



Maintenance, supervision et administration des solutions IA

IA070

Durée: 3 jours

Public :

Administrateurs, ingénieurs DevOps, ingénieurs déployant des solutions IA en production

Objectifs :

Maintenir et superviser des solutions IA en production, configurer des outils de surveillance et d'alerting, gérer les cycles de vie des modèles, administrer les infrastructures IA, assurer la sécurité et la conformité des déploiements

Connaissances préalables nécessaires :

Expérience en administration systèmes, notions de base en réseaux et bases de données, fondamentaux de l'IA

Programme :

Concepts fondamentaux de la maintenance IA

Types de maintenance : corrective, préventive, prédictive
Cycle de vie des modèles d'IA et points de défaillance critiques
Architecture de surveillance : journaux, métriques, traces distribuées, métriques de performance (précision, F1, ...)

Atelier : Installation et configuration d'une pile de surveillance complète (Prometheus + Grafana + Alertmanager) pour surveiller un modèle de classification déployé

Surveillance et suivi opérationnel

Configuration avancée Prometheus : collecte de métriques, re-étiquetage et fédération
Tableaux de bord Grafana : métriques métier, infrastructure, modèles ML
Présentation de Zabbix et ElasticStack pour la supervision IA. Avantages, inconvénients.
Surveillance IA - suivi GPU, métriques modèles, alertes sur dégradation
Métriques IA spécifiques (précision, F1, latence d'inférence)
Métriques JVM critiques (heap, GC, threads)
Alertes : règles d'alerte, escalade, corrélation d'événements
Surveillance de la dérive des données : détection automatique, seuils adaptatifs

Atelier : Déploiement d'une solution complète de suivi opérationnel



Gestion des données et stockage

Stockage et archivage des données d'entraînement et d'inférence
Organisation des datasets : structure de répertoires, conventions de nommage
Pipelines de qualité des données : validation, nettoyage, enrichissement
Sauvegarde et restauration : stratégies, planification, tests de récupération
Surveillance basique de la qualité des données : checks automatiques, alertes simples

Atelier : Configuration d'un système de stockage et de sauvegarde pour données IA - organisation des jeux de données, automatisation des sauvegardes, vérifications d'intégrité

Maintenance et cycle de vie des modèles

Indicateurs de dégradation des modèles : baisse de performance, latence
Procédures de maintenance : planification, fenêtres de maintenance
Stratégies de mise à jour : remplacement de modèles, rollback
Automatisation des tâches de maintenance : scripts, cron jobs
Gestion des incidents : procédures d'escalade, communication

Atelier : Mise en place de procédures de maintenance standardisées - scripts de vérification santé, planification des mises à jour, procédures de rollback

Administration des infrastructures IA

Gestion des ressources de calcul : allocation CPU/GPU/TPU, élasticité
Orchestration Kubernetes : déploiement de workloads IA, gestion des ressources
Stockage distribué : MinIO, Ceph, HDFS pour jeux de données volumineux
Optimisation des performances : profilage, optimisation mémoire, parallélisation
Haute disponibilité : réplication, équilibrage de charge, disaster recovery

Atelier : Déploiement d'une architecture IA haute disponibilité sur Kubernetes - auto-scaling des pods d'inférence, stockage persistant, monitoring des ressources

Déploiement et gestion des versions

Gestion basique des versions de modèles : étiquetage, documentation
Procédures de déploiement : tests préalables, validation, mise en production
Environnements : séparation dev/staging/production, gestion des configurations
Stratégies de déploiement : déploiement progressif, tests en production
Procédures de retour en arrière : critères de déclenchement, méthodes de rollback

Atelier : Mise en place d'un processus de déploiement contrôlé - environnements séparés, procédures de validation, scripts de déploiement et rollback



— Phirio —

Sécurité et conformité des déploiements IA

Sécurisation des modèles : chiffrement au repos et en transit, signature
Authentification et autorisation : OAuth, RBAC, API keys, certificats
Audit et traçabilité : logging des accès, historique des modifications
Conformité RGPD : anonymisation, droit à l'oubli, consentement
Scan de vulnérabilités : analyse des dépendances, tests de pénétration

Atelier : Configuration d'un environnement IA sécurisé - authentification multi-facteurs, chiffrement des communications, audit trail complet

Outils d'administration essentiels

Outils de journalisation : configuration et rotation des journaux d'applications IA
Supervision système : suivi des ressources, alertes système
Outils de sauvegarde : scripts automatisés, vérification d'intégrité
Outils de diagnostic : analyse des performances, détection des goulots d'étranglement
Documentation et procédures : création de runbooks, guides d'exploitation

Atelier : Installation et configuration d'outils d'administration - centralisation des journaux, monitoring système, automatisation des sauvegardes

Projet final et industrialisation

Projet pratique : Administration complète d'une solution IA - installation, configuration monitoring, procédures de maintenance, documentation
Création de procédures opérationnelles : runbooks, guides de dépannage, check-lists
Planification de la maintenance : calendrier, procédures préventives
Documentation technique : architecture, configuration, procédures d'incident