CYBERSÉCURITÉ · NIVEAU 3

Formation sécurisation des systèmes IA : protégez vos LLM et agents

La formation sécurité IA « Sécurisation technique des systèmes IA » de Yuki Solutions est un parcours de 3 jours (21 heures) pour les équipes qui connectent des modèles de langage au système d'information. Architecture, identités, prévention des fuites, gardes-fous, injection de prompt et supervision : vos équipes apprennent à déployer sans ouvrir de brèche.

DURÉE	FORMAT	PUBLIC
3 jours (21 h)	INTRA ou INTER, présentiel ou distanciel	DSI, RSSI, développeurs, DevOps
PRÉREQUIS	**TARIF**	**SESSIONS**
Culture informatique générale	dès 2 200 € HT / personne	Sur demande

Objectifs pédagogiques

✓ **Concevoir une architecture IA défendable**
Identifier les composants (modèle, orchestrateur, RAG, outils, connecteurs) et placer les contrôles au bon endroit.

✓ **Prévenir les fuites de données**
Mettre en place des mesures DLP sur les entrées et les sorties d'un système IA, et cloisonner les données sensibles.

✓ **Poser des gardes-fous**
Filtrer, valider et limiter les actions d'un modèle, avec une validation humaine sur les opérations critiques.

✓ **Maîtriser les identités et les accès**
Gérer les identités non humaines (comptes de service, jetons, clés d'API) et appliquer le moindre privilège aux agents.

✓ **Contrer l'injection de prompt**
Reconnaître les injections directes et indirectes, et tester la robustesse d'un assistant avant sa mise en production.

✓ **Superviser et tracer**
Journaliser les échanges utiles, détecter les dérives et préparer les éléments attendus par la conformité.

Programme détaillé

Le programme officiel compte six modules. Nous proposons de consacrer une demi-journée à chacun, en alternant apports et exercices sur des environnements dédiés.

Demi-journée 1 — Architecture des systèmes IA

- Anatomie d'un système IA : modèle, prompt système, contexte, RAG, outils, agents
- Modes de déploiement : SaaS, API, cloud privé, on-premise, modèles open-weight
- Cartographie des flux de données et des points de confiance
- Atelier : modélisation de la menace sur un assistant interne type

Demi-journée 2 — Identités et contrôle d'accès (NHI)

- Identités non humaines : comptes de service, clés d'API, jetons OAuth
- Moindre privilège appliqué aux agents et à leurs outils
- Rotation, stockage et révocation des secrets
- Atelier : revue des droits d'un agent connecté à une messagerie et à une GED

Demi-journée 3 — Prévention des fuites de données (DLP)

- Ce qui fuit, et par où : prompts, fichiers joints, réponses, journaux, contexte RAG
- Classification et filtrage des données en entrée et en sortie
- Cloisonnement des index RAG selon les droits des utilisateurs
- Atelier : mise en place de règles de filtrage sur un flux de test

Demi-journée 4 — Firewall AI et gardes-fous

- Filtres en amont et en aval du modèle : principe, limites, faux positifs
- Validation des sorties avant toute action (requête, envoi, écriture)

- Validation humaine sur les opérations critiques
- Atelier : configuration de gardes-fous sur un agent outillé

Demi-journée 5 — Risques techniques : injection de prompt et au-delà

- Injection directe et indirecte (documents, pages web, e-mails piégés)
- Exfiltration par les outils, empoisonnement des sources, déni de service
- Tests offensifs : construire une campagne de red teaming simple
- Atelier : attaque puis durcissement d'un assistant de démonstration

Demi-journée 6 — Monitoring et introduction à la conformité

- Que journaliser, combien de temps, avec quelles précautions sur les données personnelles
- Détection d'abus et de dérives, alertes, réponse à incident
- Lien avec l'AI Act, l'ISO/IEC 42001 et NIS2
- Cas pratique final : plan de sécurisation de votre propre système, quiz de sortie

Chaque participant reçoit le support des trois jours, une checklist de sécurisation d'un système IA, une grille d'évaluation des risques et une attestation nominative.

Public et prérequis

La formation s'adresse aux **équipes techniques** qui conçoivent, intègrent ou exploitent des systèmes IA :

- DSI, architectes et responsables d'infrastructure ;
- RSSI, analystes SOC et équipes sécurité ;
- développeurs, DevOps et data engineers.

Prérequis : une culture informatique générale est recommandée (réseau, authentification, API). Il n'est pas nécessaire d'être spécialiste de l'IA.

Modalités

Public	DSI, RSSI, architectes, développeurs, DevOps
Prérequis	Culture informatique générale recommandée
Durée	3 jours (21 heures), consécutifs ou espacés
Format	INTRA ou INTER ; présentiel ou distanciel
Langue	Français ou anglais
Tarif	dès 2 200 € HT par personne
Modalités et délais d'accès	Sous 15 jours après accord de financement
Méthodes pédagogiques	Apports, démonstrations, ateliers d'attaque et de durcissement sur environnements dédiés
Évaluation	Quiz d'entrée et de sortie, cas pratique final
Sanction	Attestation nominative
Accessibilité	Référent handicap joignable à contact@yukisolutions.fr
Indicateurs de résultats	9,1/10 de satisfaction en fin de session · 100+ professionnels formés
Financement	Organisme certifié Qualiopi : prise en charge possible par votre OPCO ; CPF selon éligibilité

Inscription et devis — un appel de cadrage de 30 min, sans engagement, pour adapter le programme à vos cas : http://rel-ia.com/contact/ · contact@yukisolutions.fr