



CHECKLIST 5 :

SÉCURITÉ DES SYSTÈMES D'IA GÉNÉRATIVE

L'IA générative introduit de nouveaux vecteurs de menaces : injection de prompts, empoisonnement des données, hallucinations, fuites d'informations confidentielles. La sécurisation de vos systèmes d'IA nécessite une approche globale qui couvre l'ensemble du cycle de vie, de la conception au déploiement. Cette checklist vous fournit les bonnes pratiques pour identifier les risques, mettre en place des mesures de protection et assurer une surveillance continue de vos systèmes.

→ Analyse des Risques

- ☐ Identifier les menaces spécifiques à l'IA générative : injection de prompt, empoisonnement des données (*data poisoning*), vol de modèle, hallucinations.
- ☐ Évaluer les risques de fuites de données confidentielles via les interactions avec les modèles.
- ☐ Anticiper l'utilisation malveillante de l'IA pour des cyberattaques (ex: *phishing* amélioré, génération de malwares).

→ Sécurisation du Cycle de Vie

- ☐ Appliquer des bonnes pratiques de sécurité dès la phase de conception (*Security by Design*) et tout au long du cycle de vie (entraînement, déploiement, maintenance).

- ☐ Suivre les recommandations de cadres de sécurité reconnus comme ceux de l'ANSSI, du NIST (AI RMF) ou le SAIF de Google.

→ Mesures Techniques de Protection

- ☐ Utiliser des environnements d'exécution cloisonnés ou des serveurs locaux (*on-premise*) pour les données et modèles sensibles.
- ☐ Mettre en place une gestion stricte des droits d'accès et des mécanismes d'authentification forte (ex: double authentification).
- ☐ Chiffrer les données sensibles, que ce soit en transit ou au repos.
- ☐ Anonymiser les données envoyées aux modèles pour éviter l'exposition d'informations personnelles ou confidentielles.

→ Tests et Surveillance

- ☐ Mener des tests de sécurité réguliers, y compris des tests de pénétration (*pentests*) et des exercices de *red teaming* pour identifier les vulnérabilités.
- ☐ Mettre en place un audit et une journalisation en temps réel des prompts et des réponses pour détecter les activités suspectes.

→ Gouvernance et Sensibilisation

- ☐ Définir une charte d'utilisation interne de l'IA, précisant les outils autorisés et les comportements à risque.
- ☐ Former et sensibiliser les collaborateurs aux risques de sécurité liés à l'IA générative.

La sécurité de l'IA générative ne se limite pas aux mesures techniques : elle requiert également une gouvernance claire et une sensibilisation de tous les collaborateurs. En appliquant le principe de "Security by Design" et en suivant les cadres reconnus (ANSSI, NIST, SAIF), vous construisez des systèmes robustes et résilients. La sécurité est un processus continu : testez régulièrement, surveillez activement et adaptez vos mesures aux nouvelles menaces.

Vos systèmes d'IA sont-ils sécurisés ?

Je propose un [audit](#) de sécurité spécifique aux systèmes d'IA générative et des [formations](#) pour sensibiliser vos équipes aux risques et bonnes pratiques. Protégez vos données et vos modèles dès maintenant.

 Contact : sebastien.chami@blackbygreen.com

 Réserver un échange : <https://calendly.com/sebastien-chami/rdv>

 Formulaire de contact : <https://blackbygreen.fr/contact/>