



# AI ACTION SUMMIT

## **AI & Cyber :** a crisis management exercise to strenghten cooperation





# AI ACTION SUMMIT

## Guiding questions

  
**RÉPUBLIQUE  
FRANÇAISE**  
*Liberté  
Égalité  
Fraternité*





# AI ACTION SUMMIT



## Table of contents

Exercise scenario .....	4
Guiding questions.....	4



# AI ACTION SUMMIT



## Exercise scenario

Several cyber authorities issue an alert related to a vulnerability in an open source project (office automation services). This vulnerability is used by an attacker to recover confidential information and put pressure (blackmail) on organisations affected. After blackmailing some organisations, the attacker decides to publish some of the sensitive data recovered from online storage.

In parallel one targeted company face an incident of production affecting one of its critical activity.

## Guiding questions

- 1- How can information sharing about vulnerabilities be improved between AI solution providers and their clients?
- 2- What processes or indicators would alert your organisation of a potential vulnerability exploitation on your AI systems? How would you communicate the discovery of such a vulnerability within your organisation?
- 3- How do you set criteria for selecting an AI model (open source, proprietary ...) or provider and its deployment (on premises, SaaS ...)? How do you assess the cybersecurity maturity of an AI provider? Or open source model use within your systems?
- 4- How do your monitoring and anomaly detection systems adapt in case of a confirmed attack on your production models? How do you analyse past data to trace an attack that has already taken place? How do you check the data perimeter this AI system has access to?
- 5- While facing such a situation, what would be the first actions performed by your organisation (internal investigations, notification of competent authorities, crisis checklist / specific set-up, communication Int / Ext ...)?
- 6- If a model is vulnerable to compromising, what steps would you take to assess the impact on your strategic customers? How do you evaluate the potential risks associated with deploying specialised AI models in production environments?
- 7- What crisis management strategies would you implement? Are they specific due to the nature of the impacted system?
- 8- How can organisations balance the need to leverage AI for enhanced security with the potential risks related to data privacy, algorithmic bias, and over-reliance on automated systems?
- 9- How do your monitoring and anomaly detection systems adapt in case of a confirmed attack on your production models? What are your specific response procedures for anomalies detected in a crisis situation?
- 10- In case of a cyberattack targeting one of your AI-enabled solution or system, what emergency measures do you implement to further isolate and secure this environment? How do you manage risks associated with external resources in such a situation?
- 11- How are you evolving your security strategies for model training and isolation in the face of emerging threats? What innovations are you considering to strengthen model protection? What kind of mechanisms you can add to avoid these types of situation?