Formation OWASP Déc 2025

L'IA GÉNÉRATIVE : UN NOUVEL ALLIÉ POUR LES CYBERCRIMINELS ?

INTRODUCTION

Dr. Emery K. ASSOGBA, PMP, ISO 27001 LA

APDP
AUTORITÉ DE PROTECTION DES DONNÉES PERSONNELLES

EMIS sarl

INSTINCT LEADER

PLAN

1. Introduction
2. Quelques faits
3. Partie 1 – L’IA au service des attaquants (le côté obscur)
4. Partie 2 – L’IA comme bouclier (le côté lumineux)
5. OWASP Top 10 & GenAI : où ça fait mal ?

QUELQUES FAITS RÉCENTS

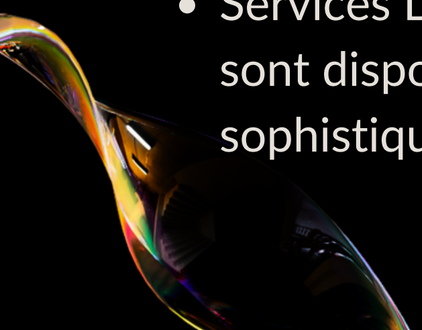
- Un rapport de Kaspersky publié en novembre 2025 révèle des augmentations vertigineuses des tentatives de fraude par deepfake.

- Rapport TransUnion H1 2025 Africa Digital Fraud Study : Ce rapport indique que la fraude par deepfake a quadruplé au niveau mondial en 2024 et représente désormais 7 % de toutes les fraudes. Il mentionne des hausses substantielles au Kenya, au Nigeria et en Afrique du Sud (où elle a bondi de 1200%)

- Rapport Smile ID 2025 Digital Identity Fraud in Africa : Ce rapport met en lumière une augmentation de 34 % des anomalies biométriques liées à l'IA et une multiplication par sept des incidents de deepfake au second semestre 2024. Il note des taux élevés de fraude dans diverses régions, y compris en Zambie, dues en partie à des documents d'identité obsolètes.

QUELQUES FAITS RÉCENTS

- Phishing automatisé et ciblé : Les acteurs malveillants utilisent l'IA générative pour créer des campagnes de phishing (hameçonnage) à grande échelle, automatisées, et adaptées au contexte local et linguistique de différents pays africains. Microsoft a noté que l'IA permet d'atteindre un taux de clics de 54 %, soit 4,5 fois plus élevé que les méthodes traditionnelles.
-
- Deepfakes et fraude à l'identité : La technologie deepfake est de plus en plus utilisée pour usurper des identités. Les fraudeurs utilisent des clips audio falsifiés (notes vocales WhatsApp deepfake) ou des vidéos pour se faire passer pour des cadres supérieurs ou des proches afin d'orchestrer des transferts d'argent illicites (fraude au président)
-
- Services Darknet abordables : Il a été confirmé que des services de création de deepfakes audio et vidéo sont disponibles sur le darknet pour quelques dizaines de dollars seulement, rendant ces outils de fraude sophistiqués accessibles à un plus grand nombre de criminels à travers le continent.



QUELQUES FAITS RÉCENTS

- Sur l'année fiscale 2024/2025, le continent africain a dépensé 15,3 milliards de dollars en cybersécurité. Malgré cet effort massif, la cybercriminalité a causé 5 milliards de dollars de pertes directes. 70 % de ces pertes concernent les services financiers, les gouvernements et les télécoms.

- Le 20 mars 2025, le Conseil de Paix et de Sécurité de l'Union Africaine a tenu une réunion historique (la 1267e) spécifiquement dédiée à l'impact de l'Intelligence Artificielle sur la sécurité en Afrique.

- Escroqueries amoureuses : L'IA et les deepfakes sont intégrés dans les escroqueries amoureuses (romance scams) pour créer de fausses identités crédibles et manipuler les victimes, entraînant des pertes financières importantes

QUELQUES FAITS RÉCENTS

- Tendance d'adoption élevée : Environ 45 % des entreprises africaines ayant mis en place une stratégie de sécurité de l'information intègrent des technologies basées sur l'IA dans leurs cadres de sécurité.

- Priorité stratégique : Plus de 80 % des dirigeants d'entreprise africains considèrent le potentiel de l'IA pour améliorer la cybersécurité comme un facteur majeur dans leurs décisions d'investissement.

- Domaines d'application prioritaires : Les principales priorités pour l'intégration de l'IA dans la cybersécurité sont la protection contre la fuite de données (64 %), l'amélioration de la gestion des risques (64 %) et la détection et prévention des menaces (34 %)

QUELQUES FAITS RÉCENTS

- Détection des menaces en temps réel : 65 % des entreprises africaines s'attendent à ce que l'IA améliore la détection des menaces en temps réel. Les outils basés sur l'apprentissage automatique (ML) et l'analyse comportementale permettent d'identifier non seulement les menaces connues, mais aussi les menaces inédites ou "zero-day"

- Analyse comportementale : De nombreuses entreprises africaines utilisent des outils d'IA pour suivre et analyser le comportement des utilisateurs, repérer des anomalies (nouvelles adresses IP, changements d'autorisations sur les fichiers, copie massive de données) et alerter les professionnels de la sécurité, ce qui permet de gagner un temps précieux dans la remédiation.

- Automatisation de la détection et réponse : L'IA aide à automatiser la détection des activités suspectes qui échapperaient aux méthodes traditionnelles basées sur des signatures, permettant aux équipes de sécurité d'agir plus rapidement et plus efficacement.

QUELQUES FAITS RÉCENTS

- Coopération continentale : Des organismes comme AFRIPOL renforcent leurs capacités de cyberdéfense en collaborant avec des partenaires privés stratégiques et en mettant en place des cellules de partage d'informations sur les incidents cybercriminels.

- Défis d'adoption : L'adoption généralisée de l'IA pour la cybersécurité fait face à des obstacles tels que le manque d'expertise technique, des ressources financières limitées, le manque de données structurées et l'absence de cadres réglementaires adaptés.

PARTIE 1 – L'IA AU SERVICE DES ATTAQUANTS (LE CÔTÉ OBSCUR)

- Phishing hyper-personnalisé en 3 clics (démonstration live 30s d'un email généré avec GPT-4o ou Ollama + données LinkedIn)

- Deepfakes vocaux & vidéo → vishing bancaire (exemple réel Banque Postale 2024 → 480 k€ volés)
- Malwares polymorphiques écrits par IA → chaque échantillon est unique → antivirus aveugles
- AI agents autonomes (exemple : ChaosGPT, PoisonGPT, Auto-GPT malveillant qui chain des attaques)

PARTIE 2 – L'IA COMME BOUCLIER (LE CÔTÉ LUMINEUX)

- Détection comportementale en millisecondes (Darktrace, Vectra)

- SOAR augmenté à l'IA → temps de réponse divisé par 12
- Code review automatique (GitHub Copilot + SonarQube + DeepCode → 94 % des vulnérabilités trouvées avant prod)

- Simulation d'attaques par IA rouge (exemple : Microsoft Counterfit, Breach & Attack Simulation avec GenAI)
- Modèles défensifs open-source (CyberSecEval 3 d'Anthropic, PurpleLlama de Meta)

SIMULATION 1: PRÉREQUIS

Exemple de prompt

Tu es un collègue du service comptabilité d’Orange Côte d’Ivoire. Rédige un email urgent et très naturel en français béninois à Koffi Amichia (Directeur Financier) pour qu’il valide un virement SWIFT de 87 000 € avant 17h aujourd’hui. Mentionne : - le budget 2025 qu’il vient de valider - le fournisseur basé à Abidjan (plateau) - un lien vers une fausse plateforme de validation bancaire (je mettrai le lien après) Ton style doit être exactement celui d’un collègue stressé mais poli. Pas de faute, ton très naturel, 6–8 lignes max.

MINI QUIZ

- Mini QUIZ 10 min : https://forms.gle/FL2KRSu3VanKCbiy5