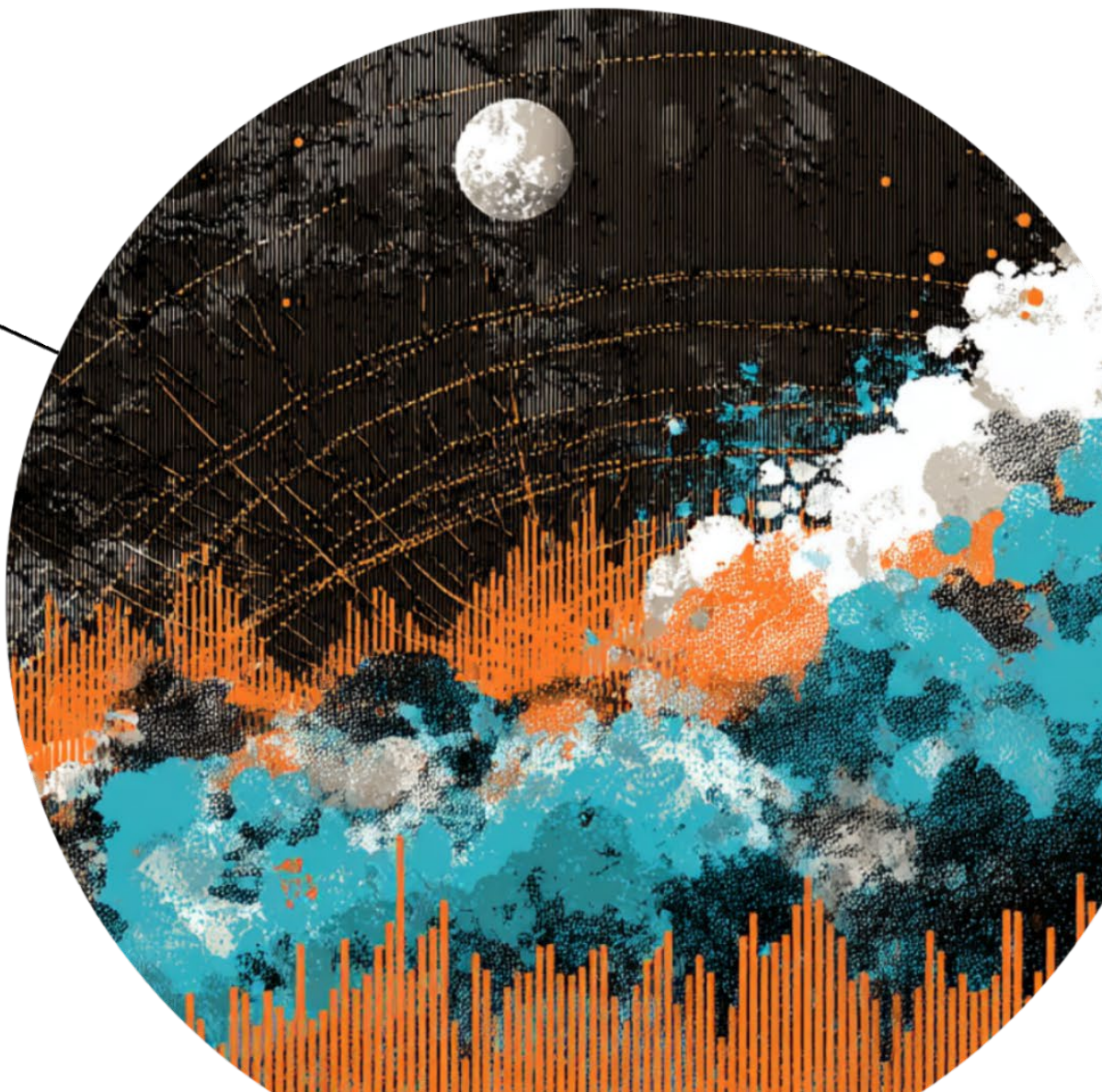


EU DISINFO LAB



DECODING DISINFORMATION IMPACT FRAMEWORKS AND INDICATORS

a comparative study



Decoding disinformation impact frameworks and indicators: a comparative study

by EU DisinfoLab

Introduction

In recent years, there have been numerous advances in research into the phenomenon of disinformation: the community of defenders has harmonised study frameworks, joined forces to increase research, and shed much light on the phenomenon. However, there are still grey areas that continue to raise questions, but also opportunities for further research. Assessing the effective or potential impact of disinformation remains one of the biggest challenges in the field: what counts as ‘impact’, and should it be judged by online reach, real-world outcomes, or shifts in attitudes and behaviour?

In response to this challenge, EU DisinfoLab [developed](#) - and has just [updated](#) - an Impact-Risk Index specifying indicators to measure the impact of individual hoaxes, intended to complement other methodologies, such as the [Breakout Scale](#), designed for influence operations. Although these are two specific frameworks for studying impact, they are not the only ones that address it. Other methodologies developed for the general study of disinformation also touch on the issue.

The aim of this publication is to compile how impact is addressed and measured by a non-exhaustive set of different frameworks, in order to identify common and divergent aspects, as well as highlight gaps. This will contribute to harmonisation within the community while opening the door to discussions and further efforts to enlighten a grey area.

Selected frameworks:

Comprehensive disinformation-lifecycle frameworks:

1. [ABCDE Framework](#)
2. [Disarm Red Framework/Disarm Blue Framework](#)
3. [CIB Detection Tree](#)

Impact-focused frameworks:

1. [Breakout Scale](#) (Ben Nimmo)
2. [Impact-Risk Index](#)¹
3. [Response-Impact Framework](#)

Understanding of impact: terminology and emphasis

Terminology matters: frameworks used to analyse disinformation or influence campaigns often tackle impact, but with different terms such as ‘effect’, ‘degree’, ‘effectiveness’, ‘changes’, and ‘impact’ – each emphasising slightly different aspects: the **reach** disinformation can achieve and the **harm** it can cause.

While the ABCDE and Disarm frameworks adopt a relatively balanced approach, considering both dimensions, the CIB Detection Tree and the Breakout Scale place greater weight on reach, though the latter does so with a notable scepticism toward metrics. The Impact-Risk Index is heavily focused on reach and engagement indicators, whereas the Response-Impact Framework offers a distinct perspective by looking at the effectiveness of countermeasures.

These definitions shape the indicators used to guide measurement. Prioritising reach implies focusing on audience reach – online and offline – while addressing harm requires emphasising online actions and shifts in public opinion, attitudes, and behaviours.

¹ Thanks to Check First, a user-friendly [Impact-Risk calculator](#) is freely available.

FRAMEWORKS		TERMINOLOGY	REACH v. HARM	QUALITATIVE INDICATORS	QUANTITATIVE INDICATORS	DESCRIPTION
Comprehensive frameworks	ABCDE Framework	Degree	Reach	X	X	Cross-platform spread, media pickup, and multilingual amplification
		Effect	Harm		X	Polarisation, discrediting institutions, risks to public health and safety, threats to fundamental freedoms, and security risks
	Disarm Framework	Effectiveness	Reach	X		Monitor and evaluate message reach/social media engagement
		Changes	Harm		X	Changes in audience knowledge, attitudes, or behaviour
	CIB Detection Tree	Impact	Reach	X		Outreach and interaction metrics
			Harm (less weight)		X	Effect on targets and polarisation
Impact-focused frameworks	Breakout Scale	Impact	Reach	X	X	Outreach on social media, mainstream media, and in real life (through communities) Metrics as an approximation
			Harm		X	Policy response, concrete action, or call for violence
	Impact-Risk Index	Impact	Reach	X	X	Virality, engagement, language, format, media, spreaders
			Harm		X	Call to action, offline effects
	Response-Impact Framework	Countermeasures impact	Reach	X		Amplification mitigation metrics
			Harm		X	Harm mitigation metrics: public awareness, attribution, action, and resilience

Table 1. How frameworks define and address disinformation impact

How to measure impact: Qualitative v. quantitative techniques and indicators

To measure these elements, most frameworks prioritise quantitative methods, as qualitative techniques are harder to apply consistently, despite the need for both. Different indicators are used across frameworks to measure impact, as summarised in Table 2.

Quantitative metrics to measure reach:

- When measuring platform amplification, the approach and level of detail vary. The ABCDE and Response-Impact Frameworks briefly mention cross-platform amplification, while others – such as Disarm, the CIB Detection Tree, and the Impact-Risk Index – offer more detailed insights, including specific KPIs on reach and engagement. Some frameworks focus on only one aspect, such as the Breakout Scale, which refers solely to the number of platforms involved.

Qualitative elements to measure reach and harm:

- Content itself is rarely examined in depth. However, a few frameworks like ABCDE highlight the importance of language, while the CIB Detection Tree and Impact-Risk Index go further by identifying content format as an amplification factor, suggesting that a greater format variety increases the opportunity for reach and impact.
- Media amplification is addressed by nearly all frameworks, typically assessing whether disinformation remains within fringe outlets or spills into mainstream media.
- The role of human amplifiers is explicitly recognised in the CIB Detection Tree, Breakout Scale, and Impact-Risk Index. The latter also highlights recurrent disinformers as a key amplification factor and, in its latest version, acknowledges the role of AI influencers or undisclosed AI personas.
- The Breakout Scale introduces a unique indicator: the number of distinct communities reached by a piece of content.
- Regarding psychological and social indicators, most frameworks recognise risks such as shifts in beliefs and attitudes, institutional distrust, polarisation, and threats to freedom of thought and expression. Frameworks like ABCDE and the CIB Detection Tree provide more granular analysis of these effects.

- Offline actions are referenced across all frameworks, though some distinguish between specific risks (e.g., threats to public health or safety) and broader societal consequences. Calls to action are considered key indicators in both the Impact-Risk Index and the Breakout Scale.

Technological considerations

Although most frameworks are technology-neutral, some acknowledge the role of technological advances in amplifying impact. The latest update to the Impact-Risk Index examines AI's role in disinformation production and dissemination – deepfake creation, large-scale translation, AI-operated accounts and AI personas/influencers, and real-time, AI-generated responses to stimulate engagement – as vectors of amplified impact.

Turning metrics around: assessing counter-measure impact as a positive approach

- Most frameworks treat high metrics as negative (signalling disinformation success), but the Response-Impact Framework adopts a positive approach by reflecting the beneficial effects of countermeasures. These include mental and behavioural changes such as greater public awareness, increased resilience, and strengthened capacity within the defender community.

	PLATFORM-RELATED INDICATORS			CONTENT-RELATED INDICATORS		MEDIA	NUMBER OF COMMUNITIES	TYPE OF AMPLIFIER			BELIF CHANGE INDICATORS		BEHAVIOURAL CHANGE INDICATORS	
	Number of platforms	Reach (views)	Engagement (shares, reactions)	Language	Format	Fringe or mainstream media		Public figures	Recurrent disinformers	AI influencers or personas	Sentiment analysis indicating polarisation	Limitation of free speech	Public health or safety risks	Limitation of free speech
ABCDE Framework	X	X	X	X							X	X	X	X
Disarm Red/Blue Framework	X	X	X								X	X	X	X
CIB Detection Tree	X	X	X	X	X	X		X	X		X	X		X
Breakout Scale	X					X	X	X					X	X
Impact-Risk Index	X	X	X	X	X	X		X	X	X			X	X
Response-Impact framework	X	X	X			X					*	*	*	*
*Positive outcome														

Table 2 Detailed indicators of impact in the different frameworks

Gaps and challenges

FRAMEWORK	SPECULATION LEVEL			ALERT SYSTEM
	LOW	MEDIUM	HIGH	
ABCDE Framework		X		
Disarm Red/Blue Framework		X		
CIB Detection Tree		X		
Breakout Scale				X
Impact-Risk Index	X			X
Response-Impact Framework	X		X	

Table 3. Frameworks' speculation level and alert systems

- While all frameworks recommend impact assessment, only the Breakout Scale and the Impact-Risk Index provide clear, actionable guidance on how to conduct these measurements. These two offer specific indicators, scoring systems, and thresholds, whereas the others remain more general or conceptual, leaving methodological implementation to the user's discretion.
- Despite mention in many methodologies, social and psychological dimensions remain underexplored and underdeveloped. They are addressed only briefly and conceptually, with no clear methodologies or practical guidelines for studying or measuring them effectively. This highlights a broader gap in integrating these considerations across existing frameworks.
- Given the inherent difficulty in establishing a direct cause-and-effect relationship between a disinformation campaign and its impact, all the frameworks reviewed involve some speculation. However, some apply precise, cautious methodologies, while others rely on broader, more interpretative approaches, resulting in inconsistent reliability when drawing conclusions about impact.

- At the same time, most frameworks lack built-in mechanisms to flag high-risk cases, which is crucial to guide prioritisation and allocate resources efficiently. Only two – the Breakout Scale and the Impact-Risk Index – include implicit warning systems to help focus efforts where impact/risk is highest.
- Most frameworks blur the line between effective and potential impact, with a few exceptions (e.g., scales that treat social metrics as proxies, or indices focused on potential influence). “Effective impact” refers to offline consequences – actions or attitude changes linked to false narratives – which is methodologically demanding and often requires establishing causality through human-subject research. “Potential impact” infers risk from digital traces such as reach, engagement, and cross-platform, or cross-language spread. The broader and more engaged the circulation, the higher the threat.
- Since most of the frameworks just focus on one side of the coin – i.e., the impact of the malicious campaigns – a major gap exists in evaluating the impact of countermeasures to disinformation and influence operations. Impact should also be measured positively, by the success of countermeasures. A more integrated approach is needed, one that includes both the attacker’s perspective and the outcomes of countermeasures of actions taken by the defenders’ community.
- In conclusion, understanding how frameworks measure impact is fundamental to filling gaps and envisioning opportunities. Current methodologies show progress but remain fragmented. Greater alignment on terminology, stronger integration of qualitative and psychological factors, improved alert mechanisms, and a clearer emphasis on countermeasures are all necessary to advance impact assessment in disinformation and influence operations.



This publication has been made in the framework of the [Vera AI project](#). Vera.ai is co-funded by the European Commission under grant agreement ID 101070093, and the UK and Swiss authorities. This website reflects the views of the vera.ai consortium and respective contributors. The EU cannot be held responsible for any use which may be made of the information contained herein.