Defending democratic information in the age of algorithmic manipulation

DI.INTRO.2 Stephen Campbell

EU CYBERNET SUMMER SCHOOL 2025

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

EU CyberNet

Federal Foreign Office

Funded by
the European Union

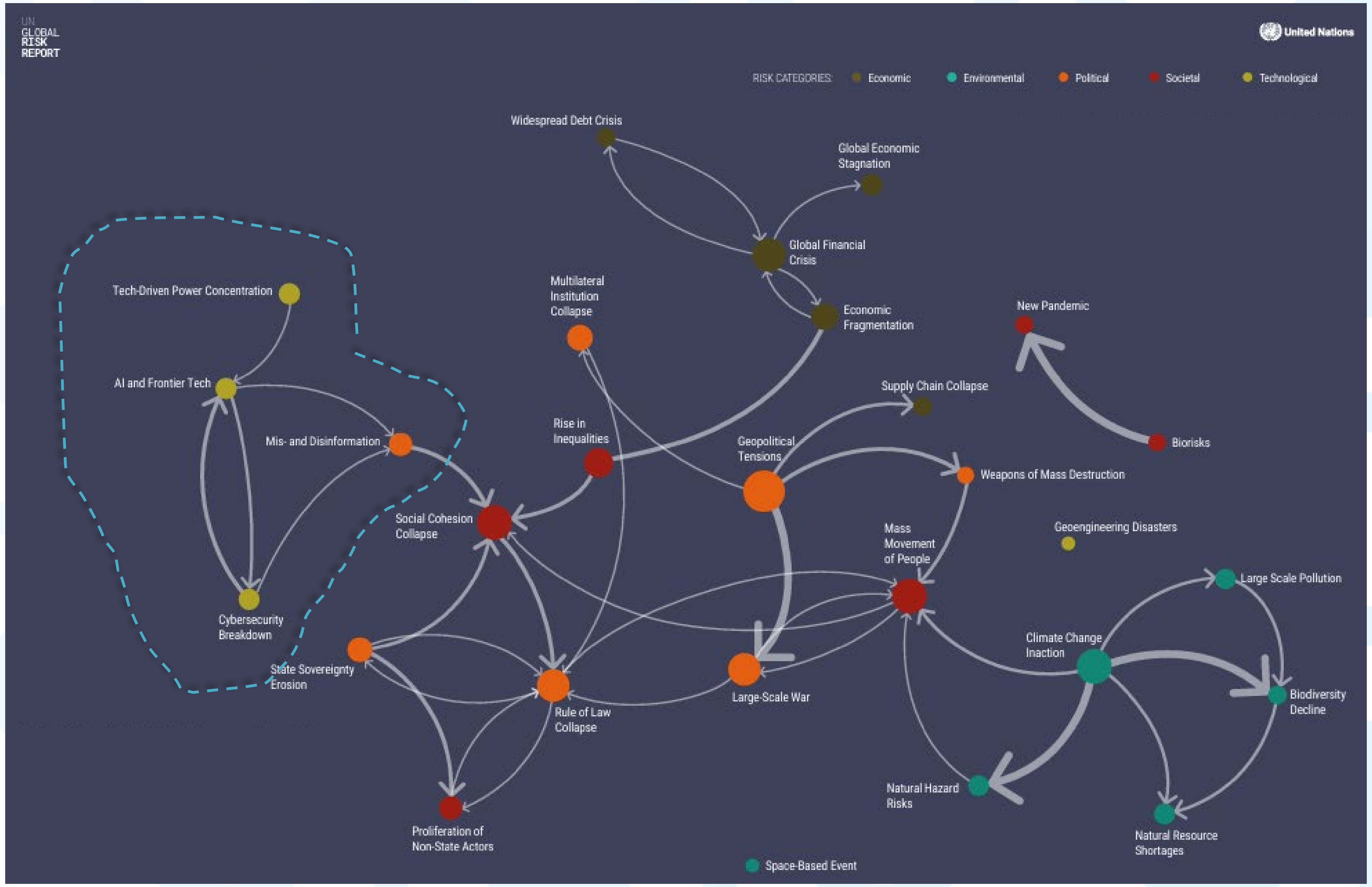
Module Outline

- Information threats during cyber crises
- Definitions and terminology
- Evolution of the information environment
- The impact of artificial intelligence
- How do we respond?

EU CyberNet
Federal Foreign Office
Funded by the European Union

UN Global Risks

- Mis/disinfo is #3 risk
- Cyber, mis/disinfo, and AI are #2, #4, #8 in least prepared risks
- For mis/disinfo biggest obstacles are gaps in data, low accountability and weak communication pathways



UN Global Risk Report 2025

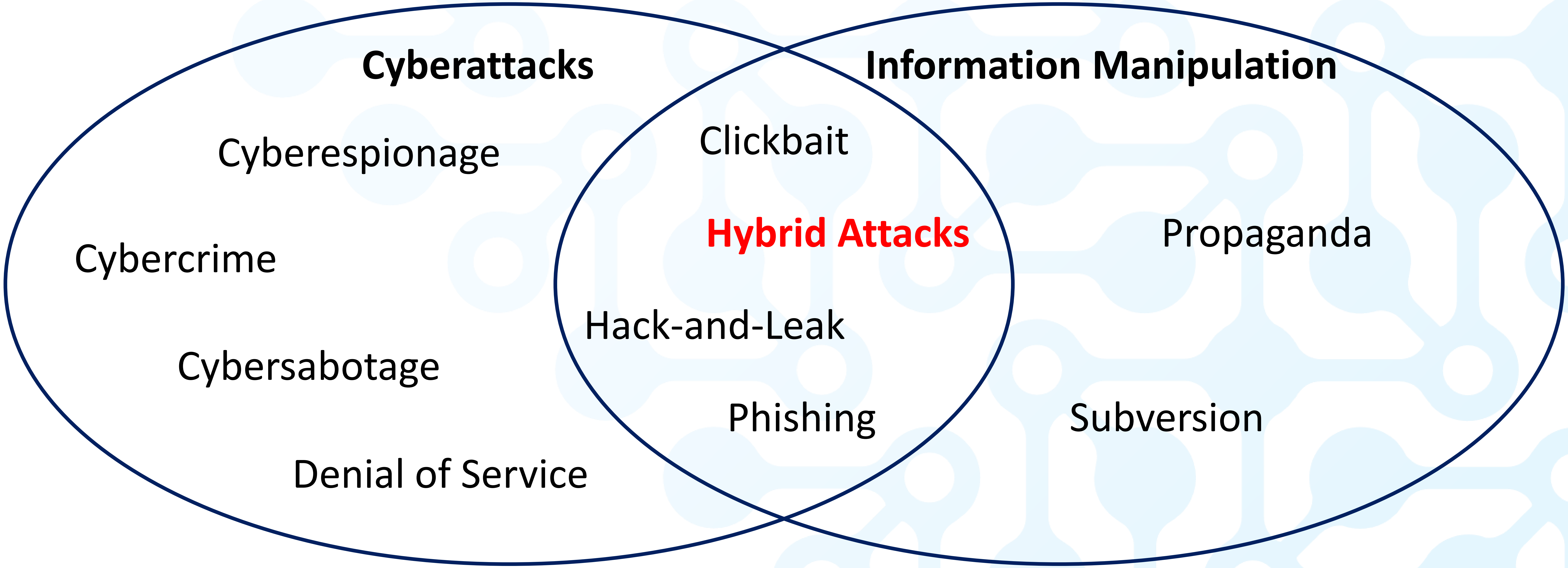
DI.INTRO.2: Defending democratic information in the age of algorithmic manipulation

EU CyberNet

Federal Foreign Office

Funded by the European Union

Threats to the Information Environment



Cyberattacks

Cyberespionage

Cybercrime

Cybersabotage

Denial of Service

Information Manipulation

Propaganda

Subversion

Clickbait

Hybrid Attacks

Hack-and-Leak

Phishing

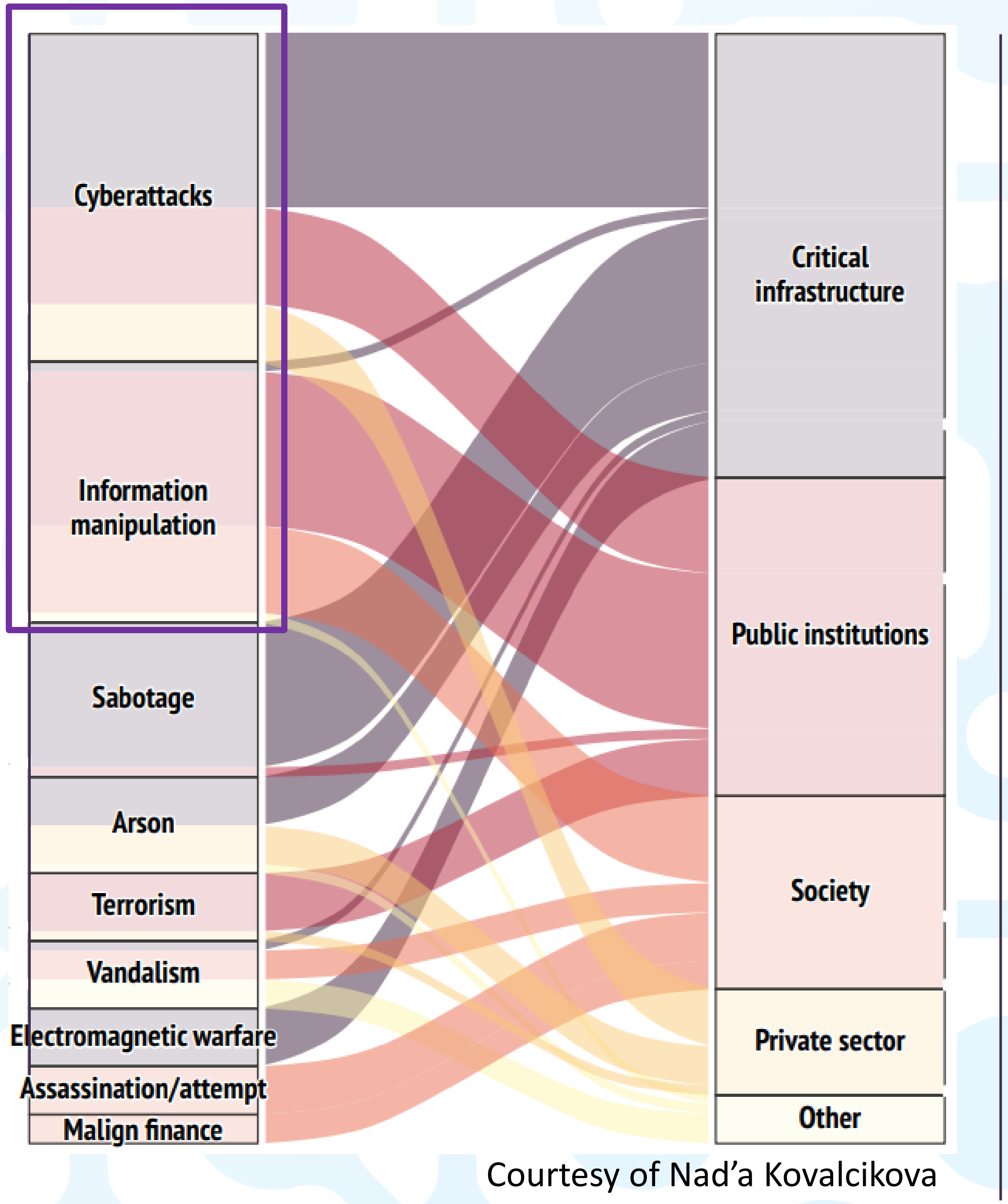
EU CyberNet

Federal Foreign Office

Funded by the European Union

Hybrid Warfare

- Staying below article 5
- Reflexive control
- Exploiting silos



Courtesy of Nad’a Kovalcikova

DI.INTRO.2: Defending democratic information in the age of algorithmic manipulation

EU CyberNet

Federal Foreign Office

Funded by the European Union

Information Campaigns are used to Promote Narratives

DI.INTRO.2: Defending democratic information in the age of algorithmic manipulation

- Reinforcing existing narratives is feasible
- Overturning established narratives is hard
- FIMI succeeds when it aggravates existing problems in the target society or amplifies a narrative that is already shifting

Courtesy of Atlantic Council

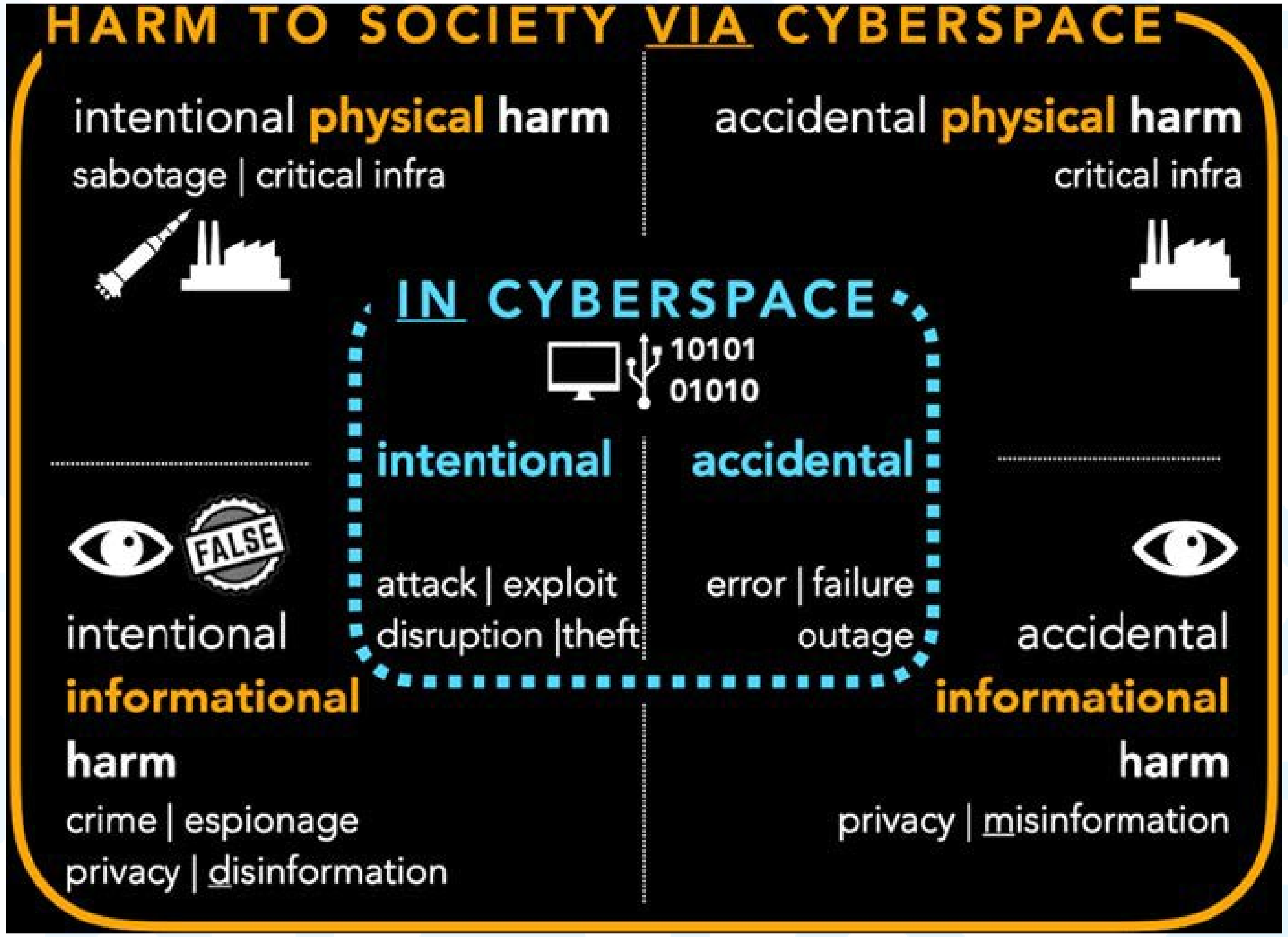
EU CyberNet

Federal Foreign Office

Funded by the European Union

Cyber Crises

- European Definition
  - Impacts security, safety or internal market across multiple member states
- Heightened Tensions
  - Intense public scrutiny and pressure to act
- Information Scarcity
  - Opportunity to challenge settled narratives



HARM TO SOCIETY VIA CYBERSPACE

intentional **physical harm**
sabotage | critical infra

accidental **physical harm**
critical infra

IN CYBERSPACE
10101
01010

intentional
attack | exploit
disruption |theft

accidental
error | failure
outage

intentional
informational
harm
crime | espionage
privacy | disinformation

accidental
informational
harm
privacy | misinformation

Courtesy of Van den Berg and Kuipers

EU CyberNet
Federal Foreign Office
Funded by the European Union

Definitions

- Disinformation
  - *Information that is deliberately false or misleading and intended to deceive*
- Misinformation
  - *Information whose false or misleading nature is unknown to those spreading it*
- Propaganda
  - *Information with a political or ideological agenda*
- Information Campaign
  - *Coordinated effort to manipulate the information environment for a strategic goal*
- Foreign Information Manipulation and Interference
  - *a mostly non-illegal pattern of behaviour in the information domain 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.*

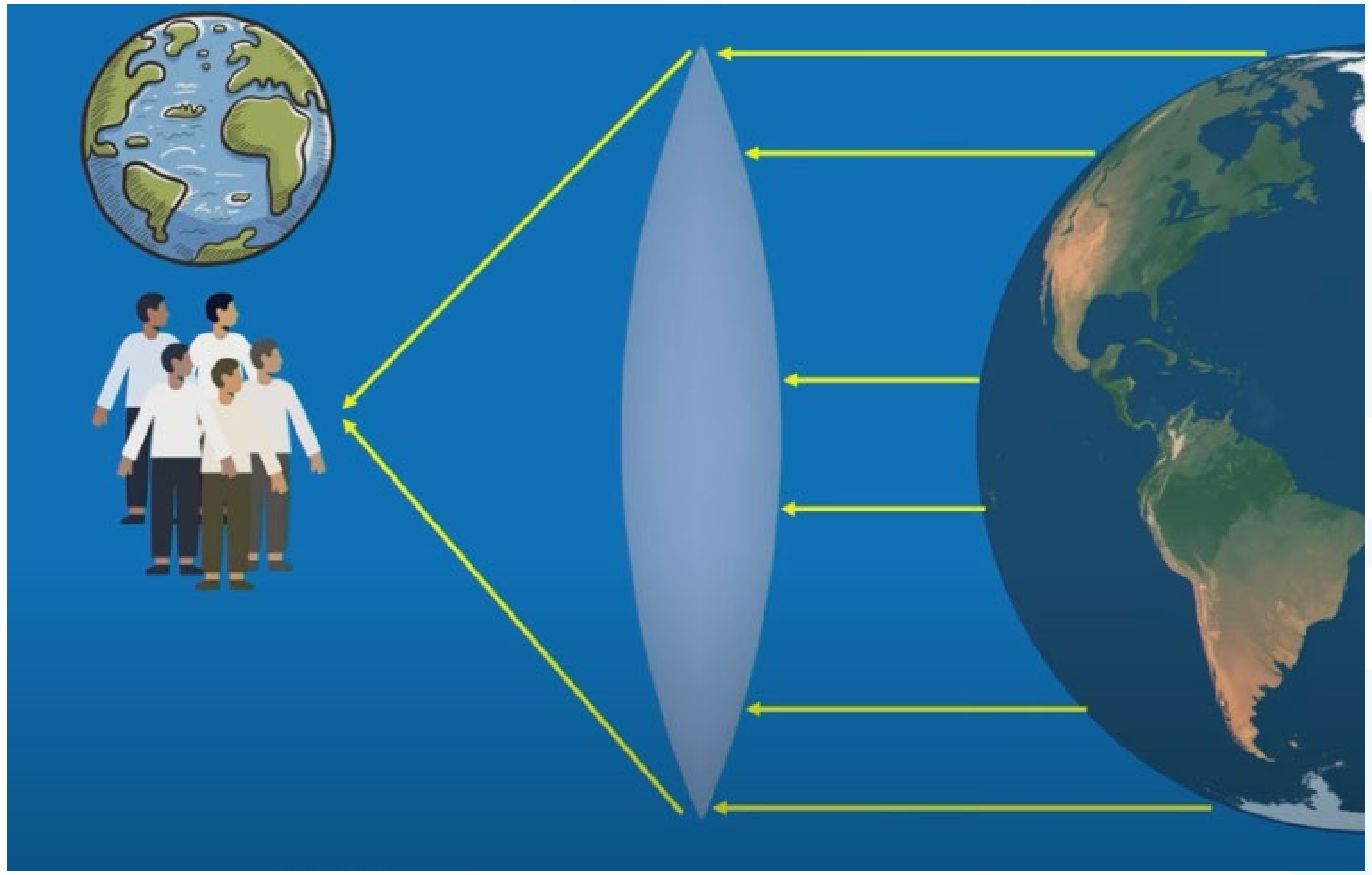
EU CyberNet

Federal Foreign Office

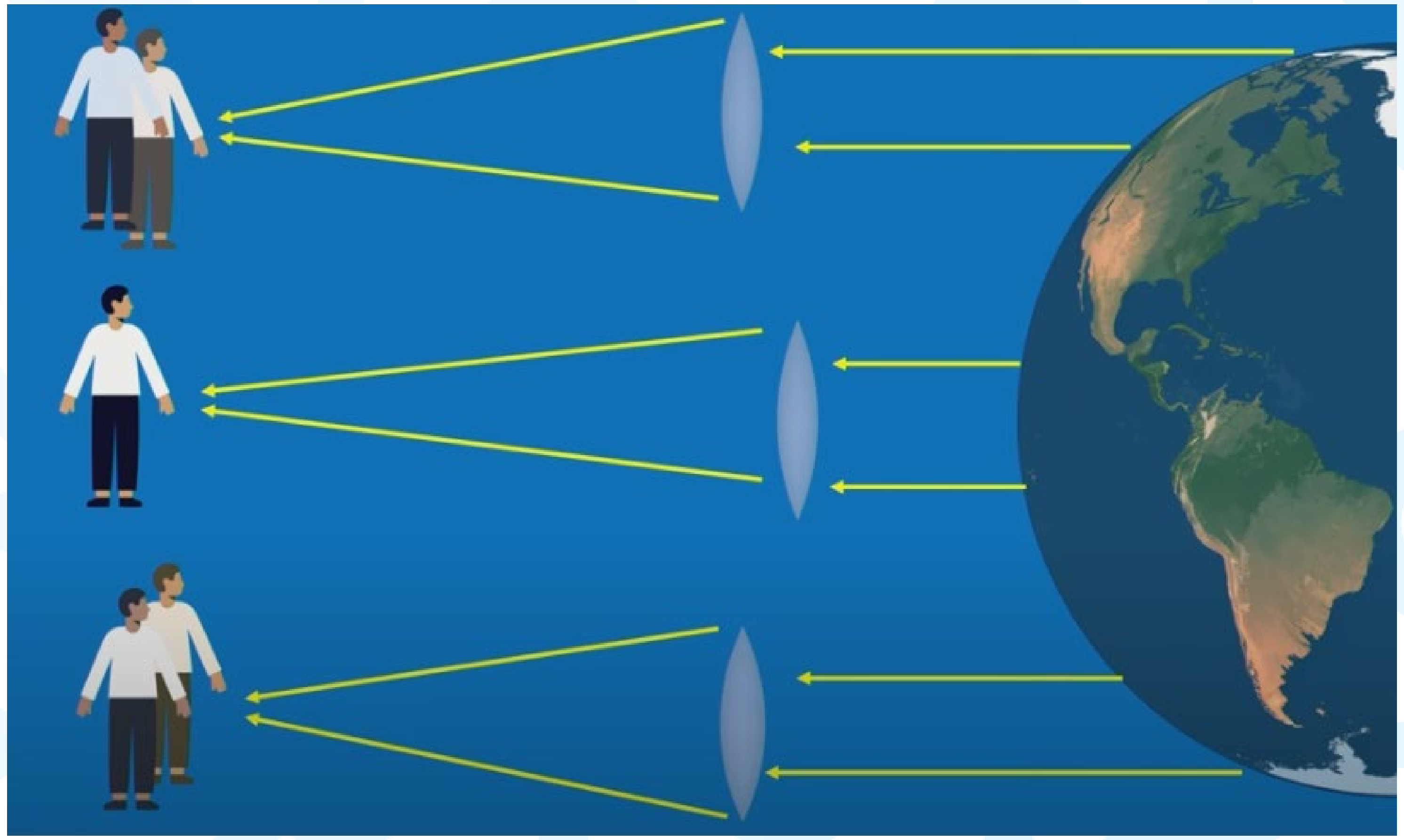
Funded by the European Union

Our Understanding of the World has Splintered

20th Century



21st Century



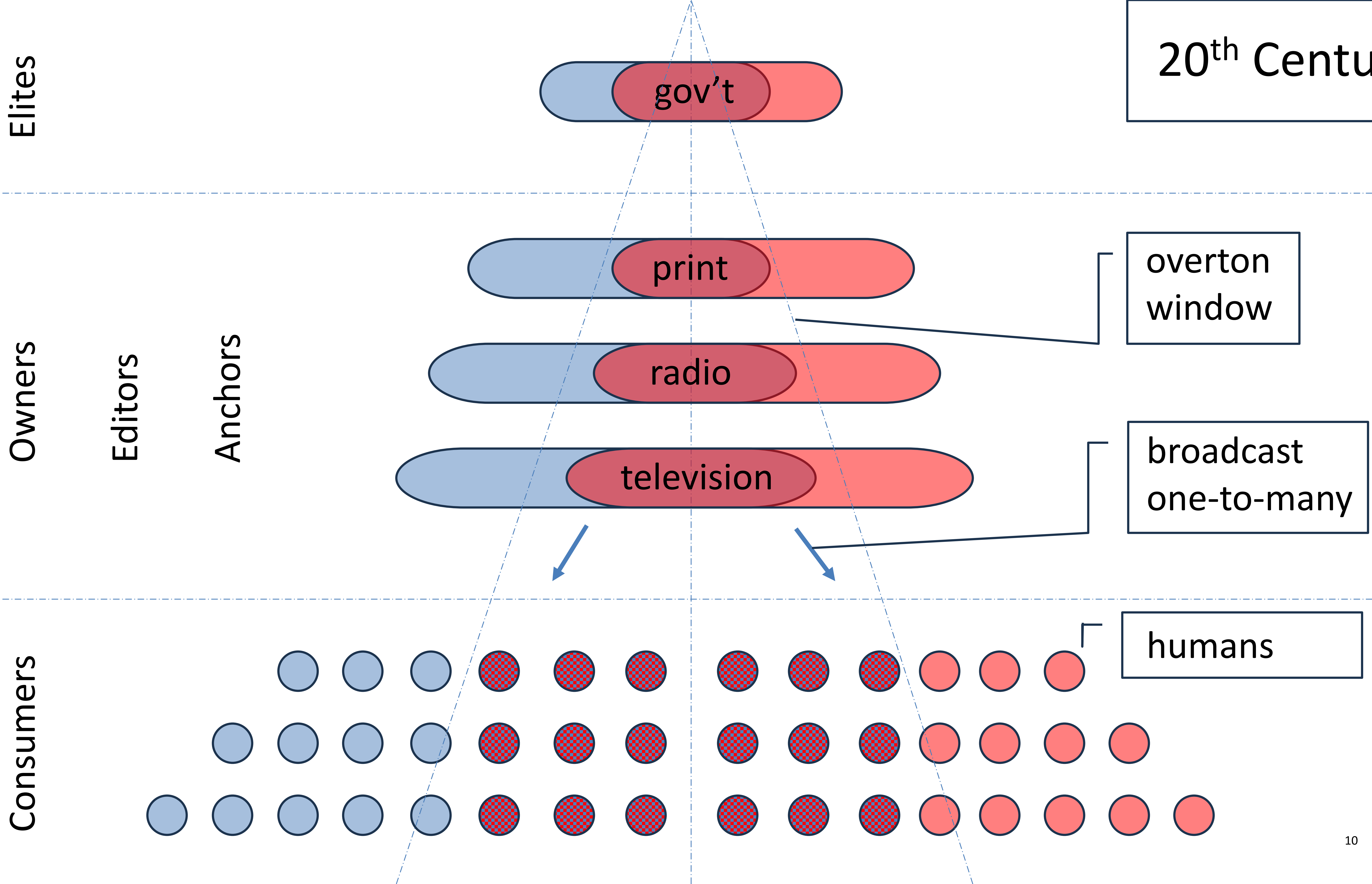
Courtesy of Eliot Higgins, CEO Bellingcat

EU CyberNet

Federal Foreign Office

Funded by the European Union

20th Century



Elites

gov't

Owners

Editors

Anchors

print

radio

television

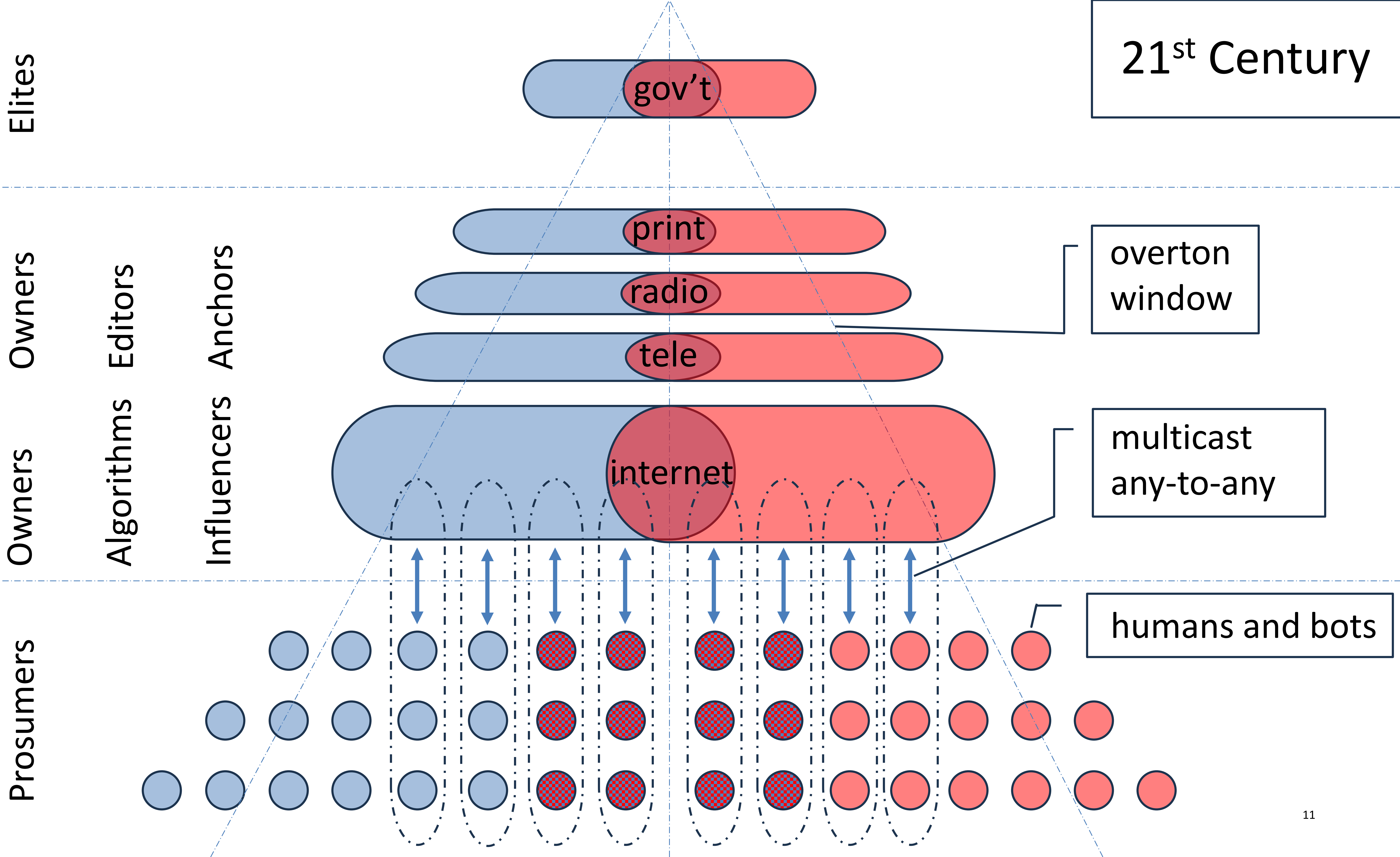
overton window

broadcast one-to-many

Consumers

humans

21st Century

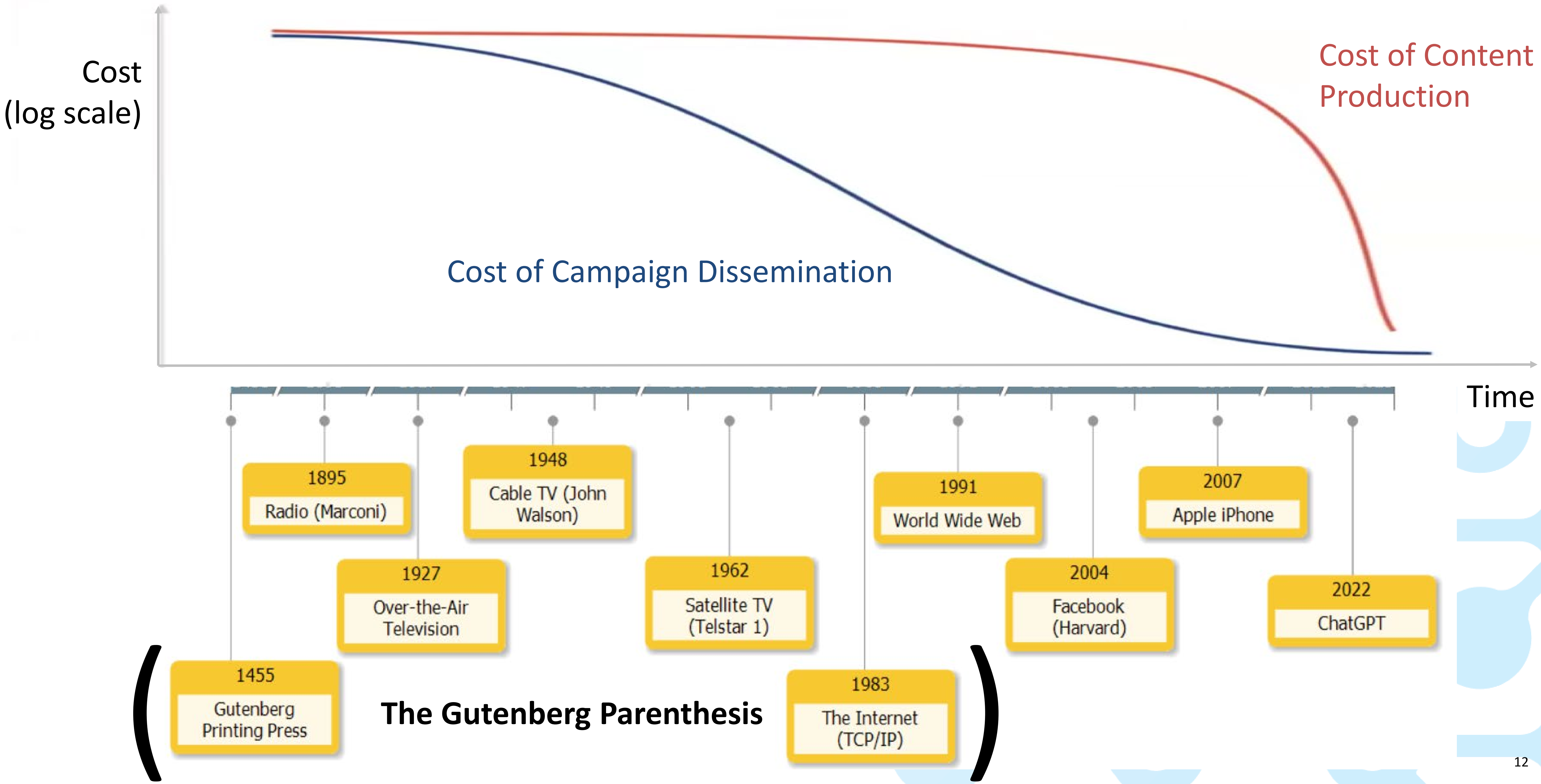


overton window

multicast any-to-any

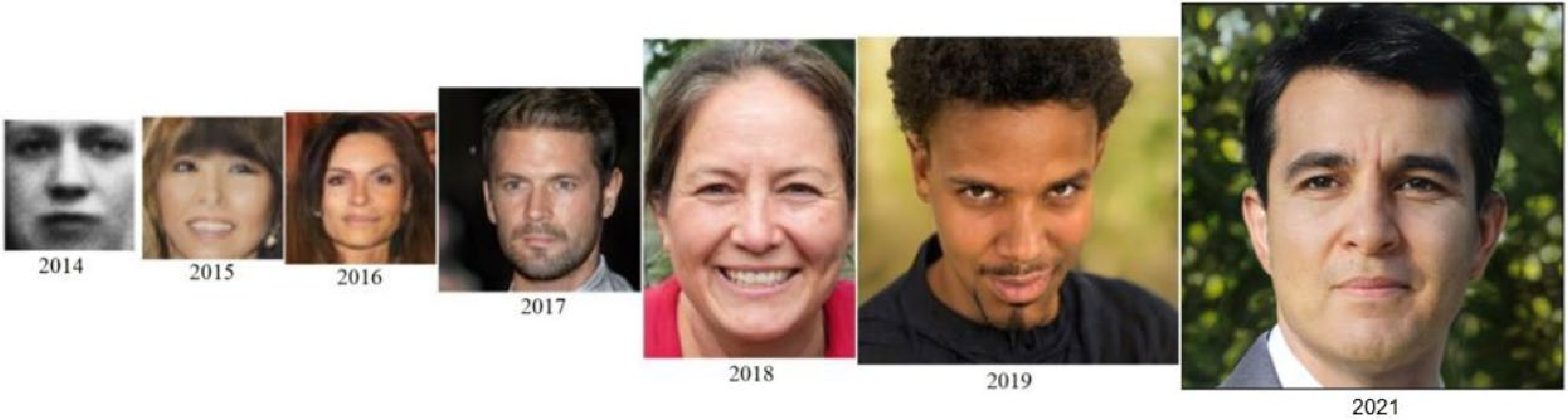
humans and bots

The Cost of Disinformation Campaigns has Plummeted



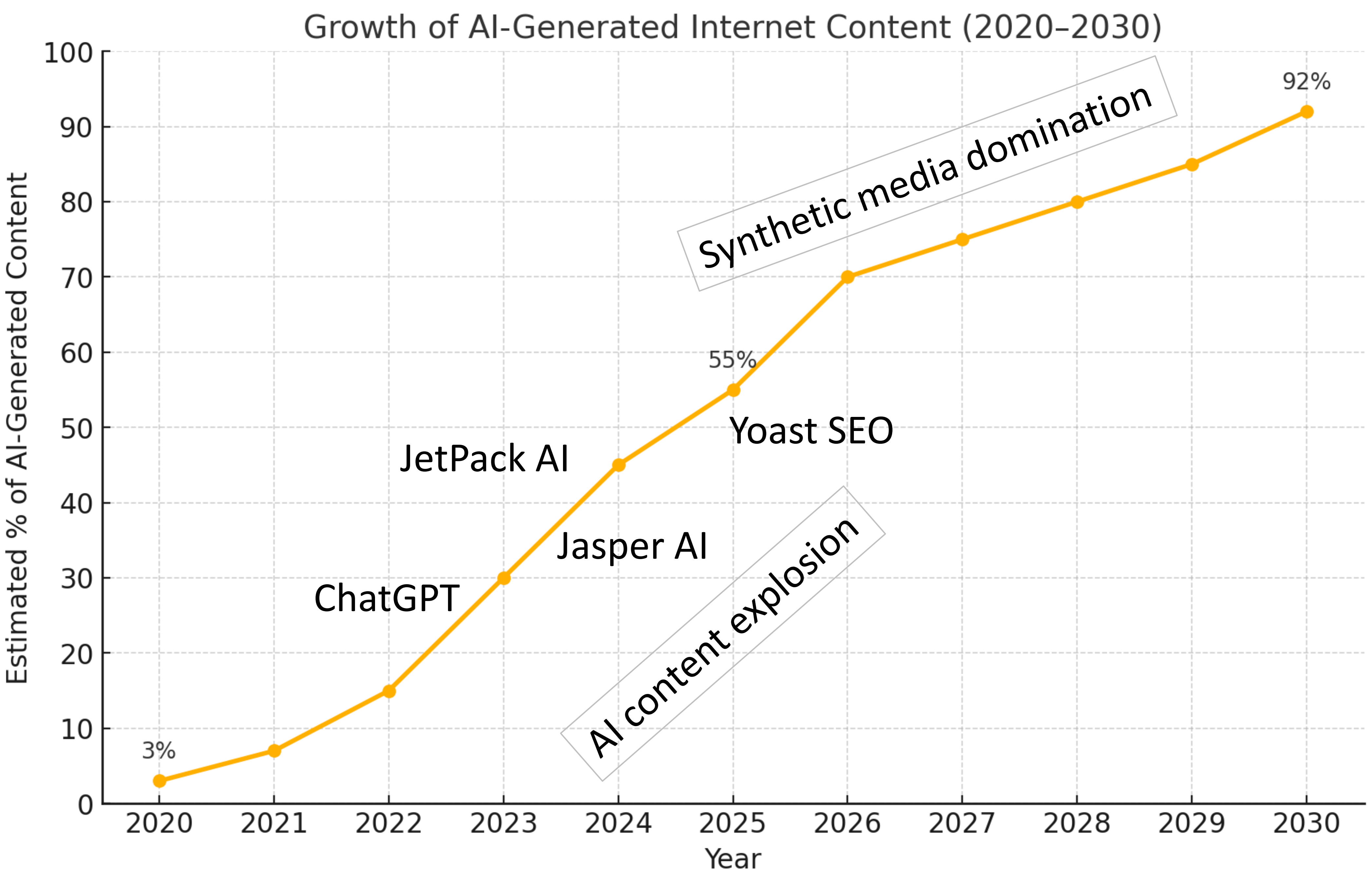
AI Can Produce Hyperrealistic Content

- Between 30% and 50% cannot distinguish real from AI-generated content



When Everything Becomes AI-Generated

- AI-generated content is growing
- Most Content Management Systems now use LLMs
- Will we reach a tipping point? AI slop takes over? Model Collapse?



DI.INTRO.2: Defending democratic information in the age of algorithmic manipulation

DI.INTRO.2: Defending democratic information in the age of algorithmic manipulation

Information Manipulation Requires a Different Response

	Cyberattacks	**Information Manipulation**
Access	Private	Public
Legality	Illegal	Gray Zone
Response	Block	Constrained
Responder	Victim	Whole-of-Society
Attribution	Desirable	Important

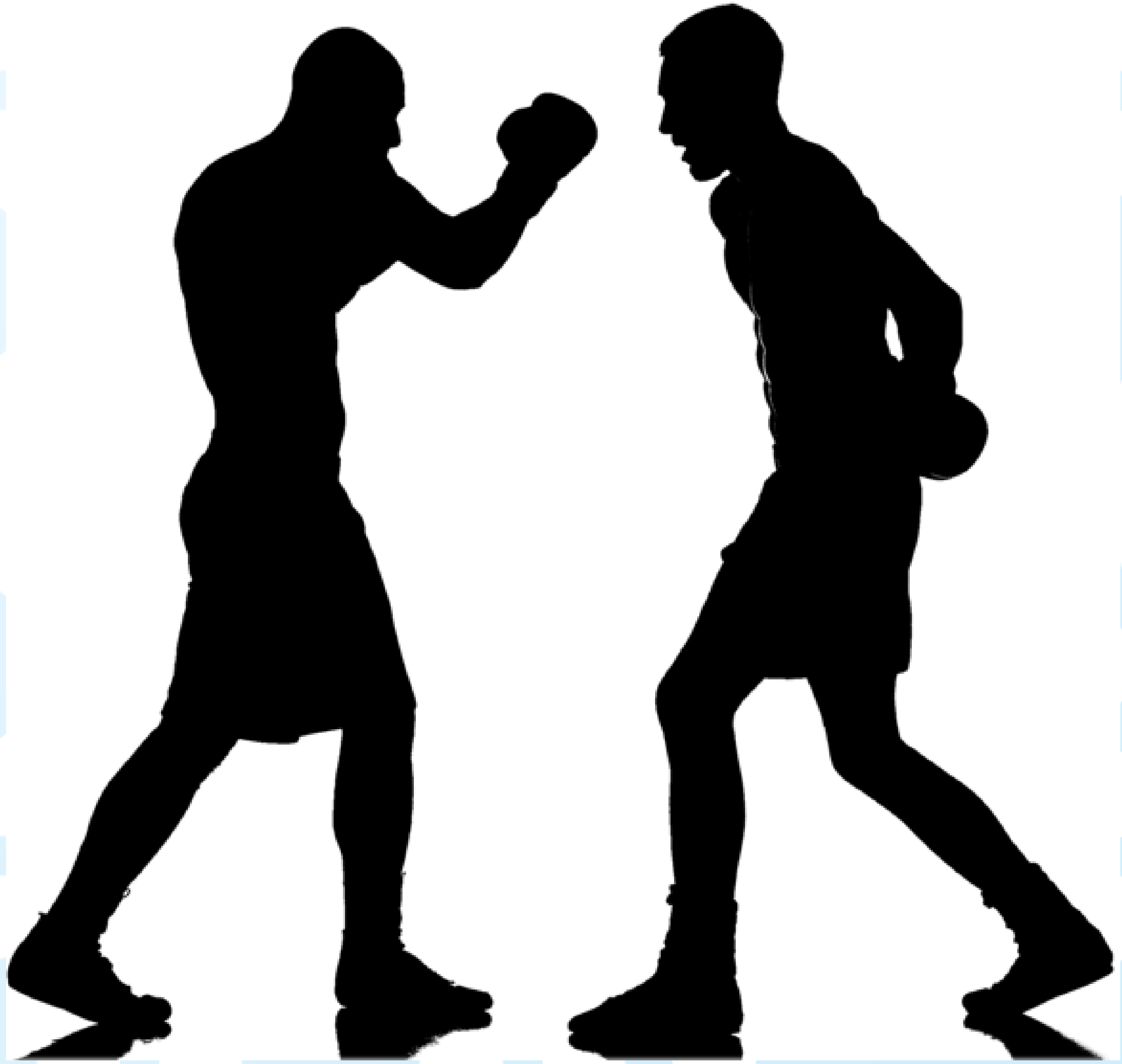
EU CyberNet

Federal Foreign Office

Funded by the European Union

So How Should We Respond?

- Shared Situational Awareness
  - *Structured information sharing*
- Unified Communications
  - *Stratcom planning*
- Fight Back
  - *Cyber Diplomacy, FIMI and Hybrid Toolboxes*



EU CyberNet

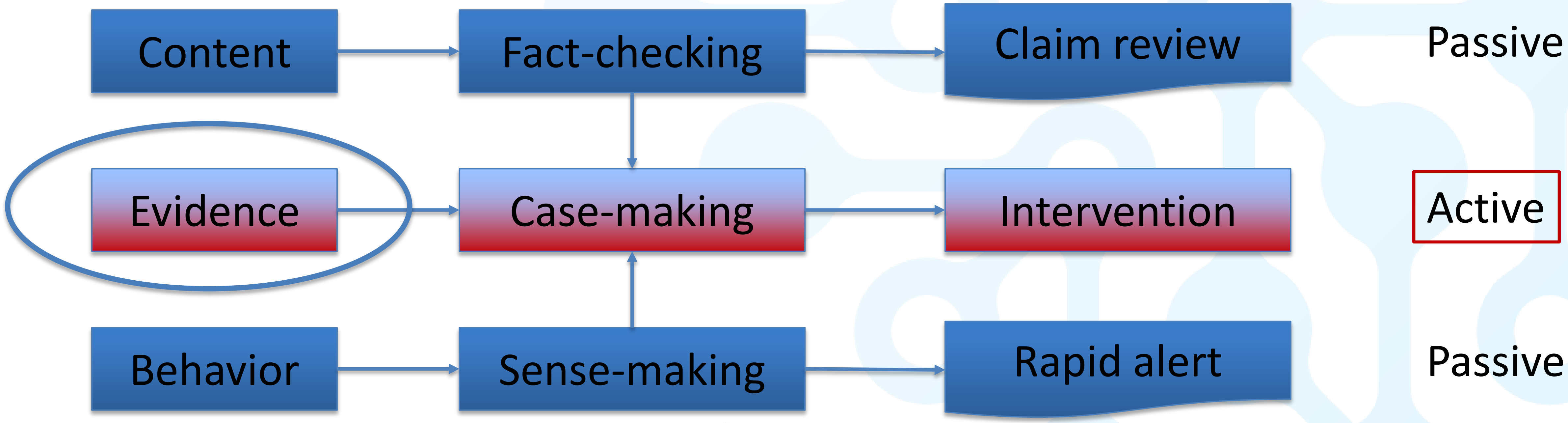
Federal Foreign Office

Funded by the European Union

BACKUP SLIDES

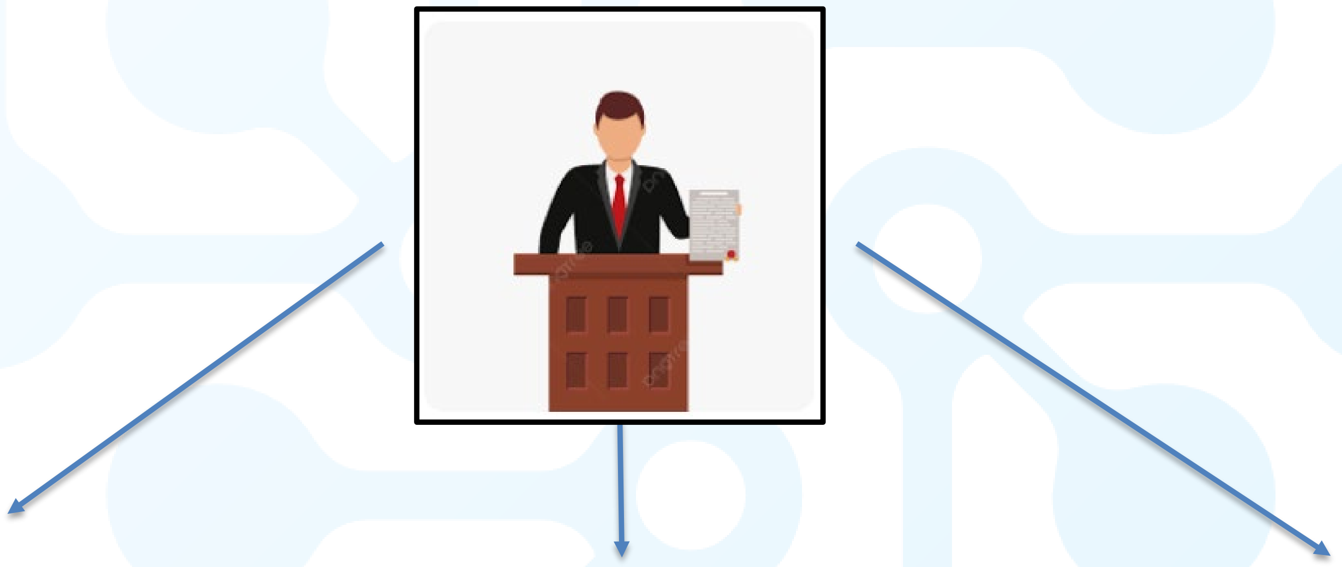
Defender Processes

Question	Focus	Process
Is that true?	Content	Fact-Checking
What’s going on?	Behavior	Sense-Making
Can you prove it?	Evidence	Case-Making



EU CyberNet
Federal Foreign Office
Funded by the European Union

Cyber	FIMI
illegal	non-illegal
private	public
no rights	free speech



takedown

indictment

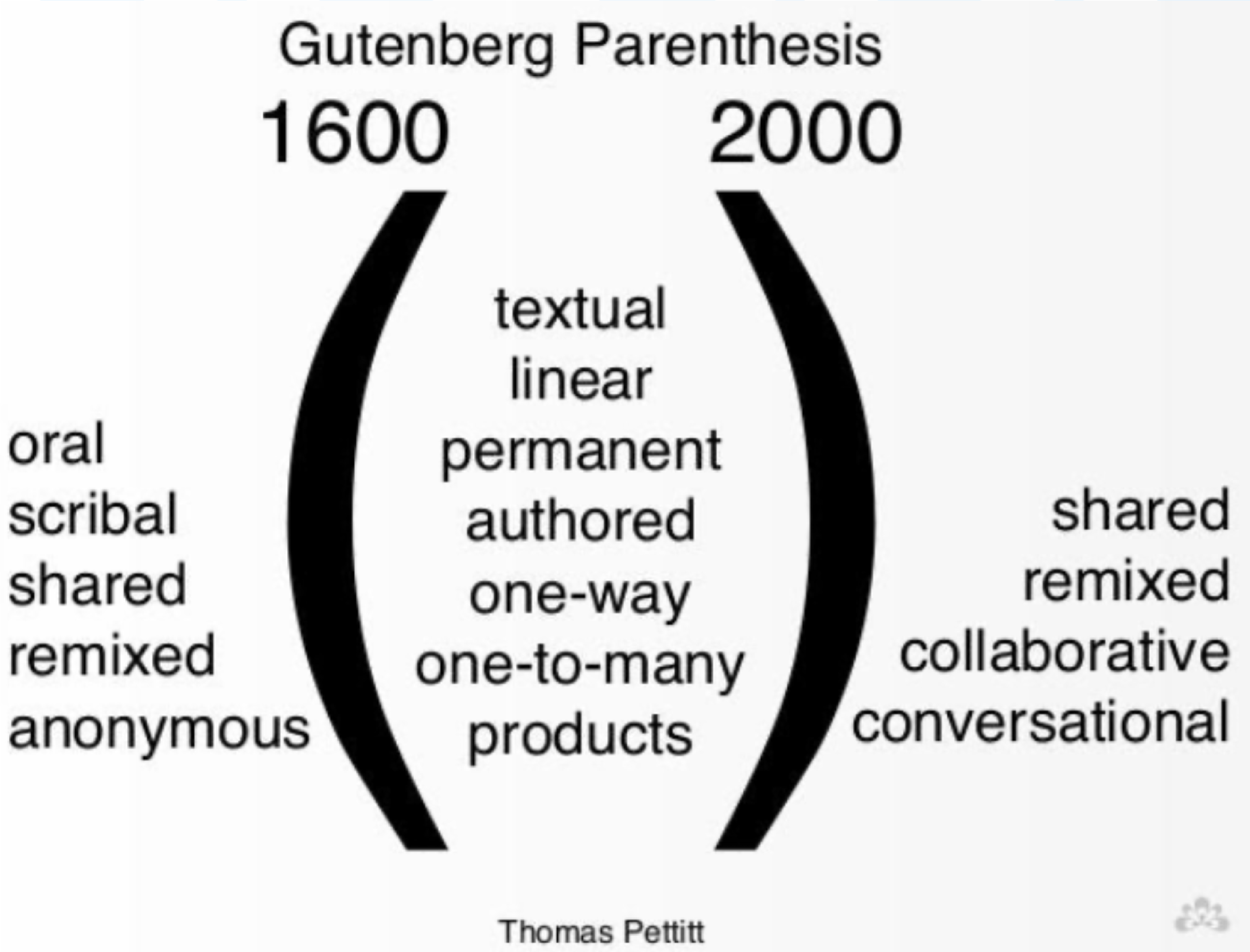
naming & shaming

EU CyberNet

Federal Foreign Office

Funded by the European Union

Learning to Talk to Each Other Again



Gutenberg Parenthesis

1600 2000

oral
scribal
shared
remixed
anonymous

textual
linear
permanent
authored
one-way
one-to-many
products

shared
remixed
collaborative
conversational

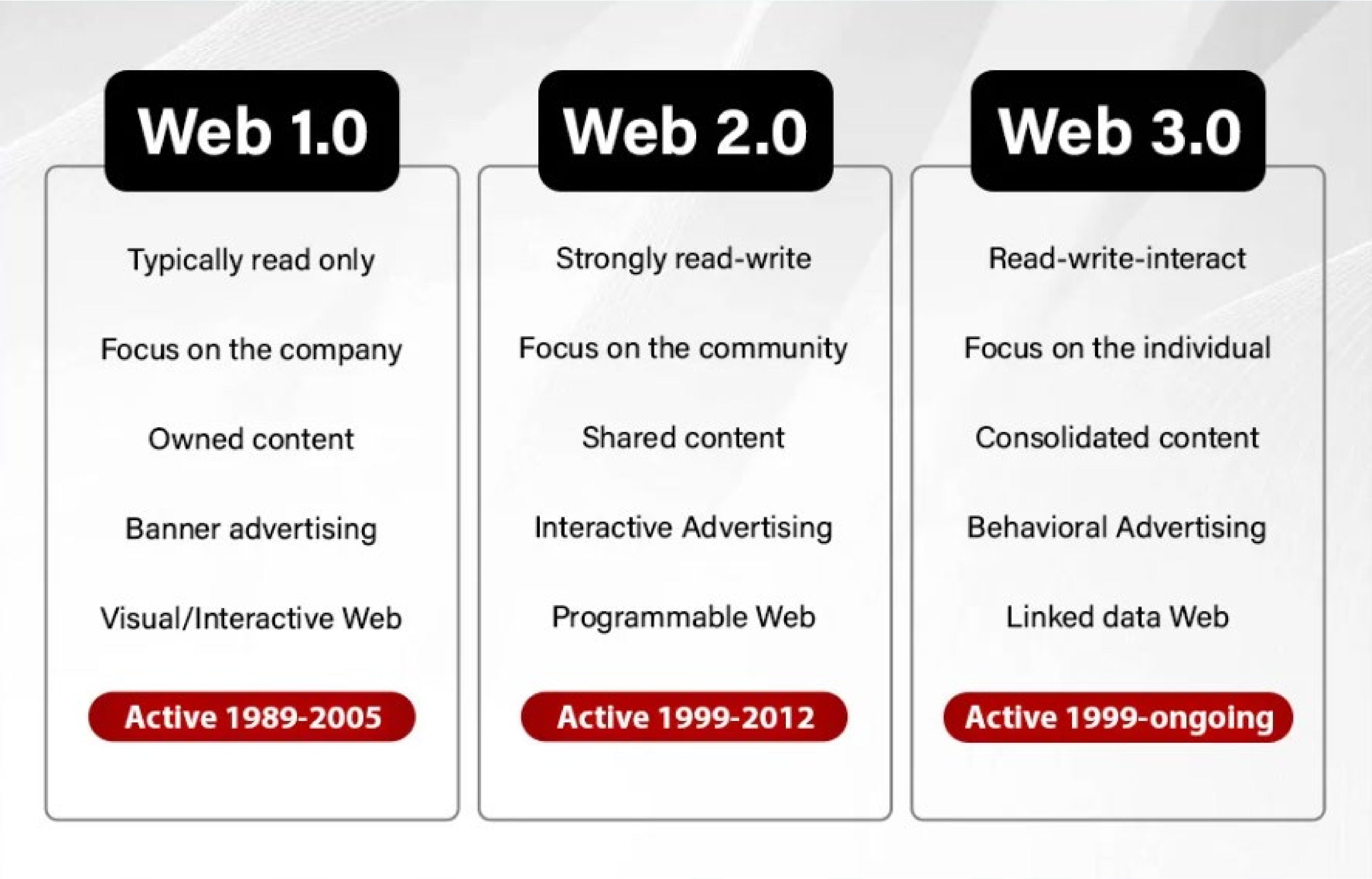
Thomas Pettitt

EU CyberNet

Federal Foreign Office

Funded by the European Union

When Everyone is a Publisher



Web 1.0	Web 2.0	Web 3.0
Typically read only	Strongly read-write	Read-write-interact
Focus on the company	Focus on the community	Focus on the individual
Owned content	Shared content	Consolidated content
Banner advertising	Interactive Advertising	Behavioral Advertising
Visual/Interactive Web	Programmable Web	Linked data Web
Active 1989-2005	Active 1999-2012	Active 1999-ongoing

EU CyberNet

Federal Foreign Office

Funded by the European Union