is to help assess the actor(s) involved in the case, and try to identify which kinds of actors produce and engage with the suspected disinformation. Sometimes actors disguise their origins and purposes, or disseminate suspicious content without malicious intent. <b>Guidelines:</b> Collect and analyze all available information to make an assessment, including secondary information, such as an attribution made by a digital platform or in a journalistic investigation, and clarify if actor(s) are: <ul style="list-style-type: none"><li>Individual(s): persons involved acting in their private capacity</li><li>Nonstate actor(s): persons affiliated with private or NG organization</li><li>Media platform: is the platform of distribution independent?</li><li>Political actor(s): does the individual act on behalf of a recognized political entity?</li><li>Foreign state(s): is the actor an agent or proxy of a foreign government?</li><li>Trolls, impersonators, fake personas etc.</li></ul>	<b>5. Tactics, Techniques / Behavior</b>  This component assesses to what extent deception, manipulation or other illegitimate communication techniques could be applied in the Observable under analysis. Also, it seeks to discover tactics used, which describe operational goals that threat actors are trying to accomplish. The component can also be used to analyze an actor's intent and evidence of manipulation or coordination, which are indicators of problematic behavior that could help shape potential countermeasures. Tactics, Techniques, and Procedures* are patterns of behaviour used by threat actors to manipulate the information environment with the intention to deceive. Tactics describe operational goals that threat actors are trying to accomplish. Techniques are actions describing how they try to accomplish it. Procedures are the specific combination of techniques across multiple tactics (or stages of an attack) that indicate intent and may be unique for different threat actors <b>Guidelines:</b> Apply different frameworks to identify the application of possible Tactics, Techniques, Procedures or Strategies used in disinformation campaigns. Add here tags describing the identified tactics and strategies. Please use: <ul style="list-style-type: none"><li>The DISARM framework analysis and taxonomy, which comprises a comprehensive definition of tactics, techniques and procedures used in disinformation campaigns</li><li>The DEPICT analysis framework, which defines "Six degrees of manipulation" often found in disinformation campaigns</li><li>CONSPIRE, the Conspiratorial Thinking Analysis framework, which defines key tactics used when exploiting conspiratorial thinking to spread and/or amplify disinformation</li><li>Stereotypes prevalent in the social and political milieu of the targets of this Observable, which may be exploited to amplify the propagation and acceptance of the story at hand</li><li>PLAAR to determine polarizing entities and concepts from news articles, which may be weaponized by the Observable's content to attract or dismay target audiences</li><li>Exploratory analysis of the content to identify whether the content at hand seeks to promote/amplify its message using false dichotomies, reappropriating or fear-mongering</li></ul>	<b>6. Objectives</b>  What are the presumed objectives of this Observable? What is the evidence from blocks 1-5 that these Objectives are plausible? How are these objectives linked to narratives? What are the manipulation objectives of this Observable? <b>Does the Observable seek to achieve one of the following objectives?</b> <ul style="list-style-type: none"><li>Discredit opponents</li><li>Dismay</li><li>Distract</li><li>Divide</li></ul> <b>Does the Observable pursue one or more of the following degrees of Manipulation?</b> <ul style="list-style-type: none"><li>Using Emotional Language to attract attention and/or support</li><li>Instigating inter-group polarization</li><li>Creating Moral Outrage to gather support and/or polarize</li><li>Building audience's follower base for fake accounts/impersonation</li><li>Sweeping conspiracy theories</li><li>Evoking Outrage through Trolling</li></ul>	<b>7. Target Audience / Degree</b>  What is the distribution of the content? Which audiences were appear to be applied in the Incident under analysis. The degree dimension attempts to gauge and describe the w information environment, possibly crossing different channels (news media, etc), targeting different linguistic, ethnic, social ( operations can reveal threat actor preferences with regards to channels in a network as source, amplifier or link to other net			
<b>2. Content</b>  What kinds of content is within this Observable? This line of questioning can help establish, for example, whether the information being deployed is deceptive or weaponized against specific target(s). The content of the incident provides details on the narratives used; the socio-political context where the incident takes place, which language, style, and content format are used. Stylistic and morphological features, along with sensational titles and writing, misleading framing of facts or the promotion of outright lies in the content can help identify if the Observable is circulated as part of disinformation campaigns. Questions to ask, include [Parment, 2020]: <ul style="list-style-type: none"><li>Channel: What is the channel wherein is this content published?</li><li>Truthfulness: Is the content verifiably untrue or deceptive?</li><li>Vulnerability: Does the content exploit a particular vulnerability of its intended audience or target?</li><li>Language(s): Which languages are used in the spread of the disinformation or other online content in question?</li><li>Synthetic: Is the content manipulated or artificial?</li><li>Expression: Is the content reasonable self-expression protected by fundamental freedoms?</li><li>Harm: Is the content harmful?</li><li>AI: Does the content appear to be drafted by generative AI?</li><li>Micro-targeting: Does the content appear to be crafted for dissemination through micro-targeting platforms?</li><li>Target: Is the content targeting a particular person, institution, organization or country?</li></ul>	<b>8. Attribution</b>  Is there any evidence about who is behind this Observable, v Evidence can be derived from the profile of the Actor or prior credible analyses published for the Actor. Sometimes the attri Observable promotes stated and documented political object those objectives or even if its originator is not a state actor or those objectives.	<b>4. Context / Theme / Narrative</b>  Based on the preliminary analysis of 1.Provenance/Actor, 2.Content and 3.Timing, what is a broader socio-political, economic context that could provide a plausible framing for the main points and/or targets of the Observable's message and/or story? Does the content align with known disinformation narratives, stereotypes against social, ethnic, racial, gender groups, conspiratorial thinking, explicit or hidden political objectives and agendas? <b>Contextualization and Narrative Analysis</b> Identify and classify the theme of the Observable's content into a broader theme of political, social, financial or cultural theme. Apply Narrative Theory to extract from the Observable's content its core narrative elements. Explore if identified narrative(s) align with known disinformation narratives? <b>Reach</b> Immediate audience Reach outside in-groups <b>Longevity</b> Does the Observable attract an interest from audiences that is persistent with time? For how long? <b>Organic Engagement</b> Do the recipients of the Observable appear to engage with it, commenting or sharing it within its publication platform or beyond, in other channels?	<b>9. Impact / Effect</b>  What is the overall impact of the Observable and whom does it affect? This question can help establish the actual harms and severity of the case. The effect of an Observable (or the impact or severity) can be measured and assessed according to different parameters such as the reach, the reach outside of in-groups, engagement, harm or behaviour-change caused offline, longevity etc. <b>Post Data</b> Felt Step Rise	<b>3. Timing</b>  What is the timing of this Observable's circulation? Are there any current events of relevance, for example election campaigns, major political developments or public controversies of relevance that may attract the attention of the public opinion?	<b>10. Vulnerabilities</b>  How do the objectives, the narrative, and the TTPs of the attack behind this incident, resonate with vulnerabilities of the target audience(s)? Provide a list of vulnerabilities, which may have been weaponized to propagate and/or amplify a particular narrative. Vulnerabilities can be identified through ethnographic studies, polls and questionnaires, media analysis, etc. Consider vulnerabilities such as: <ul style="list-style-type: none"><li>Stereotypes about and distrust towards particular social, ethnic, religious, gender etc groups</li><li>Distrust towards authorities, politicians, "the system"</li><li>Common beliefs</li><li>Polarized thinking and affective polarization</li><li>Popularity of conspiracy theories</li><li>Suspicion towards media, the political and judicial system</li></ul>	<b>11. Impact / Effect</b>  What is the overall impact of the Incident and whom does it affect? This question can help establish the actual harms and severity of the case. The effect of an Incident (or the impact or severity) can be measured and assessed according to different parameters such as the reach, the reach outside of in-groups, engagement, harm or behaviour-change caused offline, longevity etc. <b>Establish metrics and gather quantitative and qualitative evidence regarding impact through measurements, polls, etc. For example:</b> <ul style="list-style-type: none"><li>Has the incident penetrated national media?</li><li>Has it become a matter of discussion and/or conflict in political fora?</li><li></li></ul>












### Glossary

According to the European External Action Service's definition, an **observable** is a "concrete element relevant to understand how an incident unfolded – such as a tweet, a video on YouTube or an article on a website. **Observables** can be represented via the URL under which they were found or as files."

According to the European External Action Service's definition, **Foreign Information Manipulation and Interference (FIMI)** describes "a mostly non-illegal pattern of behaviour that threatens or has the potential to negatively impact values, procedures, and political processes. Such activity is manipulative in character, conducted in an intentional and coordinated manner. Actors of such activity can be state or non-state actors, including their proxies inside and outside of their own territory."

Disinformation Observables Analysis Canvas© 2024 by Marius D. Dikaiakos is licensed under Attribution-NonCommercial-ShareAlike 4.0 International

## Disinfo. Incident Analysis Canvas.

<b>1. Description</b>  <b>What is this Incident about?</b> <b>What is the story</b> conveyed by the Incident's Observable(s)? <b>What are the apparent targets</b> of the disinformation attack?	<b>5. Tactics, Techniques, Procedures</b>  <b>How</b> are the actor(s) of this Incident trying to manipulate the information environment with the intention to deceive their target audience(s)? Consider: <ul style="list-style-type: none"><li>The extent to which deception, manipulation or other illegitimate communication techniques appear to be applied in the Incident under analysis.</li><li>The tactics used, which describe operational goals that threat actors are trying to accomplish.</li><li>Analyzing an actor's intent and evidence of manipulation or coordination, which are indicators of problematic behavior that could help shape potential countermeasures.</li><li>Discovering Tactics, which describe operational goals threat actors are trying to accomplish.</li><li>Discovering Techniques, which describe how actors try to accomplish their operational goals.</li><li>Analyze possible Procedures, namely specific combination of techniques across multiple tactics (or stages of an attack) that indicate intent and may be unique for different threat actors.</li></ul> <b>DISARM:</b> use the DISARM framework analysis and taxonomy to map the TTPs identified in the Incident at hand. <b>DEPICT:</b> select one or more degrees of manipulation identified in this Incident: <input type="checkbox"/> Discrediting <input type="checkbox"/> Emotion <input type="checkbox"/> Polarization <input type="checkbox"/> Impersonation <input type="checkbox"/> Conspiracy <input type="checkbox"/> Trolling <b>CONSPIRE:</b> use the CONSPIRE framework of analysis to identify conspiratorial techniques applied in the Incident at hand. <input type="checkbox"/> Contradictory <input type="checkbox"/> Nefarious Intent <input type="checkbox"/> Persecuted Victim <input type="checkbox"/> Overriding Suspicion <input type="checkbox"/> Something must be wrong <input type="checkbox"/> Immunity to Evidence <input type="checkbox"/> Re-interpreting randomness <b>KNOWN MISINFORMATION STRATEGIES:</b> identify if known strategies apply in this Incident <input type="checkbox"/> Moral Outrage <input type="checkbox"/> Impersonation <input type="checkbox"/> Amplifying Stereotypes <input type="checkbox"/> False Dichotomies <input type="checkbox"/> Scapegoating <input type="checkbox"/> Fear Mongering <input type="checkbox"/> Spreading Conspiracies <b>Kill-chain Analysis:</b> explore if the incident can be mapped to one of the kill-chain analysis' phases <input type="checkbox"/> Step 1: Find the cracks <input type="checkbox"/> Step 5: Conceal your hand <input type="checkbox"/> Step 2: Seed distortion <input type="checkbox"/> Step 6: Cultivate Useful Idiots <input type="checkbox"/> Step 3: Wrap narratives in the kernel of truth <input type="checkbox"/> Step 7: Deny involvement <input type="checkbox"/> Step 4: Build audiences <input type="checkbox"/> Step 8: Play the long game	<b>6. Objectives</b>  Explain why are the attack(s) captured in this incident launched? What are the presumed objectives of this Incident based on evidence and analysis collected from blocks 1-5 ? How are these objectives linked to narratives? What are the manipulation objectives of this Observable? <b>Does the Observable seek to achieve one of the following objectives?</b> <ul style="list-style-type: none"><li>Discredit opponents</li><li>Dismay</li><li>Distract</li><li>Divide</li></ul> <b>Does the Observable pursue one or more of the following degrees of Manipulation, as identified in block 5?</b> <ul style="list-style-type: none"><li>Using Emotional Language to attract attention and/or support</li><li>Instigating inter-group polarization</li><li>Creating Moral Outrage to gather support and/or polarize</li><li>Building audience's follower base for fake accounts/impersonation</li><li>Spreading conspiracy theories</li><li>Evoking Outrage through Trolling</li></ul>	<b>9. Audience / Degree</b>  <b>Which</b> is the audience(s) targeted and reached by the attack behind this incident? What is the distribution of the content? The degree dimension attempts to gauge and describe the way the incident has traveled through the information environment, possibly crossing different channels (shared via Social Media, story picked-up by other news media, etc), targeting different linguistic, ethnic, social or age groups. This view on disinformation operations can reveal threat actor preferences with regards to targeted platforms and identify different roles of channels in a network as source, amplifier or link to other networks.	<b>10. Vulnerabilities</b>  How do the objectives, the narrative, and the TTPs of the attack behind this incident, resonate with vulnerabilities of the target audience(s)? Provide a list of vulnerabilities, which may have been weaponized to propagate and/or amplify a particular narrative. Vulnerabilities can be identified through ethnographic studies, polls and questionnaires, media analysis, etc. Consider vulnerabilities such as: <ul style="list-style-type: none"><li>Stereotypes about and distrust towards particular social, ethnic, religious, gender etc groups</li><li>Distrust towards authorities, politicians, "the system"</li><li>Common beliefs</li><li>Polarized thinking and affective polarization</li><li>Popularity of conspiracy theories</li><li>Suspicion towards media, the political and judicial system</li></ul>
<b>2. Target</b>  <b>What is the target</b> of the attack behind this Incident? <b>Consider:</b> <ul style="list-style-type: none"><li>Individual Person</li><li>Cultural / Religious / Ethnic/Professional Group</li><li>Institution</li><li>Organization</li><li>Country</li><li>International Alliance</li></ul>	<b>4. Narrative</b>  <b>What is the key Narrative</b> promoted and/or amplified through this Incident, based on its Description, the apparent Target of the attack, and the identified Context? What known narratives, stereotypes against social, ethnic, racial, gender groups, or conspiracy theories, explicit or hidden political objectives and agendas? <b>Perform Contextualization and Narrative Analysis:</b> <ul style="list-style-type: none"><li>Apply Narrative Theory to extract from the Incident's description and target its core narrative elements.</li><li>Identify and classify the Incident's narrative into a broader frame of political, social, financial or cultural theme.</li><li>Explore if identified narrative(s) align with known disinformation narratives?</li></ul>	<b>7. Attribution / Actor</b>  Identify who could be behind this incident? Actual individuals, groups, or organizations as well as classes of individuals, organizations, systems or groups (e.g., the finance sector) who appear to have a motive to orchestrate this Incident and/or contribute to its content and narrative. <b>Guidelines:</b> The purpose of this component is to help assess the actor(s) involved in the case, and try to identify which kinds of actors produce and engage with the suspected disinformation. Sometimes actors disguise their origins and purposes, or disseminate suspicious content without malicious intent. Collect and analyze all available information to make an assessment, including secondary information, such as an attribution made by a digital platform or in a journalistic investigation, and clarify if actor(s) are: <ul style="list-style-type: none"><li>Individual(s): persons involved acting in their private capacity</li><li>Nonstate actor(s): persons affiliated with private or NG organization</li><li>Media platform: is the platform of distribution independent?</li><li>Political actor(s): does the individual act on behalf of a recognized political entity?</li><li>Foreign state(s): is the actor an agent or proxy of a foreign government?</li><li>Trolls, impersonators, fake personas etc.</li></ul>	<b>8. Channels</b>  <b>Where</b> has this Incident been manifested? Which channels were targeted and reached? <b>Where</b> has this Incident gained traction? How different platforms have been weaponized to support the attack behind the incident? <b>Channel list (non-exhaustive)</b> <input type="checkbox"/> Web <input type="checkbox"/> Facebook <input type="checkbox"/> Twitter <input type="checkbox"/> TikTok <input type="checkbox"/> Instagram <input type="checkbox"/> Reddit	<b>12. Countermeasures</b>  Which steps or combination of measures can be taken to address the Incident at hand, if it contains misleading or manipulative content? <b>Possible Approaches</b> Debunking Refutation Steps to confine the circulation of the Incident Expose TTPs Raising awareness about misleading narrative
<b>3. Context / Theme</b>  <b>What is the broader socio-political, economic context</b> that could provide a plausible framing for the main points and/or targets of this Incident? <b>What themes</b> , stated or implied political and other agendas have the same target with the Incident and provide a reasonable frame for it?	<b>Glossary</b> European Union's External Action Service defines a <b>FIMI incident</b> as "an action perpetrated by one or more threat actor(s) pursuing specific objectives and carried out with the intent to deceive. It is composed of a combination of observables and TTPs. Multiple related incidents can be part of a campaign." According to the European External Action Service's definition, <b>Foreign Information Manipulation and Interference (FIMI)</b> describes "a mostly non-illegal pattern of behaviour that threatens or has the potential to negatively impact values, procedures, and political processes. Such activity is manipulative in character, conducted in an intentional and coordinated manner. Actors of such activity can be state or non-state actors, including their proxies inside and outside of their own territory."	<b>Disinformation Incident Analysis Canvas© 2024 by Marius D. Dikaiakos is licensed under Attribution-NonCommercial-ShareAlike 4.0 International</b>		

Source: <https://github.com/dikaiakos/FIMI-Map-Canvas/>



Funded by the European Union (grant number 101132686). UK participants in Horizon Europe Project ATHENA are supported by UKRI grant number 10107667 (Trilateral Research). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA) or UKRI. Neither the European Union nor the granting authority nor UKRI can be held responsible for them. Grant Agreement 101132686 ATHENA HORIZON-CL2-2023-DEMOCRACY-01.

# Participatory Analysis Canvases: FIMIScope



- ❑ A software toolset prototype that implements the FIMI canvases as a graphical user interface running on a browser and supported by a back-end software system which provides services for:
  - ❑ Storing, managing, and exporting data and meta-data collected during the analysis process, in a simple JSON format;
  - ❑ Supporting remote collaboration among analysts collaboratively analyzing cases, and
  - ❑ Keeping snapshots of different stages of an analysis exercise and versioning.
- ❑ FIMIScope is enhanced with a Large-Language Module tool, which retrieves JSON data collected and creates narratives describing captured FIMI incidents and campaigns, and allows analysts to interact with the collected data through a natural language, dialogue interface.
- ❑ Video demo: <https://www.youtube.com/watch?v=B0D45tWfsIA>





# ATHENA Case Studies



- ❑ The ATHENA partners have contacted a detailed FIMI analysis on 32 different case studies.
  - Each analysis included lists of observables, potential threat actors, and DISARM TTPs.

- Did the Russians dupe the US Republican Party?
- How Russian FIMI actors reacted to the US aid package.
- North Korea's disinformation campaigns.
- Iranian disinformation campaigns.
- Russian initiatives to use farmers' protests in Germany.
- Russian FIMI about the (fabricated) "Lisa" rape case.
- Russian FIMI about burning the Quran in Sweden.
- Russia generates tension along its border with Finland.
- Russian propaganda against Finland's transition to NATO member.
- Russia Today (RT), its narratives and ideological biases.
- A Wagner campaign targeting the French army in Mali.
- Russia uses bribery to spread disinformation.
- Russian interference in the Spanish election in July 2023.
- European Parliament report on FIMI in European elections.
- Altered photo showing Zelensky holding a jersey with a swastika.
- Disinformation about Zelensky's buying a villa in Florida.

- Doppelganger/RNN: Russian disinformation using media clones and more.
- Russian FIMI against the Ukrainian armed forces.
- Russian disinformation about forthcoming Russian attacks.
- Russian disinformation campaign claims "inevitable" victory against Ukraine.
- Russian disinformation about COVID vaccines.
- Chinese disinformation about the release of Fukushima water.
- Meta dismantles large-scale Chinese disinformation campaign.
- Facebook accounts in China impersonated Americans, Meta says.
- PAPERWALL: Chinese websites pose as local news to target global audiences.
- India using disinformation to discredit Pakistan and other regional powers.
- UAE FIMI targets critics at home and abroad.
- Saudi Arabia's Anti-Iran Campaign During the Trump Administration.
- Cyberwarfare and the Qatar Blockade.
- A CASE STUDY OF THE 2019-2020 IRAQI PROTESTS.
- Turkey's disinformation about a clash with the UN in the buffer zone in Cyprus.
- Turkey's disinformation about north Cyprus as an independent state.



# ATHENA Case Studies Thematics



## Russian Disinformation and Information Warfare 14 / 32

- Did the Russians dupe the US Republican Party?
- How Russian FIMI actors reacted to the US aid package
- Russian initiatives to use farmers' protests in Germany
- Russian FIMI about the (fabricated) "Lisa" rape case
- Russian FIMI about burning the Quran in Sweden

- Russian propaganda against Finland's transition to NATO member
- Russia Today (RT), its narratives and ideological biases
- A Wagner campaign targeting the French army in Mali
- Russia uses bribery to spread disinformation
- Framing the president - how Russia creates negative press on a head of state

- Russian interference in the Spanish election in July 2023
- European Parliament report on FIMI in European elections
- Doppelganger/RNN: Russian disinformation using media clones and more
- Russian disinformation about COVID-19 vaccines

## Russian Disinformation against Ukraine 4 / 32

- Altered photo showing Zelensky holding a jersey with a swastika
- Disinformation about Zelensky's buying a villa in Florida
- Russian FIMI against the Armed Forces of Ukraine (AFU)
- Russian disinformation about Russian attacks

## Chinese Disinformation and Influence Operations 4 / 32

- Meta dismantles large-scale Chinese disinformation campaign
- Facebook accounts in China impersonated Americans, Meta says
- PAPERWALL: Chinese websites pose as local news to target global audiences
- Chinese disinformation about the release of Fukushima water

## Turkish Disinformation Targeting the Republic of Cyprus 3 / 32

- Turkey's disinformation about north Cyprus as an independent state
- Turkey's disinformation about a clash with the UN in the buffer zone in Cyprus
- Varosha to Vegas: The real estate exploitation and disinformation tactics by "TRNC"



# Understanding the Use of TTPs by Adversaries



❑ Understanding the usage of TTPs across existing FIMI campaigns is critical to gaining insights into how disinformation operations evolve.

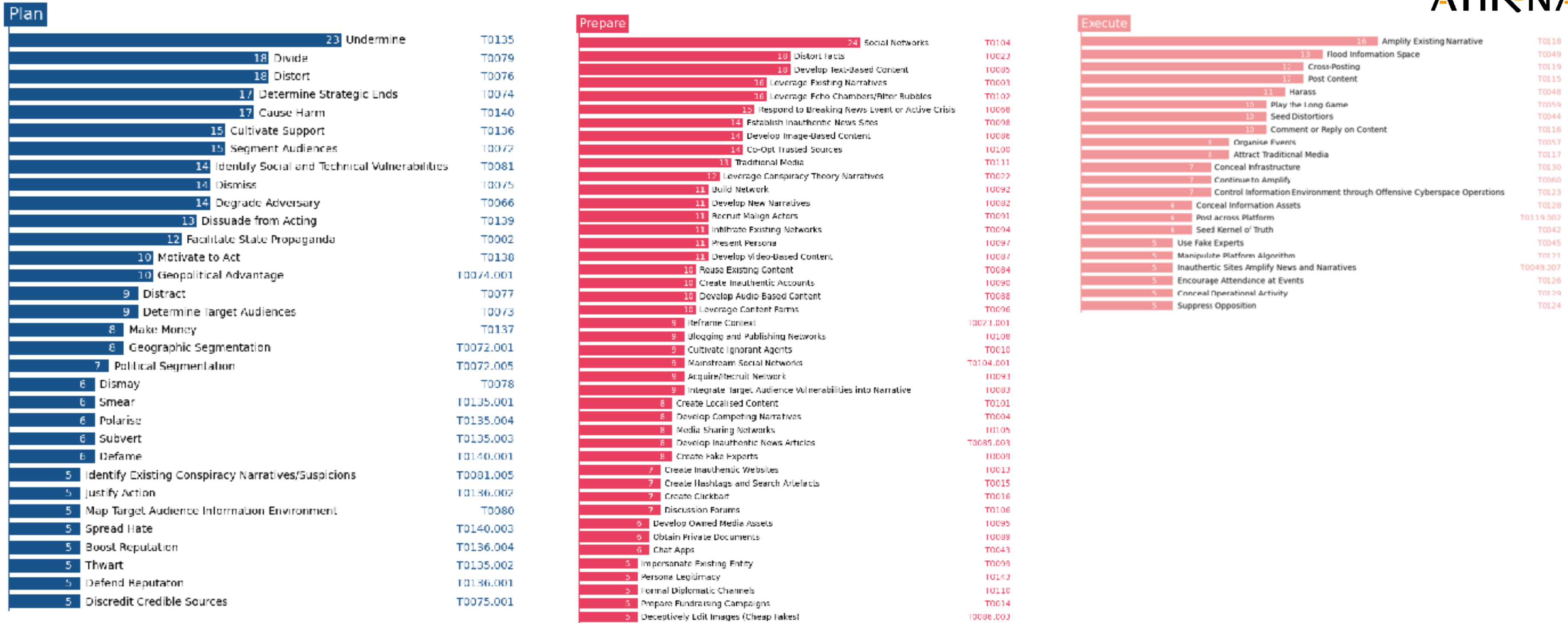
## ❑ Analysis Outline:

- TTP Frequency Analysis: across Case Studies and Thematics.
- Combination of TTPs across campaigns
- Pattern Recognition:
  - Clustering by TTP Usage
  - Frequent Itemsets

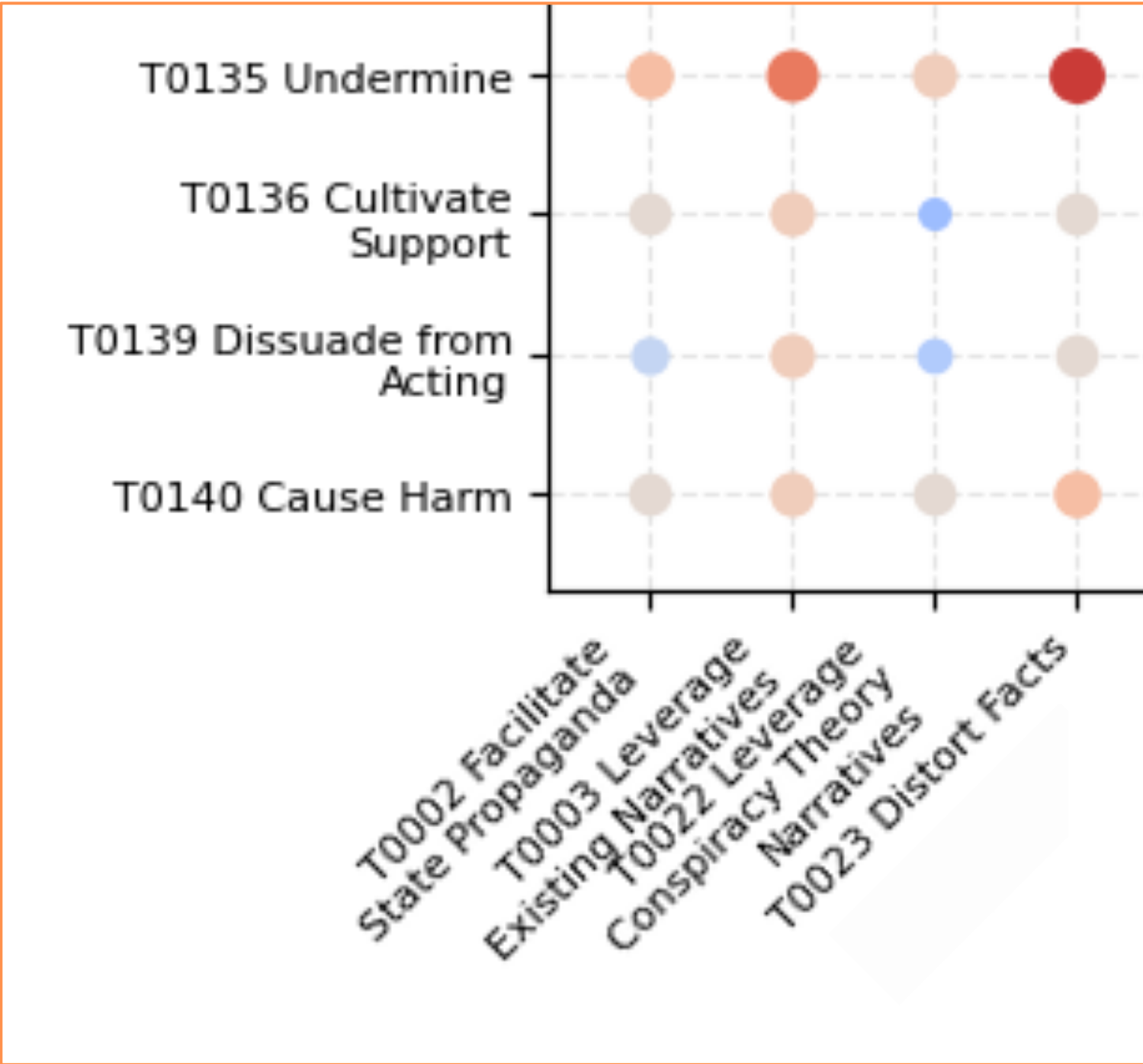




# Frequency Analysis of TTPs



# TTP Co-occurrence Analysis



# Patterns of TTP Usage Across Case Studies



- ❑ Frequent itemset analysis to identify recurring combinations of TTPs across campaigns.

## ❑ T0085 Develop Text-Based Content, T0103 Social Networks

### ❑ Occurrence: 57%

- Widely used by Russia, China, Turkey, North Korea.
- Common in election disinformation and geopolitical narratives.
- Amplifies influence via mass social media reach.

## ❑ T0023 Distor Facts, T0135 Undermine

### ❑ Occurrence: 53%

- Used to destabilize public perception.
- Creates confusion about political and public health issues.





# Patterns of TTP Usage Across Case Studies



❑ Frequent itemset analysis to identify recurring combinations of TTPs across campaigns.

❑ **T0023 (Distort Facts), T0085 (Develop Text-Based Content), T0104 (Social Networks)**

❑ **Occurrence: 47%**

- Facts are manipulated through written content.
- Disseminated via social platforms to enhance credibility.

❑ **T0102 (Leverage Echo Chambers/Filter Bubbles), T0085 (Develop Text-Based Content), T0104 (Social Networks)**

❑ **Occurrence: 43%**

- Tailors content to reinforce biases.
- Social networks entrench views, limiting exposure to alternative perspectives.

# Key Takeaways



## ❑ Understanding Evolving Disinformation Strategies:

- Disinformation campaigns are highly adaptive and strategic.
- Frequent use of "Develop Text-Based Content" (T0085) and "Social Networks" (T0104) highlights a sophisticated approach.
- **Countermeasures must evolve alongside threat actor strategies.**

## ❑ Recurring Themes across Multiple Actors:

- Russia, China, Turkey, and North Korea employ similar TTPs.
- "Distort Facts" (T0023) and "Amplify Existing Narrative" (T0118) are common.
- **International cooperation is required for effective mitigation.**



# Key Takeaways



## ❑ Amplification through Echo Chambers:

- Echo chambers and filter bubbles (T0102) reinforce manipulated content.
- Techniques like "Flood Information Space" (T0049) and "Cross-Posting" (T0119) increase reach.
- **Addressing algorithmic reinforcement of disinformation is critical.**

## ❑ Use of Multimedia to Enhance Credibility:

- Strategic combination of text-based (T0085), image-based (T0086), and video content (T0087) increases persuasiveness.
- Multimedia disinformation is harder to detect.
- AI-generated content presents an emerging and significant challenge.
- **Advanced verification technologies are necessary.**





# Generative AI and TTPs



Unit/ies		DISARM Techniques	Description	Domain of Attack	Technologies	Emerging Services / Systems			
Phase	Tactic					Future Vulnerabilities			
1. Planning and Preparation	Establish Identities and Assets	Create fake social media profiles/pages/groups (T0007)	Involves creating a foundation infrastructure of fake identities and platform aligning DISARM's focus on resources and network setup before active operations. Science: Fake journals/conferences, hijack repositories.	Social media (Facebook/T), domain registers, predatory journals, arXiv/ResearchGate	LLMs (generate fake profiles); <b>Agentic systems</b> (automate account creation).	<b>AI-curated patterns:</b> LLM-generated personas manipulate algorithmic news feeds to amplify disinformation, eroding trust in AI-mediated content and reducing the LLM reliance.			
		Create fake websites (T0013)		AI-Curated News	LLMs (generate website content)				
		Create funding campaigns (T0014)							
	Gather Information	Center of Gravity Analysis (T0005)	Researching targets, identifying vulnerabilities and analyzing the information environment to refine strategic planning, matching DISARM's emphasis a audience analysis and environment intelligence. Harvest researcher data establish presence in open data and open model repositories.		<b>Hijack legitimate accounts (T0011)</b>	LLMs (craft phishing messages).	assessments, diverting resources from real vulnerabilities. <b>Biometric Authentication Systems:</b> Facial recognition or voice ID for login. Deepfake biometrics bypass authentication, hijacking high-trust accounts.		
				<b>Narratives to achieve SDs (T0001)</b>	Developing narratives, defining objectives, and coordinating the team to ensure a cohesive operation, aligning with DISARM's strategy and content planning. Weaponize debates, create pseudoscience narratives.	Secure messaging apps. Project management software. Online collaboration platforms. Dark web forums. AI narrative managers, metaverse ecosystems. Preprint comments.	LLMs (generate tailored narratives). <b>AI Narrative Managers:</b> Systems that auto-generate news summaries for media outlets.	<b>Emotion-Aware AI:</b> Fake sentiment datasets trick AI into amplifying disinformation or harmful recommendations. <b>Open Repositories:</b> Poisoned	
				<b>Leverage existing narratives (T0003)</b>		LLMs (adapt stories)			
		<b>Competing narratives (T0004)</b>		LLMs (general narratives).					
	Testing Platform Defenses	<b>Trial content (T0020)</b>	Probing platform moderation systems by deploying trial content and refining approaches, ensuring disinformation will survive on targeted channels—central to DISARM's preparation. Submit flawed papers, manipulate citations.	Social media platforms, Online forums, Content moderation testing platforms (potentially underground), Reddit/YouTube, peer-review platforms (ScholarOne), citation databases.	LLMs (general content). <b>Pred Virality Engine</b> that forecast trends. <b>Agentic</b> to automate late tests trying out and impact.				
		<b>Manipulate online polls (T0029)</b>			<b>Agentic systems</b> (automate poll manipulation). <b>Polling System</b> time sentiment elections or pu				
2. Content Development and Deployment	Deploying Assets & Evading Detection	<b>Use concealment (T0012)</b>	Concealing the true origins and affiliations of assets to appear legitimate, aligning with DISARM's focus on hiding operational tracks and establishing plausible identities. Forge institutional affiliations, launder funding.			VPNs, Proxy servers, Anonymous browsing networks (Tor), Social media platforms (for hijacked accounts), Domain registration services (for fake websites), ORCID, institutional websites, Quantum Encryption Services, Unhackable communication channels.	<b>Agentic systems</b> (dynamic IP masking), <b>VPNs.</b> Quantum encryption used to hide disinformation operations from detection.	Synthetic identity generation, adversarial AI to mimic browsing through data laundering. <b>Algorithmic bias in reproducibility checks,</b> adversarial AI mimicking "normal" behavior.	
		<b>Backstop personas (T0030)</b>					LLMs (generate backstories). <b>Deepfake Social Avatars:</b> AI-generated human-like personas for customer service.	Disinformation actors create synthetic personas indistinguishable from humans.	
		<b>Deny involvement (T0041)</b>					LLMs (generate plausible deniability narratives). <b>AI Legal Advisors:</b> Automated systems for legal compliance.	AI advisors exploited to craft loophole-heavy deniability strategies.	
		<b>Use concealment (T0012)</b>					<b>Agentic systems</b> (dynamic adaptation to detection). <b>Self-Learning Security AI:</b> Systems that adaptively detect threats.	Adversarial AI learns to bypass detection by mimicking "normal" behavior. <b>Algorithmic bias in reproducibility checks,</b> steganography to reveal provenance of data, adversarial AI mimicking "normal" behavior.	
	<b>Deny involvement (T0041)</b>					VPNs, Proxy servers, Anonymous browsing networks (Tor), Social media platforms, Encrypted communication platforms. Self-learning security AI, GitHub, OSF, blockchain IDs.	<b>LLMs</b> (generate plausible deniability narratives). <b>Decentralized Identity Systems:</b> User-controlled digital IDs (e.g., blockchain-based).	Fake IDs created on decentralized systems, making attribution impossible.	
	<b>Twitter bots amplify (T0054)</b>					Science: Exploit open science for irreproducible research.	<b>Agentic systems</b> (manage bot networks).	<b>AI Community Moderators:</b> Automated systems for content moderation. Bots mimic human behavior to evade AI moderation filters.	
	<b>Flooding (T0009)</b>					Brass dissemination of false content to pollute the information space at scale.	Social media platforms, Online forums, News.	LLMs (generate comments/posts). <b>Holographic Influencers:</b>	Fake holographic influencers spread disinformation in immersive environments.



Funded by the European Union (grant number 101132686). UK participants in Horizon Europe Project ATHENA are supported by UKRI grant number 10107667 (Trilateral Research). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA) or UKRI. Neither the European Union nor the granting authority nor UKRI can be held responsible for them. Grant Agreement 101132686 ATHENA HORIZON-CL2-2023-DEMOCRACY-01.

# Thank you

E-mail: [mdd@ucy.ac.cy](mailto:mdd@ucy.ac.cy)