The Four Workstreams for Leaders

Building Solid Resistance to a Cyber-AI Crisis

Juin 2026

This document is not a crisis management guide, but a thinking tool — built around one question: "What should organisations do to prepare effectively against a cyber-AI crisis?"

Waiting for the context to stabilise is a pure illusion. Geopolitical tensions are mounting, technological evolution is unbridled, the threat landscape is constantly shifting, regulation grows ever stricter with divergent interpretations between states, and the very notion of proof is collapsing.

In such conditions, improvising or reacting on instinct is simply not an option — not in an interconnected, globalised world riddled with interdependencies we do not fully control.

These conditions create a fog in which reactions are difficult to anticipate. The only solution, therefore, is to focus on what we can change and influence within our own ecosystem.

Four workstreams are essential.

- Building solid cyber-AI crisis communication
- Managing your informational attack surface
- Reducing organisational vulnerabilities
- Revisiting leadership posture

Maryse Rebillot

Strategic Advisor on Cyber-AI crises and information integrity

Building **solid** cyber-AI crisis communication

A cyberattack is not a crisis. It is a trigger for disorder — an emotional and informational storm in which communication plays an essential role in containing the chaos.

It defines who speaks, what to say, to whom, in what order, when to speak and how. This is all the more critical as AI compresses the dynamics of crises.

Read our article on crisis dynamics

Yet crisis communication remains the great absent from crisis management preparations. Organisations invest in technical cybersecurity, and put in place — to varying degrees — a crisis management team and plan. But time and again, they overlook the human and organisational dimension of preparation.

This is a structural mistake **in the AI era, where you need to respond faster and more accurately** — without a solid framework and content prepared well in advance.

What are we talking about, exactly?

A systemic crisis communication strategy, adapted to new AI threats. A strategy that takes into account the entire ecosystem and its stakeholders — employees, partners, suppliers, media, authorities, clients — because attackers, whoever they are, exploit every weakness and point of entry.

Preparing this strategy is like preparing your weapons and knowing when and how to deploy them at the moment of attack and crisis.

What weapons are we talking about?

The essential foundation is a plan that details:

- cyber-AI risks
- crisis simulations to be conducted under real conditions
- the organisation's policy and posture in the event of a cyber-AI crisis
- the crisis management team — trained and rehearsed, with clearly defined roles and responsibilities
- the detection and alert system covering all types of possible anomalies
- secure alternative communication channels
- narratives and messages prepared and validated in advance
- critical content in all formats to be protected, and the type of protection planned
- security guidelines to be communicated
- up-to-date contact lists
- the media landscape (online, local, regional, national, sector, tech, cyber)

This foundation is the tip of the iceberg of human and organisational cybersecurity. It is the pressure valve that contains the strain generated across the ecosystem during a cyber-AI crisis.

The fewer human and organisational vulnerabilities your organisation carries, the more effective your cyber-AI crisis communication will be — when the moment comes.

Managing your informational attack surface

Before any crisis occurs, every piece of content published, every position taken, every official statement constitutes raw material for an adversary.

The goal isn't silence — it's deciding what you publish, where, how, and with what supporting evidence. Truth is no longer self-evident in the era of deepfakes and synthetic narratives.

Truth must be governed, proven, traceable.

Seven concrete questions to ask yourself

- Can you list the most sensitive data and content within your organisation — without which your organisation could no longer function? If so, are your employees aware of it?
- What data and sources underpin this content or position? Are they verifiable, traceable, defensible?
- What narrative and manipulation risks are you taking versus the reputational gains? In other words, is it really worth the candle?
- Who internally validates content before publication? Is there a double-verification protocol for sensitive or critical communications?
- Once published, could this content be weaponised against you in another context — geopolitical, competitive, or media-related?
- Which trusted third parties could authenticate or amplify this content to give it credibility under attack?
- Should the content be challenged, do you have the necessary evidence to defend the original version before the public, strategic stakeholders, and authorities?

This is the purpose of Truth Governance — putting in place the mechanisms that ensure the information you disseminate is verifiable and defensible, before an adversary turns it against you.

Reducing organisational vulnerabilities to manipulation

Manipulation is a tactic as old as the world.

The difference in the AI era is its pervasiveness — you don't know your enemy — its unpredictability — when and how it will strike — its reach and momentum in this interconnected world.

It exploits human and organisational blind spots, such as:

- management that ignores or downplays weak signals
- disengaged teams that no longer surface anomalies
- pressure that generates mechanical, unconsidered decisions
- ambiguous rules
- siloed structures that fragment communication and cohesion
- etc.

Read our article on weak signals

Before facing a crisis, you need to diagnose your own vulnerability across at least four dimensions, with precise questions.

- the flow of information and alerts
- internal culture and its fragilities
- dependencies and blind spots
- engagement and loyalty

On the flow of information and alerts

- Do your teams know what weak signals are?
- If so, do those signals actually reach leadership — or are they filtered, softened, or normalised along the way?
- Is there a formal channel for flagging an anomaly, an inconsistency, or unusual or inappropriate behaviour — without going through the direct chain of command?

On internal culture and its fragilities

- Are there zones of silent tension — between teams and/or hierarchical levels — that could be exploited from outside?
- Is your internal language clear and consistent, or vague and open to manipulation? Do words mean the same thing to everyone?
- How much does conformism shape decision-making? Are disagreements expressed, or self-censored?

On dependencies and blind spots

- Which suppliers, partners, or service providers have access to certain sensitive information — and with what integrity guarantees?
- Are there key individuals whose departure or compromise would destabilise the entire apparatus?
- Which decisions rest on a single source, a single tool, a single AI model — with no cross-verification possible?

On engagement and loyalty

- Do your teams understand all the cyber-AI risks to which the organisation is exposed? If so, how can you ensure this without drift in interpretation?
- Do your teams trust leadership to manage a crisis — and will they follow security guidelines scrupulously? Are they sufficiently engaged to detect and flag anomalies, even under pressure?
- What is the level of resistance to AI adoption — among whom, and why?

The most challenging question: "If an adversary wanted to destabilise our organisation from the inside, where would they start?"

This is precisely what the Resilience Culture Lab helps identify and address — not after the incident, but well before it. It is the very foundation of organisational resilience.

Revisiting your leadership posture

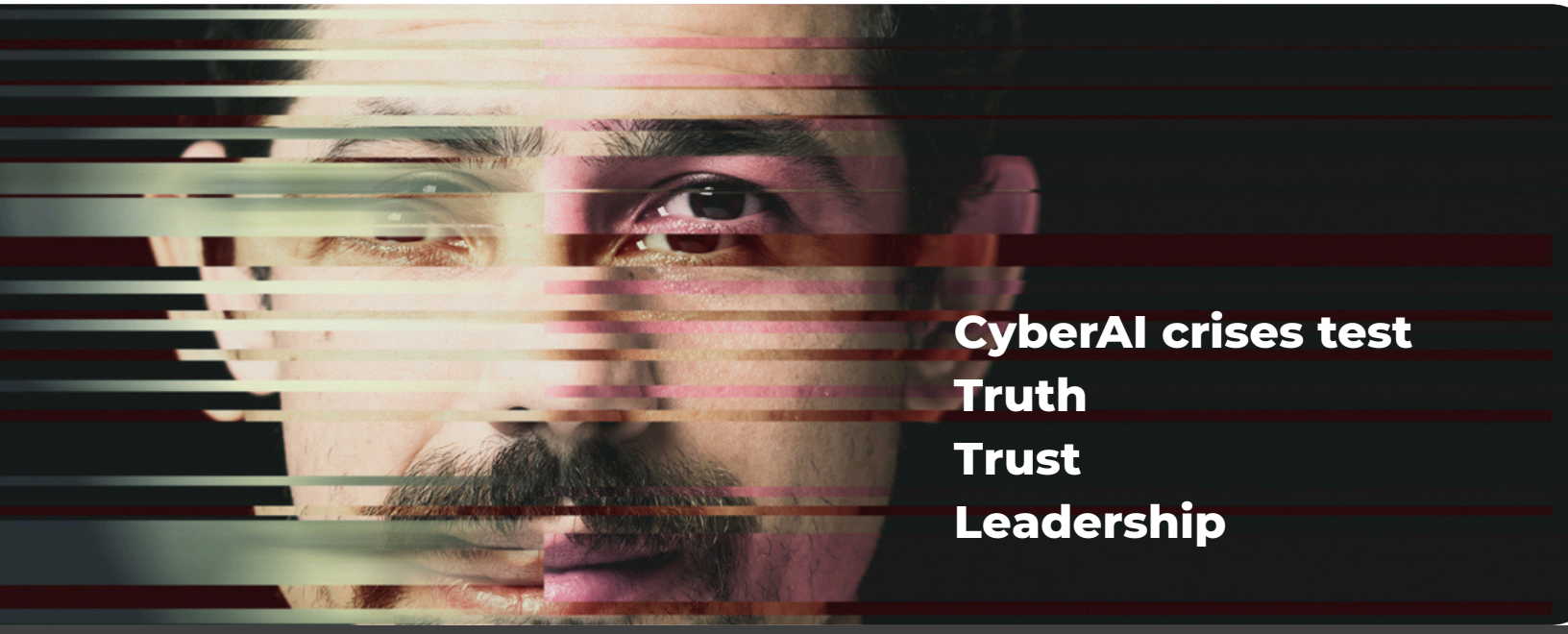
You already know this: only example counts.

You may have the most compelling discourse in the world, but...

- if it is not followed by actions and decisions aligned with your words
- if you maintain ambiguity in your rules
- if you tolerate toxic behaviours from high performers
- if you don't listen to the echo from the ground on your strategic decisions
- etc.

... You are opening the door wide to cyber-AI threats, which play on perceptions, uncertainties, and doubts.

Beyond ethics, coherence is a shield — one that generates trust and cohesion. A leader whose actions and words align is infinitely harder to discredit. This work on posture — what teams perceive, what stakeholders read, what adversaries seek to turn against you — is at the heart of Resilient & Ethical Leadership.



A few questions to ask yourself personally, in your heart of hearts.

On personal credibility under attack

- If I arrange reality to suit my purposes — on results, commitments, facts — what attack surface am I creating for an adversary who wants to make me say, via a deepfake, what I never said?
- Should a deepfake occur, will my word carry enough credibility to be believed — or will doubt set in too easily because "it sounds just like something he'd say"?

On tolerance of toxic behaviours

- If I allow high performers to act contrary to the values I publicly espouse, what message am I actually sending the organisation? And if tomorrow I must embody ethics in the face of a crisis, who will believe me?
- The gap between what I tolerate internally and what I preach publicly is exactly what an adversary seeks to expose or amplify.

On decisional isolation

- Do I leave my managers to decide alone, without a safety net, feeling abandoned in the face of difficulty? Hackers and manipulators exploit the sense of isolation — a person who feels alone is vulnerable, permeable to pressure, blackmail, and social manipulation.
- Resilience is built in the daily relationship between a leader and their teams — particularly in the moments when nothing is going right.

The overarching reflection that frames all three: "What I do when no one is watching will always end up defining who I am when everyone is watching."

5 key takeaways

1. **Waiting for stabilisation is a mistake** — in an environment of fluctuating instability, nothing stabilises. The only solution is to focus on what you can change and influence.

2. **Cyber-AI crisis communication is the great absent from crisis preparations** — this is a structural mistake in the AI era, where you need to respond faster and more accurately with a solid framework prepared well in advance.

3. **Truth must be governed, proven, traceable** — every piece of content published is raw material for an adversary. Managing your informational attack surface means deciding what you expose and how you defend it.

4. **Manipulation exploits human and organisational vulnerabilities** — managerial blind spots, silos, and disengagement are the points of entry that attackers target first.

5. **A leader's coherence is a shield** — the gap between what you tolerate internally and what you advocate publicly is exactly what an adversary seeks to expose.

In a world where reality can be manipulated, preparation is no longer optional.

Where does your organisation stand?

25 questions. 15 minutes. Immediate and anonymous results.

Take the free assessment

contact@ozngo.com