Sécurité et Gouvernance des Systèmes d’Intelligence Artificielle

Lien : https://innov-systems.com/formation/securite-et-gouvernance-des-systemes-dintelligence-artificielle

DURÉE
4 jours (28h)

RÉFÉRENCE
SEC308

CATÉGORIE
Sécurité des Systèmes et Serveurs

Innov Systems

OBJECTIFS DE LA FORMATION

À l'issue de cette formation, vous serez capable de :

- Comprendre les enjeux de sécurité, de gouvernance et d’éthique dans le déploiement des systèmes d’IA
- Identifier les principales menaces et vulnérabilités liées aux modèles d’apprentissage automatique
- Mettre en œuvre des méthodes de sécurisation et de durcissement des systèmes d’IA à chaque étape du cycle de vie
- Savoir auditer, tester et superviser la sécurité d’un système d’IA en environnement de production. Élaborer un plan d’action et des procédures de réponse aux incidents spécifiques à l’IA
- Élaborer un plan d’action et des procédures de réponse aux incidents spécifiques à l’IA

POUR QUI ?

- Développeurs IA / data engineers
- Responsables techniques et chefs de projets IA
- Architectes et responsables sécurité SI
- Experts métiers supervisant l’intégration de solutions d’IA en production

Programme détaillé

1 / COMPRENDRE LES ENJEUX DE LA SÉCURITÉ APPLIQUÉE À L'IA

- Enjeux stratégiques et réglementaires : RGPD, IA Act, responsabilité et conformité
- Typologie des systèmes d'IA (classiques, génératifs, hybrides)
- Cartographie des risques liés à l'IA : données, modèles, infrastructures

2 / STRUCTURE ET CYCLE DE VIE D'UN SYSTÈME D'IA

- Chaîne de valeur de l'IA : données → modèles → intégration applicative
- Risques spécifiques à chaque étape du cycle de vie
- Points de contrôle de sécurité dès la phase de conception

3 / MISE EN PLACE D'UNE GOUVERNANCE DE LA SÉCURITÉ IA

- Rôles, responsabilités et politiques internes
- Auditabilité, traçabilité et documentation du modèle
- Gouvernance des données et gestion du consentement

4 / ANALYSE DES VULNÉRABILITÉS DU MODÈLE

- Attaques par évasion et adversarial examples
- Inversion et exfiltration de modèles
- Empoisonnement des données d'entraînement
- Attaques sur l'apprentissage fédéré

5 / VULNÉRABILITÉS DES APPLICATIONS D'IA GÉNÉRATIVES

- Attaques par injection de prompt et jailbreak
- Fuites d'informations sensibles dans les réponses
- Attaques sur les systèmes RAG (Retrieval-Augmented Generation)

Innov Systems

- Manipulation des contextes et canaux auxiliaires

6 / CHAÎNE D'APPROVISIONNEMENT ET DÉPENDANCES LOGICIELLES

- Menaces sur les bibliothèques open source et modèles préentraînés
- Gestion sécurisée des dépendances et vérification d'intégrité
- Bonnes pratiques de déploiement et isolation des environnements

7 / MÉTHODES DE SÉCURISATION DU PIPELINE D'IA

- Sécurisation des données d'entraînement et des API d'inférence
- Chiffrement, anonymisation et contrôle d'accès
- Surveillance des flux et détection d'anomalies comportementales

8 / STRATÉGIES DE DÉFENSE ET DÉTECTION DES ATTAQUES

- Atténuation des attaques adversariales
- Utilisation d'outils open source pour la sécurité IA (IBM ART, Adversarial Robustness Toolbox, etc.)
- Bonnes pratiques de durcissement et de monitoring en production

9 / ATELIER PRATIQUE : TEST D'INTRUSION D'UN MODÈLE D'IA

- Simulation d'une attaque adversariale
- Analyse des impacts et proposition de contre-mesures
- Audit de sécurité d'un pipeline IA complet

10 / AUDIT DE SÉCURITÉ ET ÉVALUATION DE LA ROBUSTESSE

- Méthodologie d'audit d'un modèle et d'un système d'IA
- Outils et indicateurs de conformité
- Mesure de la robustesse et évaluation du risque global

11 / RÉPONSE AUX INCIDENTS ET GESTION DE CRISE IA

Inov Systems

- Scénarios d’attaques réelles et retour d’expérience
- Plan de continuité et de reprise d’activité IA
- Collaboration entre équipes sécurité, data et conformité

12 / ÉTHIQUE, TRANSPARENCE ET SÉCURITÉ À LONG TERME

- IA responsable : biais, explicabilité et contrôle humain
- Encadrement juridique et déontologique de l’IA sécurisée
- Élaboration d’une feuille de route sécurité IA à 12 mois

Approche pédagogique

- Support Ecrit et Projection
- Exposés Interactifs, Podcasts et Vidéos
- Brainstorming et Jeux de Rôle
- Cas Pratiques et Labs inclus pour leur impact opérationnel
- Test de Validation des Acquis des Connaissances

Prochaines dates programmées

17 au 20 Nov. 2026	Distanciel
16 au 19 Fév. 2027	Présentiel - Casablanca
13 au 16 Avr. 2027	Présentiel - Casablanca
08 au 11 Juin 2027	Présentiel - Casablanca
03 au 06 Août 2027	Présentiel - Casablanca

Autres dates possibles sur demande. Contactez-nous pour organiser une session intra-entreprise.

Innov Systems

Réservation & Renseignements

Téléphone : +212 522 247 210
Email : contact@innov-systems.com
Web : https//www.innov-systems.com

Innov Systems — Tous droits réservés

Innov Systems