



# Protocols for information sharing, structured data, and response

## CD.RRM.2 + CD.PROTOCOLS.1

*Stephen Campbell and 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



# Information Sharing is best when there are standards

- When different organizations use different standards to catalogue incidents and campaigns it becomes difficult to quickly share actionable intelligence
- Tools like DISARM and STIX allow for consistent machine readable standards.

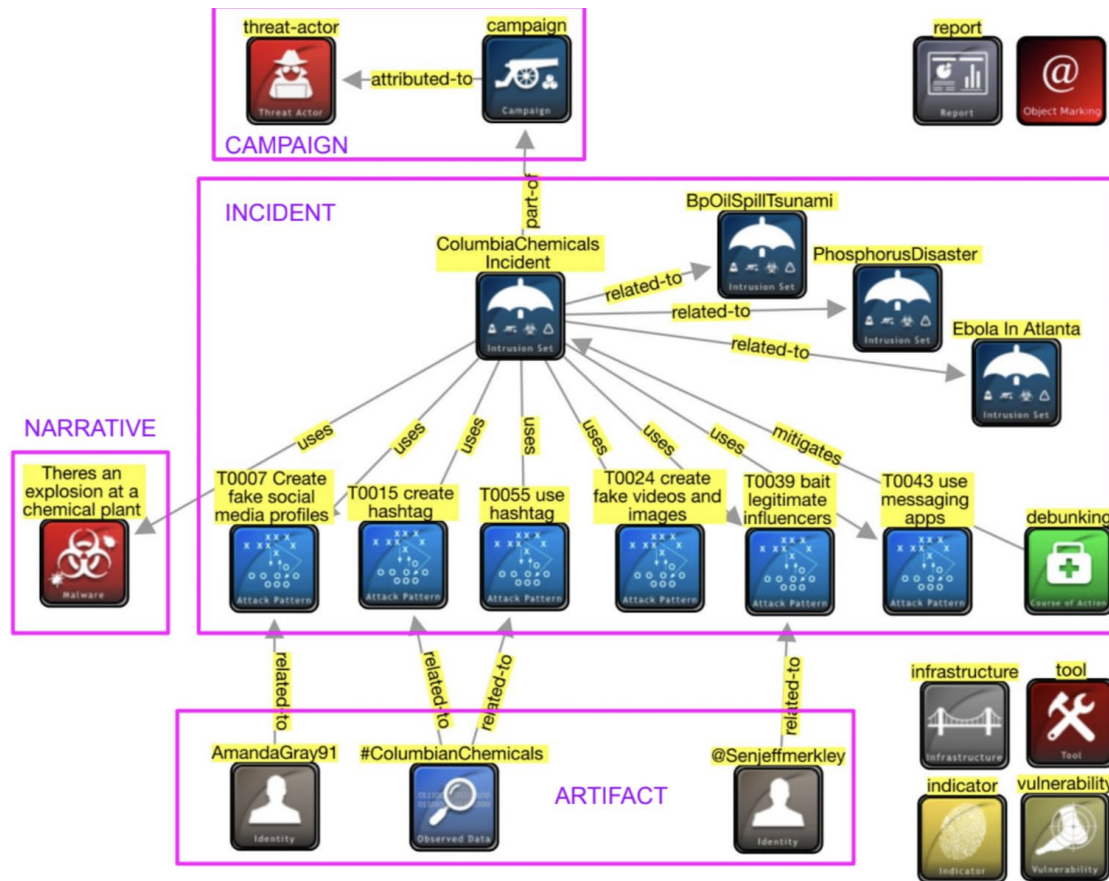
# What are Taxonomies For?

- Terms are used to tag, index, or categorize content to make it easier to be found and retrieved
- The taxonomy links the user to the desired content



*Courtesy of Heather Hedden*

# Structured Threat Information eXpression



## Assessing behaviors used using DISARM

- DISARM is a standard taxonomy of behaviors used to characterize malign influence, disinformation, and harm campaigns

## Union





CyberNet



# MITRE

## Cyber security list of Tactic Techniques and procedures

Software Execution x								
selection controls								
layer controls								
technique controls								
Initial Access 9 techniques	Execution 10 techniques	Persistence 18 techniques	Privilege Escalation 12 techniques	Defense Evasion 37 techniques	Credential Access 14 techniques	Discovery 25 techniques	Lateral Movement 9 techniques	Collection 17 techniques
Replication Through Removable Media	Native API	BITS Jobs	Process Injection (8/11)	Obfuscated Files or Information (5/5)	Credentials from Password Stores (3/3)	System Information Discovery	Replication Through Removable Media	Screen Capture
Drive-by Compromise	Windows Management Instrumentation	Hijack Execution Flow (7/11)	Access Token Manipulation (5/5)	Deobfuscate/Decode Files or Information	Network Sniffing	File and Directory Discovery	Data from Local System	Audio Capture
Valid Accounts (2/4)	Command and Scripting Interpreter (7/8)	Traffic Signaling (0/1)	Exploitation for Privilege Escalation	Modify Registry	OS Credential Dumping (8/8)	Process Discovery	Lateral Tool Transfer	Archive Collected Data (3/3)
Exploit Public-Facing Application	Exploitation for Client Execution	Valid Accounts (2/4)	Hijack Execution Flow (7/11)	Process Injection (8/11)	Brute Force (3/4)	System Network Configuration Discovery	Exploitation of Remote Services	Clipboard Data
External Remote Services	Shared Modules	Account Manipulation (1/4)	Valid Accounts (2/4)	Rootkit	Steal Web Session Cookie	System Owner/User Discovery	Taint Shared Content	Video Capture
Hardware Additions	Scheduled Task/Job (3/6)	Browser Extensions	Boot or Logon Autostart Execution (8/12)	Indicator Removal on Host (5/6)	Two-Factor Authentication Interception	Query Registry	Remote Services (6/6)	Automated Collection
Phishing (2/3)	Inter-Process Communication (2/2)	Boot or Logon Execution (8/12)	Group Policy Modification	Virtualization/Sandbox Evasion (3/3)	Unsecured Credentials (4/6)	System Network Connections Discovery	Software Deployment Tools	Data from Removable Media
Supply Chain Compromise (1/3)	System Services (2/2)	Compromise Client Software Binary	Scheduled Task/Job (3/6)	BITS Jobs	Exploitation for Credential Access	System Time Discovery	Internal Spearphishing	Man in the Browser
Trusted Relationship	User Execution (2/2)	External Remote Services	Abuse Elevation Control Mechanism (4/4)	Hijack Execution Flow (7/11)	Forced Authentication	System Service Discovery	Remote Service Session Hijacking (1/2)	Data from Network Shared Drive
		Scheduled Task/Job (3/6)	Boot or Logon Initialization Scripts (3/5)	Masquerading (5/6)	Input Capture (3/4)	Peripheral Device Discovery	Use Alternate Authentication Material (2/4)	Data from Cloud Storage Object
		Boot or Logon Initialization Scripts (3/5)	Create or Modify System Process (4/4)	Traffic Signaling (0/1)	Man-in-the-Middle (1/2)	Remote System Discovery		Data from Configuration Repository (0/2)
		Create Account (2/3)	Event Triggered Execution (10/15)	Valid Accounts (2/4)	Modify Authentication Process (3/4)	Application Window Discovery		Data from Information Repositories (1/2)
		Create or Modify System Process (4/4)		Indirect Command Execution	Steal	Network Service Scanning		Data Staged (1/2)
				Group Policy Modification		Network Share Discovery		
				Rogue Domain Controller				



# The DISARM Word Add-In (Javascript)

The screenshot shows a Microsoft Word document titled "GenocideGames.docx". The ribbon includes tabs for File, Home, Insert, Draw, Design, Layout, References, Mailings, Review, View, Developer, Zotero, Help, Acrobat, PDFelement, and Writage. The DISARM add-in is active, displaying a sidebar with the following sections:

- Choose Technique(s)**
- Search Technique(s)**
- Summaries**
- Insert Summary Red Table**
- Format Red**
- Red Tag**

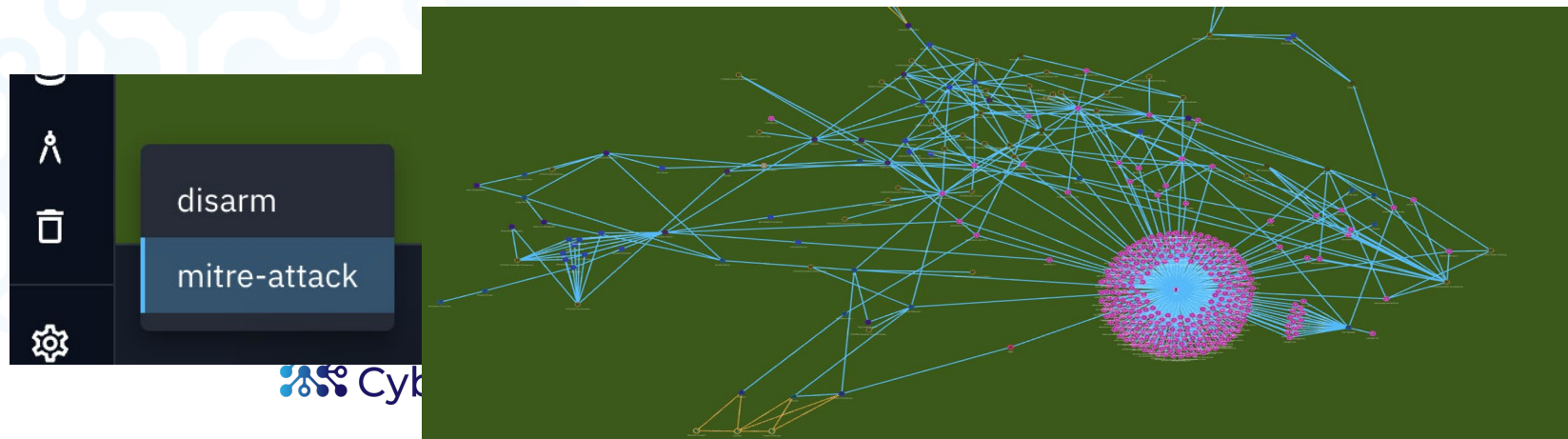
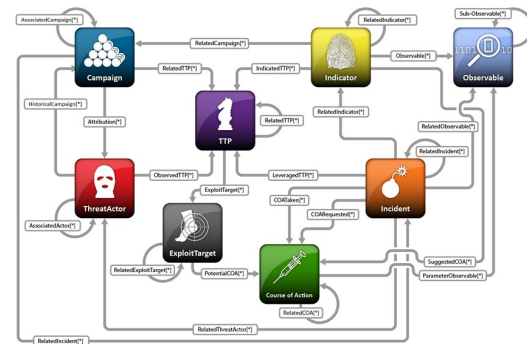
The document text reads:

#GenocideGames Network

The professors had identified a network of Twitter accounts that appeared to be designed to drown out criticism about Chinese treatment of the Uyghurs tagged with the hashtag #GenocideGames (Suppress Opposition [T0124], Flood Information Space: Flood Existing Hashtag [T0049.002]). They believed that the intent behind the network was to flood Twitter and make it more difficult for protesters to

Both interoperable with STIX (Structured Threat Information Expression)

## Both can map behaviors in threat intelligence platforms



# Infosec model: MITRE DEFEND

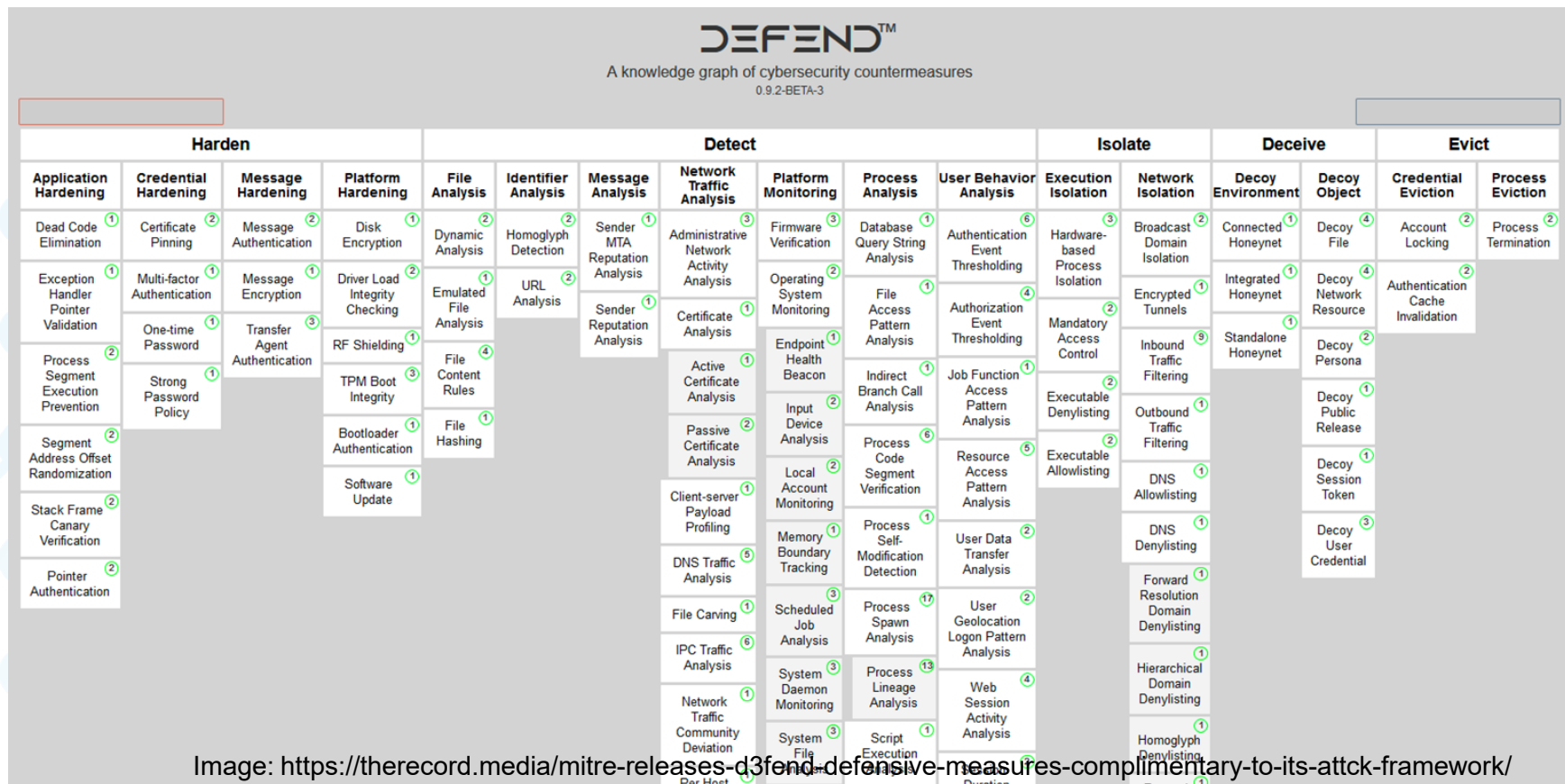


Image: DISARM Foundation

P01 Planning	P01 Planning	P01 Preparation	P01 Preparation	P01 Preparation	P01 Preparation	P01 Preparation	P01 Execution	P01 Execution	P01 Execution	P01 Execution	P01 Evaluation
T001 Strategic Planning	T001 Objective Planning	T002 Develop People	T004 Develop Networks	T003 Marketing	T006 Develop Content	T007 Channel Selection	T008 Pump Priming	T009 Exposure	T010 Physical	T011 Persistence	T012 Measure Effectiveness
C0006 Charge for social media	C0009 Educate high profile influencers on best practices	C0034 Create more friction at account creation	C0047 Coordinated inauthenticity	C0085 Reduce political targeting	C0014 Real-time updates to fact-checking database	C0097 Require use of verified identities to contribute to poll or comment	C0112 "Prove they are not an op!"	C0089 Threaten number of forwards to cut off access	C0029 Use banking to cut off access	C0131 Seize and analyse botnet servers	C0090 Fake engagement system
C0008 Create shared fact-checking database	C0011 Media literacy: Games to identify fake news	C0036 Infiltrate the in-group to discredit leaders (divide)	C0052 Infiltrate platforms	C0086 Create a hasbing and down it out (hijack it back)	C0032 Hijack content and link to truth-based info.	C0098 Revocation of "verified"	C0113 Debank and defuse a fake expert / credentials. Attract audience quality of fake expert	C0022 Content moderation: Censorship?	C0130 Mentoring elders, youth, credit. Learn vicariously.	C0033 Deplatform Account?	C0040 "Bomb" link shorteners with lots of calls.
C0010 Enhanced privacy regulation for social media	C0028 Make information provenance available	C0040 third party verification for people	C0053 Delete old accounts / Remove unused social media accounts	C00216 Use advertiser controls to stem flow of funds to bad actors	C0071 Block source of pollution	C0099 Strengthen verification methods	C0114 Don't engage with payloads	C0123 Bot control	C0135 Deplatform message groups and/or message boards	C0047 Make amplification of social media posts explicit (e.g. can't like/retweet after n days)	C0047 Make amplification of social media posts explicit (e.g. can't like/retweet after n days)
C0012 Platform regulation	C0029 Create fake website to issue counter narrative and counter narrative through physical merchandise	C0042 Address truth contained in narratives	C0056 Get off social media	C0027 Content censorship in non-relevant domains e.g. Pinterest antivax	C0072 Content censorship in non-relevant domains e.g. Pinterest antivax	C0100 Hashtag jacking	C0115 Expose actor and intentions	C0124 Don't feed the trolls	C0136 Microtarget most likely targets then send them counter messages	C0048 Add random links to network graphs	C0048 Add random links to network graphs
C0013 Rating framework for news	C0030 Develop a compelling counter narrative (truth based)	C0044 Keep people from posting to social media immediately	C0059 Verification of project before posting (counters funding campaigns)	C0074 Identify identical content and mass deplatform	C0075 normalise language	C0102 Make repeat voting harder	C0116 Provide proof of involvement	C0125 Prepare the population with pre-announcements	C0137 Pollute the AB-testing data feeds	C0049 Poison the monitoring & evaluation data	C0049 Poison the monitoring & evaluation data
C0016 Censorship - not recommended	C0031 Dilute the core narrative - create multiple permutations, target / amplify	C0046 Marginalise and discredit extremist groups	C0062 Free open library sources worldwide	C0076 Prohibit images in political discourse channels	C0077 normalise language	C0103 Create a bot that engages / distract trolls	C0118 Repurpose images with new text	C0126 Social media amber alert	C0138 Spam domestic actors with lawsuits		
C0017 Repair broken social connections	C0060 Legal action against for-profit engagement factories	C0048 Name and Shame Influencers	C00162 collect data/mapp constellations of Russian "civil society". Unravel/target the Potemkin villages	C0078 Change Search Algorithms for Disinformation Content	C0079 Prohibit images in political discourse channels	C0105 Buy more advertising than the adversary to shift influence and algorithms	C0119 Engage payload and debunk. Provide link to facts.	C0128 Create friction by marking content with ridicule or other "declarations"	C0139 Weaponise youtube content matrices	C0043 (botnet) DMCA takedown requests to waste group time	
C0019 Reduce effect of disinformation enablers	C0070 Block access to disinformation resources	C0051 Counter social engineering training	C0058 Report crowdfunder as violator	C0080 Create competing narrative	C0081 Highlight flooding and noise, and explain motivations	C0106 Click-bait centrist content	C0120 Open dialogue about design of platforms to produce different outcomes	C0151 "fight in the light"	C0143 Buy out troll farm employees / offer them jobs	C0045 Pollute the data voids with wholesome content (Kittens/ Babyshark?)	
C0021 Encourage in-person communication	C0092 Reputation scores for social media influencers	C0058 Report crowdfunder as violator	C0067 Denigrate the recipient/ project (of online funding)	C0073 Active defence: run T003 "develop people": not recommended	C0082 Ground truthing as automated response to pollution	C0107 Content moderation	C0121 Tool transparency and the resilience of a-risk populations.	C0158 Use training to build the resilience of a-risk populations.	C0169 develop a creative content hub		
C0022 Inoculate: Positive campaign to promote feeling of safety	C00164 compatriot policy	C0067 Denigrate the recipient/ project (of online funding)	C0073 Active defence: run T003 "develop people": not recommended	C0084 Modify disinformation narratives, and rebroadcast them	C0085 Mute content	C0109 De-escalation	C0122 Tool transparency and the resilience of a-risk populations.	C0154 Ask media not to report false info	C0178 Fill information voids with non-disinformation content		
C0024 Promote healthy narratives	C00207 Run a competing disinformation campaign - not recommended	C0093 Influencer code of conduct	C00155 Ban incident actors from funding sites	C0086 Distract from noise with additive content	C0087 Make more noise than the disinformation	C0110 Monetize centrist SEO by subsidizing the difference in greater clicks towards extremist content	C0123 Regulate alert when large batches of new URLs get registered together	C0188 Newsroom/Journalist training to counter SEO influence	C0182 malware detection/quarantine/deletion		
C0026 Shore up democracy based messages	C00222 Tabletop simulations	C00160 find and train influencers	C00189 Ensure that platforms are taking down flagged accounts	C0095 Limit access to alterable documents	C00165 Limit access to alterable documents	C0111 Present sympathetic views of opposite side	C0195 Redirect Method	C0189 Newsroom/Journalist training to counter SEO influence	C0184 Media exposure		
C0027 Create culture of civility		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0112 Present sympathetic views of opposite side	C0196 Redirect Method	C0203 Stop offering press credentials to propaganda outlets	C0190 open engagement with civil society		
C0073 Inoculate populations through media literacy training		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0113 Present sympathetic views of opposite side	C0197 Redirect Method	C0204 Strengthen local media	C0194 Provide an alternative to Russian information by expanding and improving local media		
C0096 Strengthen institutions that are always truth tellers		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0114 Present sympathetic views of opposite side	C0198 Redirect Method	C0205 Respected figure (influencer) disavows misinfo	C0211 Use humorous counter-narrative		
C00153 Take pre-emptive action against actors/ infrastructure		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0115 Present sympathetic views of opposite side	C0199 Redirect Method	C0206 Respected figure (influencer) disavows misinfo	C0212 build public resilience by making civil society more vibrant		
C00159 Have a disinformation response plan		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0116 Present sympathetic views of opposite side	C0200 Redirect Method	C0207 Respected figure (influencer) disavows misinfo	C0213 build public resilience by making civil society more vibrant		
C00161 Coalition Building and Third-party Inducement		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0117 Present sympathetic views of opposite side	C0201 Redirect Method	C0208 Respected figure (influencer) disavows misinfo	C0214 build public resilience by making civil society more vibrant		
C00170 elevate information as a critical domain of statecraft		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0118 Present sympathetic views of opposite side	C0202 Redirect Method	C0209 Respected figure (influencer) disavows misinfo	C0215 build public resilience by making civil society more vibrant		
C00174 Create a healthier news environment		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0119 Present sympathetic views of opposite side	C0203 Redirect Method	C0210 Respected figure (influencer) disavows misinfo	C0216 build public resilience by making civil society more vibrant		
C00176 Improve Coordination amongst stakeholders: public and private		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation	C0120 Present sympathetic views of opposite side	C0204 Redirect Method	C0211 Respected figure (influencer) disavows misinfo	C0217 build public resilience by making civil society more vibrant		
C00205 strong dialogue between the federal government and private sector to encourage better reporting		C00189 Ensure that platforms are taking down flagged accounts	C00197 remove suspicious accounts	C0096 Distract from noise with additive content	C0097 Make more noise than the disinformation						
C00220 Develop a monitoring and intelligence plan						C00202 Set data "honeypots"					
C00221 Run a disinformation red team, and design mitigation factors						C00210 Use encrypted apps for confidential communication					
C00223 Strengthen Trust in social media platforms						C00219 Add metadata to content that's out of the control of disinformation creators					

# Responses

# Raising Adversary costs D-Rail

Linking the cyber kill chain, DISARM blue, EU DisinfoLab's emerging cost-effectiveness evaluation framework

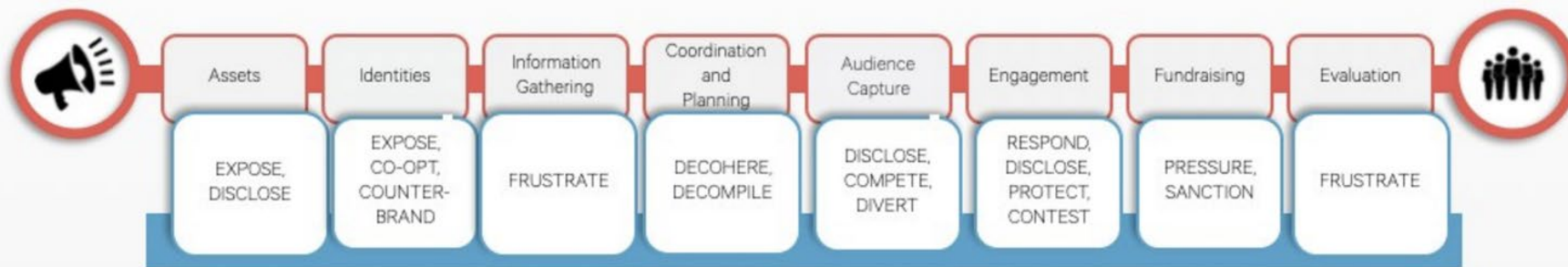




Image: **DISARM Foundation**

T01 Planning	T01 Planning	T01 Preparation	T01 Preparation	T01 Preparation	T01 Preparation	T01 Preparation	T01 Preparation	T01 Execution	T01 Execution	T01 Execution	T01 Execution	T01 Execution
T0101 Strategic Planning	T0101 Objective Planning	T0101 Develop People	T0101 Develop Networks	T0101 Motivating	T0101 Develop Content	T0101 Channel Selection	T0101 Platform Pruning	T0101 Exposure	T0101 Exposure	T0101 Physical	T0101 Persistence	T0101 Evaluation
C00006 Charge for social media	C00009 Educate high profile influencers on best practices	C00034 Create more friction at account creation	C00047 Coordinated inauthenticity	C00065 Reduce political targeting	C00014 Real-time updates to fact-checking database	C00097 Require use of verified identities to contribute to poll or comment	C00112 "Prove they are not an op"	C00089 Throttle number of forwards	C00129 Use banking to cut off access	C00131 Seize and analyse botnet servers		C00090 Fake engagement system
C00008 Create shared fact-checking database	C00011 Media literacy: Games to identify fake news	C00036 Infiltrate the in-group to discredit leaders (divide)	C00052 Infiltrate platforms	C00066 Co-opt a hashtag and drown it out (hijack it back)	C00032 Hijack content and link to truth-based info.	C00098 Revocation of "verified"	C00113 Debunk and deliver a fake expert / credentials. Attack audience quality of fake expert	C00122 Content moderation. Censorship?	C00130 Mentoring elders, youth, credit. Learn vicariously.	C00133 Deplatform Account?		C00140 "Bomb" link shorteners with lots of calls
C00100 Enhanced privacy regulation for social media	C00028 Make information provenance available	C00040 third party verification for people	C00053 Delete old accounts / Remove unused social media accounts	C00216 Use advertiser controls to stem flow of funds to bad actors	C00071 Block source of pollution	C00099 Strengthen verification methods	C00114 Don't engage with payloads	C00123 Bot control		C00135 Deplatform message groups and/or message boards		C00147 Make amplification of social media posts expire (e.g. can't like/repost after x days)
C00102 Platform regulation	C00029 Create fake website to issue counter narrative and counter narrative through physical merchandise	C00042 Address threat contained in narratives	C00056 Get off social media	C00059 Verification of project before posting (counters funding campaigns)	C00072 Content censorship in non-relevant domains e.g. Pinterest antivax	C00100 Hashtag jacking	C00115 Expose actor and intentions	C00124 Don't feed the trolls		C00136 Microtarget most likely targets then send them counter messages		C00148 Add random links to network graphs
C00103 Rating framework for news	C00030 Develop a compelling counter narrative (truth based)	C00044 Keep people from posting to social media immediately		C00059 Verification of project before posting (counters funding campaigns)	C00074 Identify identical content and mass deplatform	C00101 Create participant friction	C00116 Provide proof of involvement	C00125 Prepare the population with pre-announcements		C00137 Pollute the AB-testing data feeds		C00149 Poison the monitoring & evaluation data
C00106 Censorship - not recommended	C00031 Dilute the core narrative - create multiple permutations, target / amplify	C00046 Marginalise and discredit extremist groups		C00062 Free open library sources worldwide	C00075 Normalise language	C00102 Make repeat voting harder	C00117 Downgrade de-amplify label promote counter to disinformation	C00126 Social media amber alert		C00138 Span domestic actors with lawsuits		
C00107 Repair broken social connections	C00060 Legal action against for-profit engagement factories	C00048 Name and Shame Influencers		C00162 collect data/map constellations of Russian "civil society". Unravel/target the Potemkin villages	C00076 Prohibit images in political discourse channels	C00103 Create a bot that engages / distract trolls	C00118 Repurpose images with new text	C00128 Create friction by marking counter with ridicule or other "declarations"		C00139 Weaponise youtube content matrices		
C00109 Reduce effect of disinformation enablers	C00070 Block access to disinformation resources	C00051 Counter social engineering training			C00078 Change Search Algorithms for Disinformation Content	C00105 Buy more advertising than the adversary to shift influence and algorithm	C00119 Engage payload and debunk. Provide link to facts.	C00120 Open dialogue about design of platforms to produce different outcomes	C00151 "Light in the light"	C00143 (botnet) DMCA takedown requests to waste group time		
C00121 Encourage in-person communication	C00092 Reputation scores for social media influencers	C00058 Report crowd/funder as violator			C00080 Create competing narrative	C00106 Click-bait centrist content	C00121 Tool transparency and literacy for channels people follow.	C00158 Use training to build the resilience of at-risk populations.		C00144 Buy out troll farm employees / offer them jobs		
C00122 Inoculate: Positive campaign to promote feeling of safety	C00164 compatriot policy	C00067 Denigrate the recipient/ project (of online funding)			C00081 Highlight flooding and noise, and explain motivations	C00107 Content moderation	C00122 Tool transparency and literacy for channels people follow.	C00159 Use training to build the resilience of at-risk populations.		C00145 Pollute the data voids with wholesome content ("Kitten" Babyshark?)		
C00124 Promote healthy narratives	C00207 Run a competing disinformation campaign - not recommended	C00093 Influencer code of conduct			C00082 Ground truthing as automated response to pollution	C00108 De-escalation	C00123 Tool transparency and literacy for channels people follow.	C00160 Develop a creative content hub				
C00126 Shore up democracy based messages		C00155 Ban incident actors from funding sites			C00084 Modify disinformation narratives, and rebroadcast them	C00110 Monetize centrist SEO by subsidizing the difference in greater clicks towards extremist content	C00124 Tool transparency and literacy for channels people follow.	C00161 Develop a creative content hub		C00147 Full information voids with non-disinformation content		
C00127 Create culture of civility		C00160 find and train influencers			C00085 Mute content	C00111 Present sympathetic views of opposite side	C00125 Tool transparency and literacy for channels people follow.	C00162 Develop a creative content hub		C00148 malware detection/quarantine/deletion		
C00173 Inoculate populations through media literacy training		C00189 Ensure that platforms are taking down flagged accounts			C00086 Distract from noise with additive content	C00112 Present sympathetic views of opposite side	C00203 Stop offering press credentials to propaganda outlets	C00163 Develop a creative content hub		C00149 malware detection/quarantine/deletion		
C00096 Strengthen institutions that are always truth tellers		C00197 remove suspicious accounts			C00087 Make more noise than the disinformation	C00113 Present sympathetic views of opposite side	C00204 Strengthen local media	C00190 open engagement with civil society				
C00153 Take pre-emptive action against actors' infrastructure					C00088 Make more noise than the disinformation	C00114 Real-time updates to fact-checking database		C00194 Provide an alternative to Russian information by expanding and improving local content				
C00159 Have a disinformation response plan					C00089 Honeyopt social community	C00115 Expose actor and intentions		C00200 Respected figure (influencer) disavows misinfo				
C00161 Coalition Building and Third-party Indocent					C00090 Honeyopt social community	C00116 Provide proof of involvement		C00211 Use humorous counter-narrative				
C00170 elevate information as a critical domain of statecraft					C00091 Honeyopt social community	C00117 Downgrade de-amplify label promote counter to disinformation		C00212 build public resilience by making civil society more vibrant				
C00174 Create a healthier news environment					C00092 Reputation scores for social media influencers	C00118 Repurpose images with new text						
C00176 Improve Coordination amongst stakeholders: public and private					C00093 Influencer code of conduct	C00119 Engage payload and debunk. Provide link to facts.						
C00205 strong dialogue between the federal government and private sector to encourage better reporting					C00094 Force full disclosure on corporate sources of research	C00120 Open dialogue about design of platforms to produce different outcomes						
C00220 Develop a monitoring and intelligence plan					C00095 Platform adds warning label and decision point when sharing content	C00121 Tool transparency and literacy for channels people follow.						
C00221 Run a disinformation red team, and design mitigation factors					C00096 Distract from noise with additive content	C00122 Tool transparency and literacy for channels people follow.						
C00223 Strengthen Trust in social media platforms					C00097 Ensure that platforms are taking down flagged accounts	C00123 Tool transparency and literacy for channels people follow.						
					C00098 Make more noise than the disinformation	C00124 Tool transparency and literacy for channels people follow.						
					C00099 Honeyopt social community	C00125 Tool transparency and literacy for channels people follow.						
					C00100 Hashtag jacking	C00126 Social media amber alert						
					C00101 Create participant friction	C00127 Regulate alert when large batches of new URLs get registered together						
					C00102 Make repeat voting harder	C00128 Create friction by marking counter with ridicule or other "declarations"						
					C00103 Create a bot that engages / distract trolls	C00129 Use banking to cut off access						
					C00104 Real-time updates to fact-checking database	C00130 Mentoring elders, youth, credit. Learn vicariously.						
					C00105 Buy more advertising than the adversary to shift influence and algorithm	C00131 Seize and analyse botnet servers						
					C00106 Click-bait centrist content	C00132 Bot control						
					C00107 Content moderation	C00133 Deplatform Account?						
					C00108 De-escalation	C00134 (botnet) DMCA takedown requests to waste group time						
					C00109 De-escalation	C00135 Deplatform message groups and/or message boards						
					C00110 Monetize centrist SEO by subsidizing the difference in greater clicks towards extremist content	C00136 Microtarget most likely targets then send them counter messages						
					C00111 Present sympathetic views of opposite side	C00137 Pollute the AB-testing data feeds						
					C00112 Present sympathetic views of opposite side	C00138 Span domestic actors with lawsuits						
					C00113 Present sympathetic views of opposite side	C00139 Weaponise youtube content matrices						
					C00114 Real-time updates to fact-checking database	C00140 "Bomb" link shorteners with lots of calls						
					C00115 Expose actor and intentions	C00141 "Prove they are not an op"						
					C00116 Provide proof of involvement	C00142 Platform adds warning label and decision point when sharing content						
					C00117 Downgrade de-amplify label promote counter to disinformation	C00143 (botnet) DMCA takedown requests to waste group time						
					C00118 Repurpose images with new text	C00144 Buy out troll farm employees / offer them jobs						
					C00119 Engage payload and debunk. Provide link to facts.	C00145 Pollute the data voids with wholesome content ("Kitten" Babyshark?)						
					C00120 Open dialogue about design of platforms to produce different outcomes							
					C00121 Tool transparency and literacy for channels people follow.							
					C00122 Tool transparency and literacy for channels people follow.							
					C00123 Bot control							
					C00124 Don't feed the trolls							
					C00125 Prepare the population with pre-announcements							
					C00126 Social media amber alert							
					C00127 Regulate alert when large batches of new URLs get registered together							
					C00128 Create friction by marking counter with ridicule or other "declarations"							
					C00129 Use banking to cut off access							
					C00130 Mentoring elders, youth, credit. Learn vicariously.							
					C00131 Seize and analyse botnet servers							
					C00132 Bot control							
					C00133 Deplatform Account?							
					C00134 (botnet) DMCA takedown requests to waste group time							
					C00135 Deplatform message groups and/or message boards							
					C00136 Microtarget most likely targets then send them counter messages							
					C00137 Pollute the AB-testing data feeds							
					C00138 Span domestic actors with lawsuits							
					C00139 Weaponise youtube content matrices							
					C00140 "Bomb" link shorteners with lots of calls							
					C00141 "Prove they are not an op"							
					C00142 Platform adds warning label and decision point when sharing content							
					C00143 (botnet) DMCA takedown requests to waste group time							
					C00144 Buy out troll farm employees / offer them jobs							
					C00145 Pollute the data voids with wholesome content ("Kitten" Babyshark?)							

# EEAS FIMI Toolbox



# Ethical responses: United Nations Global Principles For Information Integrity

**Societal Trust and Resilience-** people need to be able to trust information presented, we need to limit actors operating for strategic, political or financial gain











**Healthy Incentives-** advertisers and platforms need to be driven by business incentives that are not actively harmful (eg. fermenting rage for attention)

**Public Empowerment-**Technology companies should empower users to provide input and feedback on all aspects of trust and safety, privacy policy and data use, recognizing user privacy rights

**Independent, Free and Pluralistic Media-** a free press

**Transparency and Research-** transparency from tech platforms that allows outside researchers to see how data is being used

**Table 1. Overview of Case Studies<sup>1</sup>**

Type	Intervention	How much is known?	How effective does it seem?	How easily does it scale?
	1. Supporting local journalism	Modest	Significant	Difficult
	2. Media literacy education	Significant	Significant	Difficult
	3. Fact-checking	Significant	Modest	Modest
	4. Labeling social media content	Modest	Modest	Easy
	5. Counter-messaging strategies	Modest	Modest	Difficult
	6. Cybersecurity for elections and campaigns	Modest	Modest	Modest
	7. Statecraft, deterrence, and disruption	Modest	Limited	Modest
	8. Removing inauthentic asset networks	Limited	Modest	Modest
	9. Reducing data collection and targeted ads	Modest	Limited	Difficult
	10. Changing recommendation algorithms	Limited	Significant	Modest



Public information



Government action



Platform action



Source  
Funded by  
the European Union

