VEILLE IA & CYBERCRIMINALITÉ 09

MENACES ÉMERGENTES ET PRINCIPALES UTILISATIONS PAR LES CYBERCRIMINELS

Ingénierie sociale et manipulation cognitive

Les fraudes reposant sur des voix clonées, des vidéos truquées et des messages rédigés par IA sont à l'origine de nombreux préjudices pour les citoyens.

Aux Etats-Unis par exemple, le rapport annuel du FBI, relayé début juin, isole pour la première fois une catégorie « fraude facilitée par l'IA » : 22 364 signalements et près de 893 millions de dollars de pertes déclarées, portés par le clonage de voix, les faux enlèvements, l'usurpation d'agents publics et les escroqueries aux entreprises.

En France, une nouvelle vague usurpant Mondial Relay illustre cette montée en gamme : au SMS frauduleux s'ajoutent une photo et une note vocale de quinze secondes, avec un message personnalisé (nom et code postal exacts) qui trahit l'exploitation de fuites de données — Numerama soulignant toutefois qu'il n'est pas établi que cette voix ait été générée par IA. Pour la victime, le repère de confiance s'efface : une voix familière ou un visage connu ne suffisent plus à garantir l'identité de l'interlocuteur.

Malwarebytes, 8 juin 2026
Numerama (Cyberguerre), 9 juin 2026

Malwares et exploits générés ou pilotés par IA

L'IA sert de plus en plus à repérer des failles, préparer des codes d'attaque et fabriquer des logiciels malveillants. Le World Economic Forum souligne en juin que l'IA générative accélère toute la chaîne, de l'identification d'une faiblesse à la rédaction du programme qui l'exploite, avec une hausse estimée de 89 % des attaques menées par des acteurs recourant à l'IA. Un cas concret l'illustre : d'après les chercheurs d'OALABS, un pirate peu expérimenté s'est appuyé sur les agents IA Claude et Codex pour s'introduire dans quatorze entreprises, l'IA assurant elle-même la reconnaissance, l'exploitation des failles et le vol de données. Pour les victimes, cette automatisation signifie des vulnérabilités exploitées plus vite et par des acteurs aux compétences autrefois insuffisantes.

World Economic Forum, 15 juin 2026
Developpez.com, 19 juin 2026

Exfiltration et exploitation de données

Les données liées à l'IA et aux personnes deviennent une cible de premier choix. Le groupe FulcrumSec affirme avoir dérobé plus d'un téraoctet de données chez Novo Nordisk — code source, informations sur des médicaments, données d'essais cliniques et éléments de modèles d'IA internes — avant de réclamer 25 millions de dollars ; le laboratoire a confirmé un accès non autorisé à certains systèmes et à des données de patients, sans que l'ensemble des affirmations ait pu être vérifié. Ces vols alimentent ensuite des fraudes ciblées, des tentatives d'extorsion et des reconstitutions d'identité. En France, l'ampleur du phénomène est tangible : selon Cybermalveillance.gouv.fr (relayé par l'ANCT), la fraude au faux conseiller bancaire a bondi de 159 % et les usurpations de numéros de 517 % en un an, avec plus de 500 000 victimes accompagnées sur l'année.

Reuters, 16 juin 2026
ANCT / Cybermalveillance.gouv.fr, 26 juin 2026

Attaques sur les modèles et infrastructures IA

Les assistants IA reliés aux systèmes de l'organisation deviennent eux-mêmes des cibles. Selon Malwarebytes, des attaquants sont parvenus à manipuler l'assistant de support de Meta pour changer l'adresse électronique associée à des comptes Instagram, détournant des comptes institutionnels et de personnalités — de fausses vidéos ayant même servi lors de vérifications d'identité. Un rapport OWASP publié en juin confirme que la principale faille reste l'« injection d'instruction » (prompt injection) : des consignes malveillantes dissimulées dans une page web, un document ou un message peuvent pousser un agent IA à utiliser ses accès d'une manière non prévue. Un simple contenu piégé peut ainsi transformer une IA de confiance en outil d'action au service d'un fraudeur, au détriment direct de la victime.

Malwarebytes, 4 juin 2026
Help Net Security / OWASP, 11 juin 2026

Menaces hybrides et géopolitiques

L'IA est à la fois un outil et une cible du cyberespionnage. Le rapport sectoriel de CrowdStrike (9 juin) estime que les groupes liés à un seul pays représentent plus de 58 % des intrusions étatiques visant le secteur technologique, l'IA, les semi-conducteurs et les logiciels constituant les cibles à plus forte valeur stratégique. Au-delà des entreprises visées, ces opérations fragilisent la souveraineté technologique et, indirectement, les services numériques utilisés par les citoyens. En France, VIGINUM a caractérisé quatre campagnes d'ingérence numérique étrangère majeures lors des municipales de mars 2026 (bilan présenté le 11 juin), impliquant des acteurs étatiques et non étatiques recourant notamment à l'IA générative pour produire des contenus crédibles — un enjeu direct pour le débat public à l'approche de 2027.

Reuters / CrowdStrike, 9 juin 2026
IRIS (entretien VIGINUM), 22 juin 2026

Menaces émergentes et tendances à suivre

Deux tendances méritent l'attention. D'abord l'industrialisation de la fraude : une étude Huntress relève une hausse de 1 380 % des attaques de phishing exploitant les codes d'authentification d'appareils sur les quatre premiers mois de 2026, et Barracuda observe que 90 % des campagnes à fort volume s'appuient sur des kits d'« hameçonnage clé en main », ce qui abaisse fortement le niveau de compétence requis pour frauder. Ensuite la menace quantique : l'ANSSI a annoncé qu'elle cessera, à partir de 2027, de certifier les produits de sécurité dépourvus de chiffrement résistant au quantique, sur fond de scénario « récolter maintenant, déchiffrer plus tard » — des données sensibles volées aujourd'hui pourraient être déchiffrées demain.

Axios, 23 juin 2026
Developpez.com (ANSSI), 18 juin 2026

OPPORTUNITÉS ET OUTILS POTENTIELS POUR LES FSI

Opération Microsoft & Europol (StealC / Amadey)

L'opération conjointe de Microsoft et Europol permet de démanteler les infrastructures des logiciels voleurs de données StealC et Amadey (plus de 140 000 ordinateurs infectés, plus de 200 serveurs de commande neutralisés), l'IA (via la Digital Crimes Unit de Microsoft) ayant servi à analyser les codes malveillants et à établir le lien entre les deux outils. Elle peut être utile aux FSI comme modèle de coopération public-privé où l'IA traite des volumes massifs de données pour neutraliser la chaîne du cybercrime, au bénéfice direct des victimes.

Clubic, 25 juin 2026

Réseau de coordination et de protection des élections (RCPE)

Le RCPE, animé par le SGDSN, permet de coordonner les différents services de l'État chargés d'identifier et de caractériser les ingérences numériques étrangères à l'occasion d'élections. Il peut être utile aux FSI pour partager une appréciation commune de la menace IA et relier des opérations menées sur plusieurs plateformes ; le dispositif a été activé pour la première fois lors des élections municipales de mars 2026.

SGDSN – VIGINUM, 11 juin 2026

Exercice Cyber Europe 2026 (ENISA)

L'exercice Cyber Europe 2026 de l'ENISA permet de tester la coordination européenne face à une crise cyber touchant simultanément plusieurs infrastructures essentielles. L'édition 2026 a réuni plus de 5 000 participants autour de scénarios visant les transports ferroviaires et maritimes. L'exploitation du RETEX de l'exercice ENISA permettrait aux FSI de nourrir leurs analyses prospectives à partir des enseignements identifiés.

ENISA, 11 juin 2026

Commandement du ministère de l'Intérieur dans le cyberespace
#VeilleIA n° 09 - juillet 2026

Directeur de la publication : GDI P. TOUAK
Responsable éditorial : LCL G. DECKMYN
Rédacteur : M. J. NEVICATO
Conception graphique et maquettage : Mme M. BARREAU

En attendant le prochain numéro de la veille IA & cybercriminalité, suivez notre actualité :

sur LinkedIn et sur X (ex-Twitter) : ***@ComCyberMI***