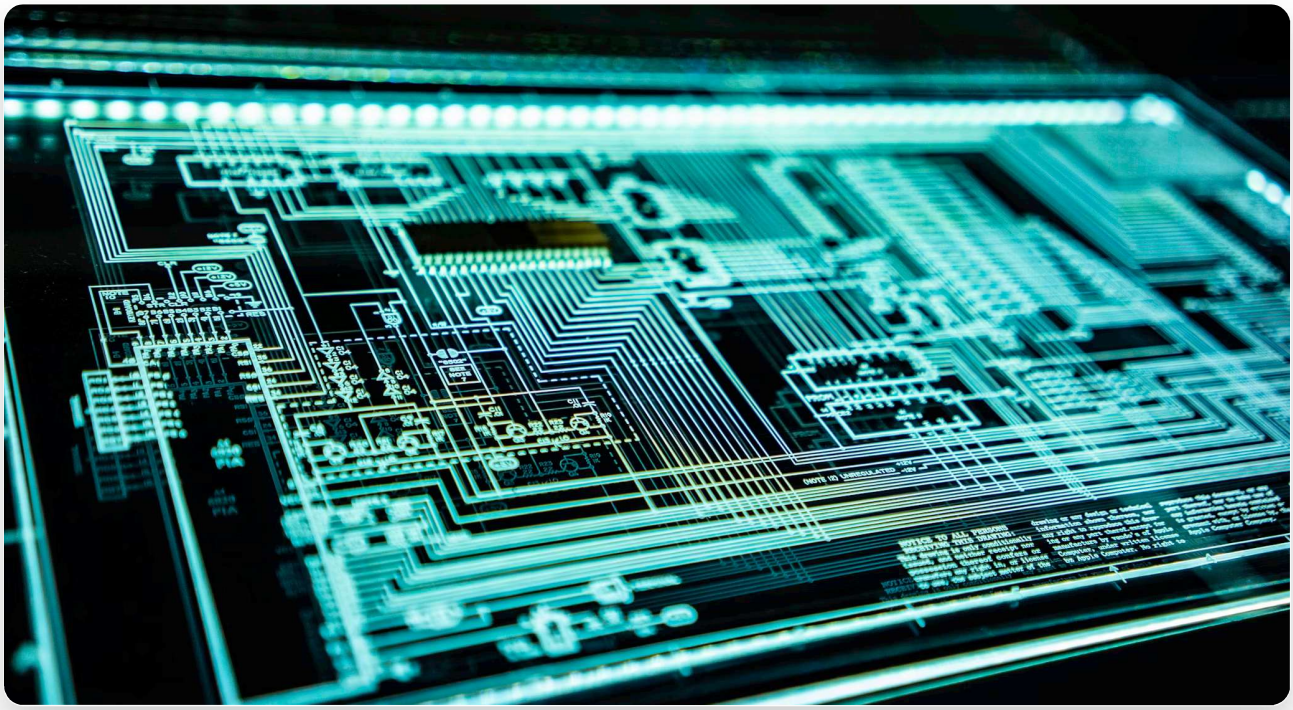
BTS SIO - OPTION SISR

VEILLE TECHNOLOGIQUE

L'IA Générative au service des Cyberattaques



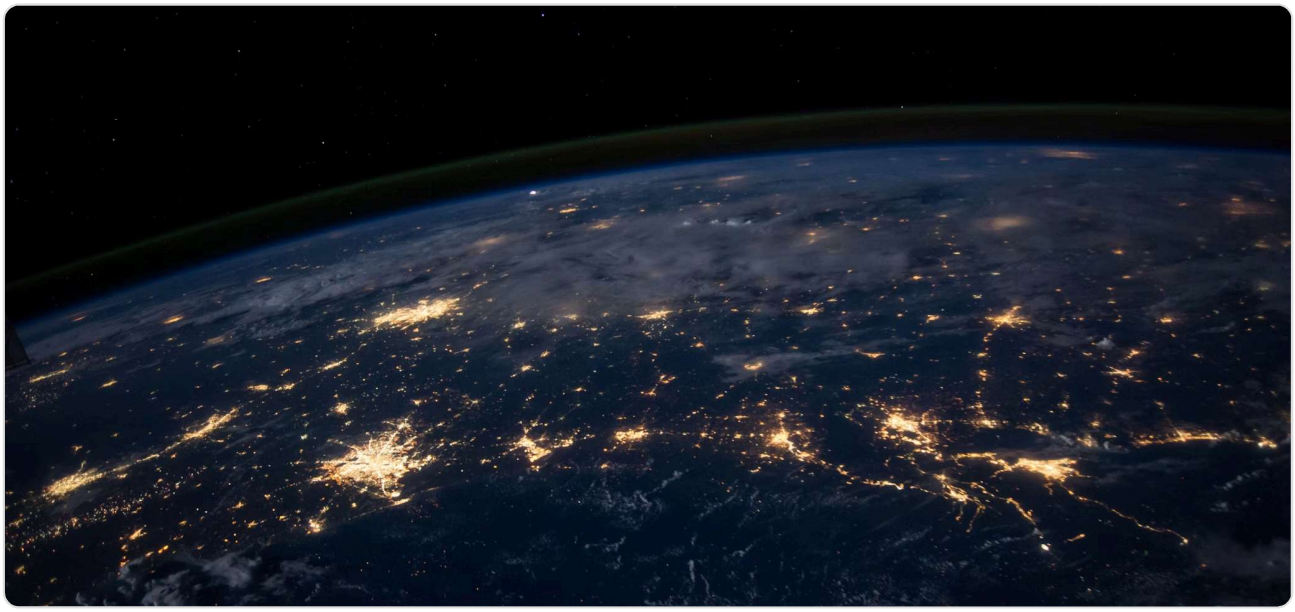
Présenté par : Salimou TOURAY

Session : 2025

OBJECTIF DE LA VEILLE

Comprendre comment l'Intelligence Artificielle (ChatGPT, Deepfakes) transforme les méthodes des attaquants pour contourner les défenses classiques, et analyser l'évolution des menaces d'Ingénierie Sociale.

SOMMAIRE



1. MÉTHODOLOGIE DE VEILLE

Pour réaliser cette veille, j'ai mis en place un processus de surveillance automatisé afin de filtrer les informations pertinentes parmi le bruit médiatique.

Outils utilisés :

- **Feedly :** Agrégation des flux RSS (ANSSI, The Hacker News).
- **Google Alerts :** Mots clés "IA Phishing", "Deepfake Security".
- **Newsletter :** CERT-FR (Hebdomadaire).

Figure 1 : Interface de monitoring et d'agrégation de flux.

I. Présentation Générale

L'IA Générative permet de créer du texte, des images ou de la voix à partir de simples instructions. Si des outils comme ChatGPT sont bénéfiques pour la productivité, ils ont offert une "arme clé en main" aux cybercriminels.

Le problème majeur pour les administrateurs réseaux (SISR) n'est plus seulement la faille technique, mais la **manipulation parfaite de l'humain** rendue possible par l'IA.

II. LA RUPTURE TECHNOLOGIQUE

Deux technologies principales changent la donne :

- **LLM (Large Language Models) :** Capables de comprendre et générer du texte humain avec une cohérence parfaite et multilingue (Ex: GPT-4).
- **GAN (Generative Adversarial Networks) :** Réseaux de neurones qui créent des images ou voix synthétiques (Deepfakes).

III. Attaque Classique vs IA

Avant l'IA, le phishing était souvent détectable par des fautes ou un style approximatif. Aujourd'hui, la barrière tombe.

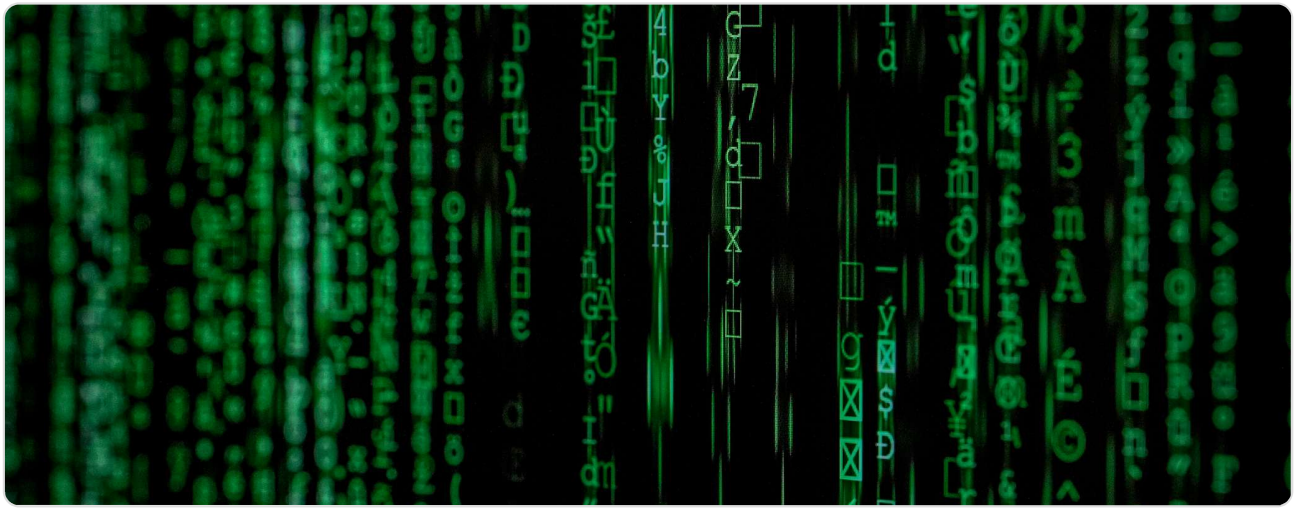


Figure 2 : L'IA génère du code polymorphe et des textes parfaits.

Attaque Classique	Attaque par IA
Orthographe approximative	**Syntaxe parfaite**, ton professionnel
Envoi de masse générique	**Ciblage contextuel** (via LinkedIn)
Traduction automatique douteuse	Langue native et argot métier

IV. LE "SPEAR PHISHING"

L'IA permet d'industrialiser le harponnage (attaque ciblée). Un attaquant peut demander : *"Analyse le profil LinkedIn de ce DAF, et rédige un mail urgent de son fournisseur habituel."* Le résultat est indétectable par les filtres classiques.

V. Les Deepfakes (Vishing)

Le **Voice Phishing** utilise l'IA pour cloner la voix d'une personne à partir de 3 secondes d'audio.

> **Scénario réel :** Février 2024, une multinationale perd 25M$ à Hong Kong. Un employé a viré les fonds après une visioconférence où **tous les participants** (sauf lui) étaient des Deepfakes générés par IA.

Figure 3 : Usurpation d'identité biométrique par Deepfake.

VI. HISTORIQUE (2022-2024)

- **Nov 2022 :** Sortie de ChatGPT. Premiers scripts malveillants générés.
- **2023 :** Apparition de "FraudGPT" et "WormGPT" sur le Dark Web (IA sans éthique pour pirates).
- **2024 :** Démocratisation des Deepfakes vocaux et vidéo en temps réel.

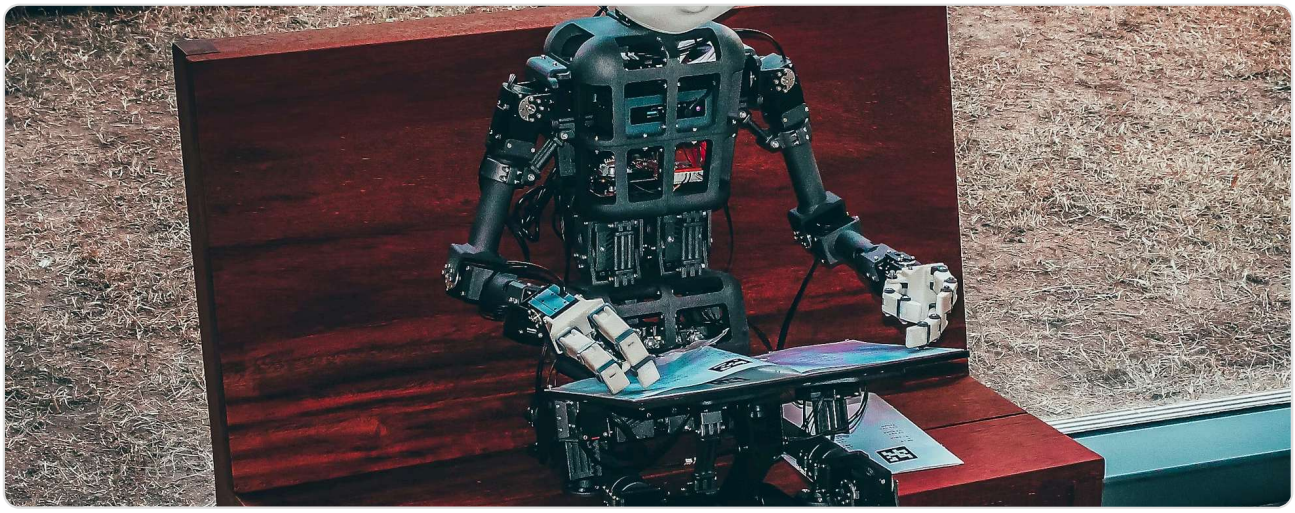


Figure 4 : La course aux armements : IA Offensive contre IA Défensive.

VII. L'Avenir & Conclusion

L'avenir de la cybersécurité est une guerre "IA contre IA". Les humains ne sont plus assez rapides pour analyser les logs. Les entreprises devront déployer des **IA Défensives** (SOC) capables d'analyser le style d'écriture ou les micro-anomalies vocales.

Conclusion : L'IA a abaissé la barrière d'entrée pour les cybercriminels. Pour un SISR, la protection technique ne suffit plus : la sensibilisation des utilisateurs est la nouvelle priorité absolue.

Sources : ANSSI (Rapports 2023), The Hacker News, Wired, CNN Business.