PROGRAMME DE FORMATION

Sécurité des IA et des agents connectés au réseau

Évaluez les risques d'un assistant ou agent IA connecté au système d'information et mettez en place des protections vérifiables avant sa mise en service.

DURÉE	NIVEAU	FORMAT	ORGANISATION
2 jours	**Intermédiaire à avancé**	**Présentiel ou à distance**	**Intra-entreprise**

Finalité de la formation

Formation de deux jours pour sécuriser les modèles, assistants et agents IA ayant accès à des API, documents ou outils d'administration réseau.

Public concerné : Administrateurs réseau, ingénieurs sécurité, développeurs d'intégrations IA et responsables techniques

Prérequis : Comprendre les API web, l'authentification, la segmentation réseau et le principe du moindre privilège.

Objectifs opérationnels

- Cartographier les données, outils et privilèges d'un agent IA.
- Identifier les scénarios de fuite, d'injection et d'abus d'outils.
- Protéger les clés, secrets et comptes de service.
- Valider les entrées, sorties et actions sensibles.
- Mettre en place journalisation, surveillance et réponse à incident.

Compétences développées

- Construire un modèle de menace adapté à une intégration IA.
- Définir une architecture réseau segmentée et des droits minimaux.
- Concevoir une validation humaine pour les actions à impact.
- Tester les contrôles et documenter les risques résiduels.

Programme détaillé

Le rythme alterne apports ciblés, démonstrations, manipulations guidées et mise en situation sur un environnement de laboratoire isolé.

Jour 1 - Menaces et architecture de confiance

Matin - Surface d'attaque d'une solution IA

- Composants : modèle, interface, API, données, outils et connecteurs.
- Fuites de données, prompt injection et contenu non fiable.
- Abus de fonction, escalade de privilèges et actions non prévues.
- Méthode de cartographie des actifs, flux et niveaux de confiance.

Travaux pratiques : Construire le modèle de menace d'un assistant ayant accès à une documentation et à une API réseau.

Après-midi - Identités, secrets et segmentation

- Comptes de service, jetons courts et rotation des secrets.
- Coffre de secrets, séparation des environnements et moindre privilège.
- Segmentation réseau, filtrage sortant et contrôle des destinations.
- Traçabilité des appels d'outils et attribution des actions.

Travaux pratiques : Définir une matrice de droits et des flux réseau minimaux pour l'agent.

Jour 2 - Contrôles applicatifs, tests et réponse

Matin - Encadrer les entrées, sorties et actions

- Validation de schéma, filtrage et normalisation des paramètres.
- Liste d'outils autorisés et séparation lecture/écriture.
- Approbation humaine, simulation et limites transactionnelles.
- Traitement des erreurs et prévention des données sensibles dans les traces.

Travaux pratiques : Ajouter des contrôles à un agent vulnérable et vérifier leur efficacité.

Après-midi - Audit et gestion d'incident

- Jeux de tests adverses, critères d'acceptation et non-régression.
- Journalisation exploitable, alertes et surveillance des comportements.
- Désactivation d'urgence, révocation des secrets et investigation.
- Politique d'usage, responsabilités et dossier de sécurité.

Travaux pratiques : Auditer puis sécuriser un agent doté d'outils d'administration dans un laboratoire isolé.

Projet ou atelier de synthèse

Audit de sécurité et durcissement d'un agent IA capable d'interroger des outils d'administration réseau.

Méthodes et moyens pédagogiques

- Apports techniques courts suivis d'une application immédiate.
- Démonstrations commentées et analyse collective des choix réalisés.
- Travaux pratiques individuels ou en binôme dans un laboratoire isolé.
- Supports numériques, fichiers d'exemple, modèles de procédures et corrigés.
- Adaptation possible aux technologies et contraintes de l'organisation après cadrage.

Évaluation des acquis

- Questionnaire de positionnement en début de formation.
- Évaluation formative pendant les manipulations et corrections collectives.
- Mise en situation finale évaluée à l'aide de critères annoncés.
- Questionnaire de validation des connaissances et bilan individuel.
- Attestation de participation remise à l'issue de la formation.

Environnement technique requis

- Poste avec navigateur, éditeur de texte et outils d'appel d'API.
- Environnement de laboratoire isolé fourni par le formateur.
- Aucune connexion à une infrastructure de production pendant les exercices.

Accessibilité et adaptation

Le programme, le rythme et les modalités peuvent être adaptés aux besoins des participants et aux contraintes d'accessibilité identifiées lors du cadrage. Contactez Provence Cloud avant la session afin de préparer les aménagements nécessaires.

Livrables remis aux participants

- Support de formation au format numérique.
- Fichiers, configurations ou scripts utilisés pendant les travaux pratiques.
- Fiches de contrôle et bonnes pratiques applicables en entreprise.
- Corrigé ou solution de référence du projet de synthèse.

Contact

Organisme : Provence Cloud - ESN basée à Salon-de-Provence

E-mail : Contact@provencecloud.com

Téléphone : 06 64 47 14 64

Modalités : Formation intra-entreprise, en présentiel ou à distance, sur devis et après analyse du besoin.