Governare l’autonomia degli agenti nella Cyber Security

Federico Cerutti

RICHMOND
CYBER RESILIENCE
FORUM

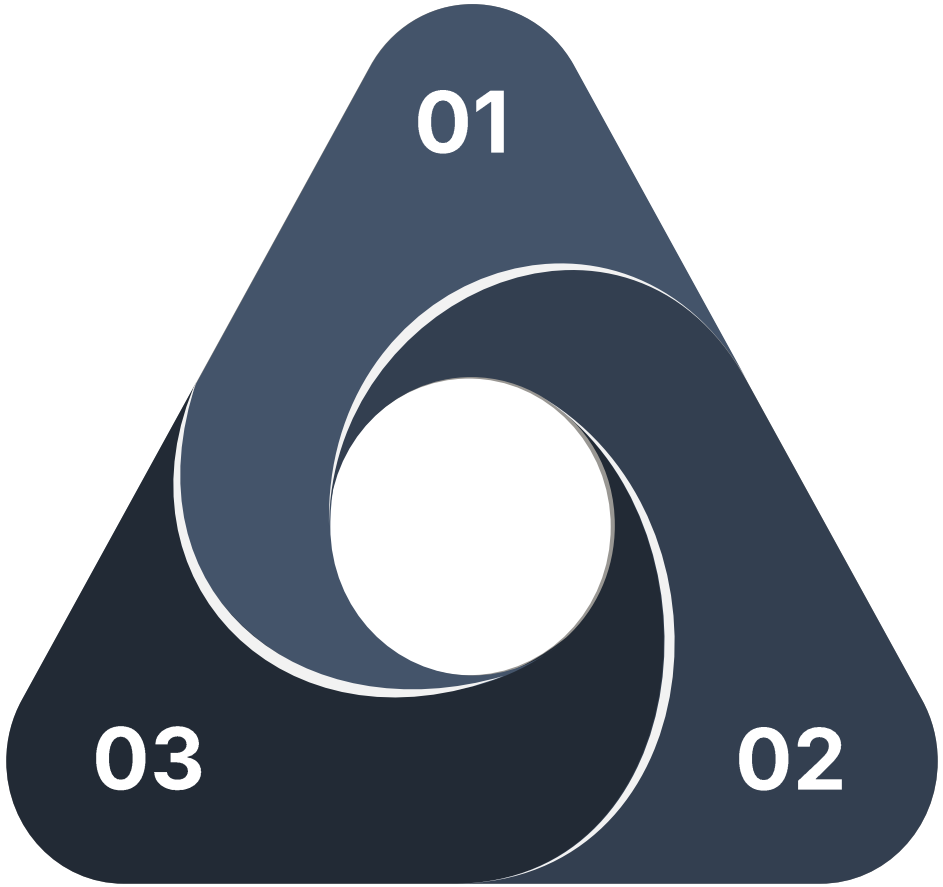
RIMINI
10-12 MAGGIO 2026
SPRING EDITION

VOTA LA CONFERENZA

Image by ChatGPT

Images from Pexel

Investigations



Crisis Management

Cyber Threat Intelligence

Image from Pexel

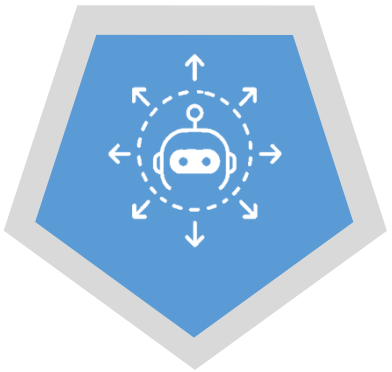
Image from Pexel

Image by Gemini

Il consiglio sbagliato si discute

L'azione sbagliata si paga

Perdita del controllo



Compromissione

Goal drift

Manipolazione della memoria

Abuso delle autorizzazioni

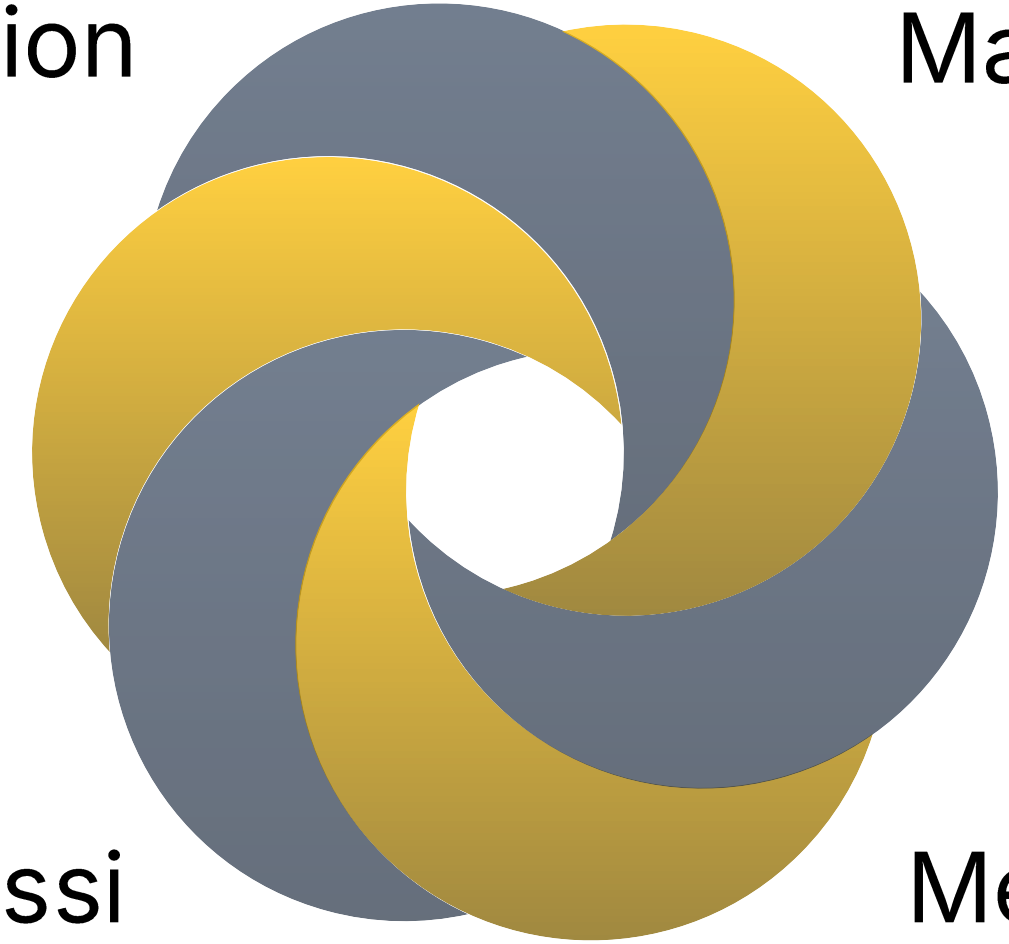
Icons by ChatGPT

La conoscenza decide l'azione

Image by ChatGPT with a concept from Pexel

Escalation

Mandato



Evidenze

Fonti

Permessi

Memoria

Non è
un *tool*
da usare

È
firmare
una delega

Il valore dell'agente è nella conoscenza che lo guida.

Grazie

RICHMOND ITALIA
HUMAN 2 HUMAN EVENTS

VOTA LA CONFERENZA